Dynamics of 5th Grade Students Engineering Service Learning Projects

Ms. Jessica E S Swenson, Tufts Center for Engineering Education and Outreach
Dr. Merredith D Portsmore, Tufts University
Dynamics of 5th Grade Students’ Evaluation of Solutions for Engineering Service Learning Projects

Introduction

Young students are natural problem solvers and engineers and therefore we are interested in understanding the ways in which they tackle engineering problems. The engineering design process is a representation of the practices engineers engage in when solving design problems and is commonly included as part of K-12 engineering education activities. Service learning projects are used in undergraduate engineering education to motivate students and give them real world experience [1]. Following this style of curricula we piloted an engineering service learning curriculum in one 5th grade classroom. Students were asked to find a problem in their school and solve it using the engineering design process. We were interested in two things. First, what types of problems do students choose to pursue? Second, how do students go about evaluating their solution? In this paper, we will characterize the types of problems the students chose and how the type of problems were related to their approach to evaluating a prototype of their solution.

Literature Review

Engineering in the Classroom

Most problems students see while in school are well-defined ones. Well-defined problems are those in which the initial conditions and type of solution is well known [2]. As engineering is entering elementary school curricula, well-structured problems are common. One of the most widely used formalized curriculums in elementary school classrooms are the Engineering is Elementary (EIE) units [3]. These units combine learning in engineering, technical literature, and science. The literature section of the unit asks students to think critically about a story of a child in a foreign country with an engineering problem. The story introduces the type of engineering being explored and the problem the students will eventually solve. Students are then asked to analyze how the protagonist uses the engineering design process to solve the problem. The science section introduces students to use of scientific data to inform engineering design. Finally, the engineering section of each unit teaches the students about the engineering design process and asks them to construct a model out of found materials to solve the problem posed in the book. In the Engineering is Elementary unit “Lerato Cooks Up a Plan” [4] students are given a basic model of a solar oven made from a shoebox and aluminum foil. After undertaking a science investigation testing the thermal insulation properties of found materials, students are given those materials so they can design the insulation of their ovens. They then test them by measuring their internal temperature over time. Students are then asked to improve their designs. This type of engineering problem engages students in science investigation and engineering design practices. However, a number of the parts of the engineering design process are done for the students within EIE. The problem is defined for the student (build a solar oven) as are the requirements (the oven should be about the size of a shoebox and reach the maximum temperature possible) and the ways that they should evaluate their design. We are interested in understanding how students engage in more open-ended, ill-defined engineering problems that require them to create the problem definition, develop criteria to evaluate their problem, and create tests or evaluations to access if they have met those criteria.
Ill-defined problems are those more commonly found in professional engineering practice and have multiple unknowns, no clear problem-solving process and may require the integration of knowledge from multiple domains [2]. This mismatch from K-16 to professional practice make students ill-prepared to handle the problems they encounter in industry and the real world [2]. Over the past 15 years universities have used engineering service learning to give students more experience with ill-defined, open-ended problems. Service learning curriculums are designed to give students a more real-world experience by communicating and working with a client. The team works to define the problem, criteria, and methods for evaluating their solution. One example program is the Engineering Projects in Community Service (EPICS) program created by Purdue University [5, 6]. Each semester, students are paired with a community partner organization to assist it in solving a problem. Past projects include designing museum exhibits, helping communities encourage recycling, and aiding local non-profits in increasing their community presence through technology. A number of other universities have adapted this model as it is highly engaging and motivating for students and has been shown to attract underrepresented populations [7]. EPICS has been transposed for high school students [8] but to date there have not been efforts to bring it to lower grades.

With the goal of creating a motivating and engaging elementary school curriculum, we worked with a 5th grade teacher to transpose the ideas of university level service learning curriculums to create a new unit.

K-12 Students’ Evaluation of Design Artifacts

In more well-defined engineering design problems, tests for the final design are prescribed. However, the open-ended nature of service learning projects means that one of the tasks that students need to do is to develop ways of evaluating their final design.

To date, we know little about how K-12 students evaluate ill-defined problems. Most of what we know about how students evaluate their designs comes from studies about students learning scientific practices where the researchers or teachers aid the students in defining the problem and evaluations. In these units, students construct and evaluate models (replications or representation of existing devices or phenomenon) either physically [9, 10, 11, 12] in simulation software [13] or a combination of both [14]. A number of studies have found that giving students basic functioning models and then asking them to redesign them to accomplish a larger goal motivates students to evaluate and iterate [9, 12, 13].

In two more open-ended studies by Penner Lehrer and Schauble [10, 11], students were asked to build a model that works like their elbow in order to learn the scientific practices of model building and revision. In the first study, first and second grade students first built models that focused on the visual similarities to elbows rather than accurately functioning like them. Only after two exercises examining the functions of their own elbows did students use their own arm movements to evaluate the validity of their models. In the second study, third graders did not ignore function and “continually compared and evaluated their models with respect to a real elbow” (p. 438). These studies show young students initially have difficulty constructing criteria.
to evaluate the function of their models and instead focus on aesthetics. Yet, when given a model or goal to compare their model to, they’re able to successfully evaluate their models.

Kolodner’s Learning by Design™ units [9] focus on learning authentic scientific practices through design based challenges. Students conduct two types of evaluation: running investigations or experiments and testing their designs. Students conduct experiments to learn scientific concepts to aid them in creating better designs. Designing experiments are what Kolodner and her fellow researchers identify as a “ritualized activity.” Students and teachers learn the sequence of events and how to identify the variable that needs to be varied and how to keep the others constant. Students are supported throughout the unit with pages in their Design Diaries outlining how to construct an investigation and gather data. Once students extract the implications of their experimental results, they apply these “rules of thumb” to their designs. These designs are then tested and the quantitative and qualitative data analyzed in order to determine what further investigations need to be conducted. This study shows when students are given guidance and a set of procedures to follow to evaluate they are able to construct evaluations and, with help, extrapolate implications from the results. Students also are able to collect qualitative and quantitative results when evaluating design solutions.

Sadler, Coyle and Schwartz [12] also designed a series of engineering design challenge units based on engineering competitions at undergraduate institutions. Students in grades five through nine start with basic, barely functioning models and are tasked with redesigning and improving the model to make the best design according to the criteria established by the class. The design goal “must be easily recognized by students so that there is no ambiguity in feedback concerning the performance of their design” (p. 313) and the authors created tests that were “highly reliable” (p. 324) so students can consistently receive quantitative results to encourage iterations. The authors also note students had difficulty varying only one variable at a time and did not keep records of trials over time. Data about how students perform these tests is not explicitly given, but it reports that when given straightforward tests that produce quantitative results students can easily interpret the results for further iteration.

Schauble [13] tasked fifth and sixth grade students to find out which features did and did not make a car travel faster in a micro-world software. The students were shown how to design an experiment and predict the outcome and given eight weeks to use the software program. None of the children used a systemic approach nor tested all the variables, and many constructed the same designs over and over. Despite the fact that students were able to discover which variables were causal and which were not, they had a difficult time creating theories from their experimental results or interpreted their results falsely. This study informs us that students have difficulty making evidence-based theories from their design evaluations. Therefore, if evaluating an engineering solution, students might have difficulty extrapolating from the results the issues with their design and what needs to be changed. The study also reports students did not run systematic experiments and did not use their logbooks to record their data. When evaluating engineering solutions, we should not expect students to test systematically or record data.

Schwarz and White’s [14] study focuses on modeling in the context of force and motion. Students were given predictive questions and performed real-world experiments to develop theories and build computer models. After students were taught about the importance of evaluation to
improve models and given five common criteria used to evaluate models, Schwarz and White found students “had a strong understanding and use of model evaluation criteria” (p. 185) throughout the remainder of the unit. In interviews after the unit, students discussed how they used the criteria introduced during the curriculum to evaluate their models using data and experiments. This study shows students are able to use criteria and understand its value when designing and evaluating models.

Previous studies focused on design tasks in a science context either gave students criteria or developed them in a teacher-led discussion. The research that has been done is not explicit on how students conduct these tests and in what ways teachers support them when doing so but inform us that evaluation is understood when students compare models, are given criteria, or given simple tests with quantitative results. These units all demonstrate students have the ability to iterate or redesign based on testing and feedback when units are highly structured or scaffolded by teachers leading all students in a similar design task. We want to look to explore how students develop evaluations for their solutions on their own when given an open-ended, ill-defined problem.

**Evaluation in Engineering**

There are many places within the engineering design process where engineers and design teams evaluate. Before evaluating, it is necessary to have a set of criteria that includes requirements, specifications, and constraints of the design. Criteria defines what the solution should do and how it should do it. Examples of criteria include size, cost, functions, and needs of the user. Engineers evaluate prototypes to measure if their solution has solved the problem and how well it has met the criteria. Depending on the product, a combination of user testing and testing for reliability and performance are used to evaluate the product. User testing focuses on the customer’s needs and wants and how they react to the product through observations and interviews. In testing for reliability and performance, engineers determine test the product through experiments and data analysis. This could be performed within the team, with real users in a focus group, or using testing equipment to take measurements (e.g. a car crashing into a wall). For the purposes of this analysis, we looked for instances in which students evaluated their prototypes or discussed ways to evaluate their prototypes. We examined the type of criteria students wished to meet and the results students obtained to determine the beginnings of the different types of engineering evaluations students used to evaluate their prototypes.

**Population**

Our engineering service learning unit took place in Mr. Walsh’s 5th grade classroom in an urban suburb of a major Northeastern city as a part of the Student Teacher Outreach Mentorship Program (STOMP). Mr. Walsh attended professional development as a part of the W-STOMP program, a one-year project that focused on women and girls, and chose to teach service learning, a new curriculum for the STOMP program. The classroom consisted of 20 students: 12 boys and eight girls. As part of the STOMP program, these lessons were given once a week for an hour, and co-taught with the teacher by two university students, or STOMP fellows. Mr. Walsh directed the class and the STOMP fellows served as the engineering and
technical experts. The researcher (first author) also served as a STOMP fellow while taking data and aided the teacher with curriculum direction.

**Engineering Service Learning Curriculum**

The engineering service learning unit took place over nine weeks as nine one-hour sessions. The first week the students watched a video of professional engineers tackling an engineering design problem \(^{[21]}\) and then walked around their school to pinpoint problems. These problems were then collected on a classroom whiteboard and each student chose the problem that was most meaningful to them to write on a Post-it. The following week the teacher randomly selected four students to form a group, forming five groups total. The groups then brainstormed solutions to each member’s selected problem and were asked to decide which problem they wanted to solve using engineering. The next two days were dedicated to making a stop-motion action movie using SAM software \(^{[23]}\) to explain their problem and proposed their solution to the class. Weeks seven and eight were spent designing, building, and testing. The materials students used to build their prototypes consisted of whatever was available in the classroom and supplies obtained by request from the STOMP fellows. The last day of the unit was dedicated to reflection about the process.

**Data Collection**

The primary method of data collection was video of in-class student group work and larger classroom discussions. Pictures of student artifacts were also collected to document the stages of their prototypes.

**Data Analysis**

The data was analyzed by transcribing selected video clips in which students were actively engaging in evaluation or discussing how to go about evaluating their prototypes. Pictures of student work were examined along with the video clips. A case study approach was used to examine and understand each group’s design and evaluation process.

**Characterization of Problems**

The students in Mr. Walsh’s class found five unique problems and developed solutions to solve them. The complete list of groups, problems, and solutions can be found in Table 1.
As we reviewed the students’ problems and solutions, a spectrum of types of problems has emerged. A number of the problems students chose required modifying the behavior of their fellow schoolmates—whether it be encouraging them to flush toilets in the bathroom or stay to the right side when going up and down the stairs. We categorize these problems as human-behavior-oriented involving multiple end-users because they attempt to modify the behavior of fellow students. The other end of the spectrum are more device-based problems such as the Guminator, the tool the Gum Group designed to remove gum from the bottom of desks.

### Table 1: Five groups and their problems

<table>
<thead>
<tr>
<th>Group</th>
<th>Problem</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gum Group</td>
<td>Students stick gum to the bottom of desks.</td>
<td>The Guminator, comprised of two forks to scrape off gum, a small plastic container to catch the gum, a knife covered in velcro remove remainder, and a spray bottle of Goo Gone to help loosen stiff gum.</td>
</tr>
<tr>
<td>Locker Group</td>
<td>Backpack straps from top lockers hang outside the locker get trapped in the locker below it.</td>
<td>Rare earth magnets and bright orange tape on the end of backpack straps to hold them out of the way.</td>
</tr>
<tr>
<td>Stair Group</td>
<td>Multiple classrooms on the same stairs get mixed up and people can get hurt.</td>
<td>Arrows and dashed line taped on stairs to direct lines of students up the right side of the stairs and down the left.</td>
</tr>
<tr>
<td>Toilet Flushing Group</td>
<td>Students forget to flush the toilets in the bathrooms.</td>
<td>Signs saying “Please Flush” posted in the bathroom and individual stalls to encourage fellow students to flush.</td>
</tr>
<tr>
<td>Trash Group</td>
<td>Trash does not make it into the trash cans.</td>
<td>Decorated trash cans and recycling bins to turn discarding items into the bins a game.</td>
</tr>
</tbody>
</table>

For this paper, we will present instances in which we see the beginnings of students engaging in an evaluation like professional engineers would—specifically, occurrences in which students discussed how they might evaluate their prototype, what criteria their prototype needed to meet, observations they made in the middle of a test, or results they collected from a test. The three groups presented were chosen to show a range of students’ ideas of evaluation of different problem types.

### Data
The Gum Group

The Gum Group was comprised of Matt, Mark, Jonathon, and Kendra. They decided to tackle the age-old problem of chewing gum stuck to the bottom of desks. No criteria were explicitly made before prototype construction began. This group started construction of their prototype at the beginning of day two while making stop-motion animation movies. It was comprised of two plastic forks with a plastic bag taped to it (see Figure 2).

![Figure 2: Guminator Iteration 1](image)

Shortly after making it, Matt and Jonathon decided to try it out. While testing, they discovered their prototype had some flaws.

Matt: Is it working?
Mark: Is it working? Yep it works.
Matt: Now we just gotta make it so falls into the bag easier.
Jonathon: Yeah, I could I could do that.

By testing their prototype, they found the forks removed the gum but the way the plastic bag was taped onto the device did not allow the gum to drop into the bag as they intended. Matt’s comment about the bag indicates he has created an implicit criterion that the gum should not be touched by the user and wants it to fall off the fork into some kind of container. He is thinking about the user’s needs and not only the problem of removing gum. From the feedback from this test they modified the prototype by replacing the bag that was giving them trouble with a more structured plastic container covered with a piece of cardboard into which they cut a slot that the gum could fall through (shown in Figure 3). They also changed the attachment of the two forks so they would stay together but slide past each other.
Their second test revealed that the new attachment sleeve, which was made out of paper, was too stiff and did not allow the forks to slip past each other. They modified the prototype by using a slipperier material–duct tape–that would allow the forks to slide (shown in Figure 4). This modification shows the group is not only concerned about the function of the prototype (taking gum off of desks) but the ease of use for the user (the user doesn’t touch the gum and can easily move the forks).

While performing a test with the prototype shown in Figure 4, a researcher started probing the team about their plans for making other models.

Researcher: So are you going to…are you going to make any other models then since you have more time next week?
Matt: Uuuuuuhhh.
Researcher: And then test them all against each other?
Matt: Actually, the first thing that we did was a…ummm…we tried different liquids to get it [the gum] fall off.
Researcher: Ahah.
Matt: That didn't work.
Researcher: So...
Matt: Maybe if we coated the fork with a liquid it would work better.

Going off an idea they had about the hardness of the gum and the difficulty that brings in removing it, the team decided to try different liquids to loosen the gum to see if they would help remove better. Matt reported that their trials did not produce successful results. Yet, this still prompted more brainstorming about how to use liquids in a different way to aid in the construction of their prototype.

The final iteration of the Gumminator prototype is shown in Figure 5. Besides the two forks and plastic container, the team also added a spray bottle of a cleaning product, Goo Gone, and a plastic knife with the blade part covered in the hook side of Velcro. The spray bottle of Goo Gone is meant to loosen the gum before it is picked up and the knife is meant to be a scrubber to scrub off the gum residue left behind once the gum is removed. After the team showed the prototype to an STOMP fellow, he asked to see how it works. They began to show how it works to the STOMP fellow and were picking out a piece of gum to remove. The following exchange occurred between two group members about which piece of gum they would choose.

Kendra: Can I try it? Can I scrape it?
Matt: That one doesn’t work Kendra.
STOMP fellow: Yeah I can see it.
Matt: It’s too flat.
Kendra: But we need to replace all kinds of gum. Flat or not.

Matt, having tried many pieces of gum with the device, doesn’t believe the device will be able to remove the flat piece of gum. Kendra states the device should be able to remove all pieces of gum, arguing that it should be a requirement of the device.

Later in the demonstration, both students mention the need for different materials to improve the effectiveness of their prototype.
Matt: That’s going to be hard to get off.
STOMP fellow: It’s going to be hard, but you should be able to take it all off, right?
Matt: Well, yeah. See what she’s doing. It’s harder [the gum].
Matt: I wish we had silverware.

Matt, still frustrated with how the stiffness of the gum made it more difficult to get off, wished their group had silverware, a much harder material than plastic forks they were using. Later in the demonstration, Kendra made a similar comment. After a number of modifications to improve the usability of the device, the students are now concerned with the function of the device and the plastic fork’s ability to scrape off gum.

The students were unable to remove the flat piece of gum and move onto another one. While attempting to remove the piece of gum, the top fork became unhinged from the rest of the device. This error allows the students to discover something new about their prototype.

Matt: That wasn’t supposed to happen.
Kendra: No, it’s going to work.
Matt: That wasn’t supposed to happen.
Matt: That never happened.
STOMP fellow: That never happened?
Matt: Nope.
STOMP fellow: Okay.
Kendra: Aren’t you going to spray it first?
Matt: Actually that’s a better idea.
STOMP fellow: That might be a better idea.
Matt: And then you…
STOMP fellow: Because now it gets too bulky.
Matt: Slip it back on.

As Matt tested his prototype the top fork became unhinged from the rest of the prototype. At first, Matt tried to cover for the prototype’s error but then realized removing the bulkiness from the prototype made it easier to manipulate and thus remove the gum. By trying out and using their prototype, the students discover how to use the device in different and better ways.

We claim that we see the Gum Group engaged in a few kinds of evaluation: evaluation of materials, evaluation of the function, and evaluation of the tool arrangement. We see that their evaluation results prompted the construction of four distinct iterations of their design: two forks and a plastic bag, two forks with a plastic container and paper sleeve, two forks with a plastic container and duct tape sleeve, and lastly the addition of a knife with a rough surface and a spray bottle of Goo Gone. Evaluation also caused the team to verbalize implicit criteria and construct criteria as they are evaluating. From the transcript, the team started with two criteria: that their device works and the user does not have to touch the gum. From evaluation, they added the device works on all types of gum and removes the entire piece of gum to that list. The activity of evaluating prompted the implicit criteria to be verbalized and based on the device’s performance, new criteria to be created.
The Backpack Group

The Backpack Group was comprised of Hunter, Ross, Jacob and Mackenzie. The problem they decided to address concerned how the straps of a backpack placed in a top locker hang out over the bottom locker. Because of this, the straps can occasionally get caught in the bottom locker because the student owning that locker does not realize they are hanging down. The students set out to create some way that the hanging straps would be noticed. Their prototype consisted of bright orange tape wrapped around the bottom of each strap and a magnet; the idea behind the former was to get the student’s attention; the latter, to attach the magnets to the outside of the top locker, thus getting the straps out of the way. A close-up of the prototype and how it was used are shown in the figures below.

Figures 6 and 7: The Backpack Group’s Prototype

On the first day of prototyping, a researcher asked the group how they thought they were going to test their prototype.

Researcher: So is there a way you're going to test this so in a…in a bigger sense? Like now just this one, but like…
Ross: I was thinking, like, we could take maybe ten people from different classes or just from our class, and we could…um…take a, like, have them put in somebody else's backpack so that they—we would take people who hadn't attached the magnets to or tape yet.
Researcher: So maybe like....
Ross: So they wouldn't…so they wouldn't know about it so then we would ask them to put to…um…we would take two people and then one person would…that would be the person putting in the backpack with the magnets and the tape on into the locker and then the other person would…you'd see how many times a pair, and the bottom person noticed.
Ross suggested testing the backpack with students, who are the customers or clients this team was working for, who know nothing about the prototype, and seeing how many of them notice the straps. Essentially, he was proposing a user test with quantitative data to verify their prototype works. Ross’s proposed test never occurred.

The next session after the students’ modified their prototype by increasing the strength of the magnets, the team and a STOMP fellow walked through how the prototype could be used.

    STOMP fellow: And imagine alright now you're going to close the locker. Now somebody else is going to come in— (Figure 8)
    Jacob: Wait, wait, wait. What if this happens...what if this happens? (Figure 9)
    Hunter: Yeah the—
    STOMP fellow: With both lockers open at the same time?
    Hunter: But you have to, like, force that though. Like, if it's if it's say it's right here just say...right here you go like this and this person opens it. Wouldn't they notice this bright orange tape? (Figure 10)
    Jacob: Yeah, no, I know they would.

While the team was describing the use of their prototype, Jacob stopped the group and suggested a scenario to the team: What if the person in the bottom locker doesn’t notice the strap and shuts their door on it? Hunter stood up for the prototype, stating that would only happen if it were forced and asked the group to back him up that the person who owned the bottom locker would see the tape.

We claim the Backpack Group engaged in mechanical evaluation but did not evaluate the human-behavior aspect of their solution. The mechanical testing prompted them to iterate on their materials and add extra features to their design. While the group never tested their prototype with users outside of the group, one student proposed a method of testing the human-behavior aspect of their solution like user testing with quantitative data that professional engineers would engage in.

*The Stair Group*

The Stair Group was comprised of Steven, Mike, Derek and Bella. Their problem involved managing traffic on the staircases throughout the school. From the students’ description, there
were collisions in the stairwell when multiple classes were on the same staircase. The students’ first prototype (shown in Figure 11) was a sign with the words “Please wait till other line comes down! Then you may go. Thank you! (Please stay on the right side)” with a stoplight-like drawing in which the bottom red circle red signified stop. After they hung up the sign, a STOMP fellow questioned them about how they were going to test it.

![Figure 11: The Stair Group’s first prototype](image)

STOMP fellow: How are you guys going to test if these signs work?
Steven: Um we're going to ah...we're going to see...we're going to see—when the teachers go up we're going to see...we're going to ask them, How do you like our new kind of—
STOMP Fellow: Signs? Okay.
Steven: Like a miniature interview.

Steven suggested asking the teachers once they have gone up and down the stairs and seen the signs how they like the signs and conduct a miniature interview. He might have also been suggesting that they would observe them using the sign on the stairs. It seems that Steven had identified teachers as the primary client because he proposed interviewing them and not any of the students. Another group member, Mike then volunteered his answer to the STOMP fellow’s question.

Mike: Say, like, aaaaahh, let's see, like the way...the way, like, if it works that's a sign that it's doing good. So if it keeps going for like at least this school year or something and then if it keeps working we can use it until next year.
Mike: And then if it...if it doesn't work so then, um, we...
STOMP fellow: You can figure it out from there.
Mike: Yeah we could figure it out so if we if we keep doing um [this] and like so if we keep doing [this] it probably would work if we like mmmm...
Mike: Like if they...pretend like you're coming down and like there's ..... I can't really explain that much because ah it's like um...
Mike: Because, um, since since...this sign is...might still be there they're going to be tired of listening to it and then like the the...
STOMP fellow: Yeah it will get ineffective because people will stop, absolutely.
Mike: Mm-hmm and then and then and then no one's not going to listen to that sign that's here and that sign up there.
STOMP fellow: Right.
Mike: And if we keep doing...if we put it all around those staircases there and over here, it's not going...they're just going be tired of .... just, like, doing it.

Instead of thinking about how to test the sign’s current effectiveness, Mike hypothesized the sign would work and though about the long-term use of the sign. At first he seemed unable to articulate his original idea and then stated if the sign is up for a long time it would eventually be ignored and become ineffective. He hypothesized if the sign worked and they duplicated it and they were all over the school, people would get used to seeing them everywhere and stop paying attention to them, and they would no longer solve the problem.

Discussion

Our analysis shows students engaging in evaluation of their ideas and prototypes. The problem type and students’ evaluation of their solution appear to be related as evidenced by the number of evaluations and redesign each group performed. Groups with device-based problems were able to test and evaluate immediately after designing, which provided them with instant feedback that prompted redesign. Groups that had more human-behavior problems required a large amount of people to interact with their designs during evaluation. Without instant feedback, these groups did not iterate as frequently.

The Gum Group was observed to repeatedly evaluate their device and make iterative changes to their design. We claim that these students’ evaluations were similar to that of professional engineers’ testing for performance, function, and robustness. They were able to pinpoint specific features of the device that needed to be changed, specifically, the plastic bag in their first test, the fork attachment sleeve in their second test, and the brittle plastic forks in the final test. These features were then modified on the prototype, improving the device each time. Their process differed from professional engineers in that a number of their criteria were developed from their evaluations instead of explicitly stated before their design process. Testing their device on pieces of gum allowed the gum group to better understand their problem and therefore more specifically define their problem, creating more criteria and solutions to meet those criteria. For instance, the students noticed that once a piece of gum was removed, a residue remained on the surface. They created the scrubber as part of their device to remove this extra gum. While never explicitly stating usability requirements before testing, comments made during evaluation indicated the team, like professional engineers, had ideas about user preferences. For example, they recognized the user would not want to touch the piece of gum. Like professional engineers, the team also used research in their design by engaging in testing to determine the effect liquids had on the hardness of gum. The results from all these tests and evaluations caused the group to iterate multiple times to eventually create a very unique device that solved the problems they identified and met the criteria that emerged as they were engaged in prototyping.

The Backpack Group created a device-based solution that also involved dealing with a human-behavior across multiple users. The students were able to conduct a number of tests on the mechanics of the design that allowed them to identify issues, such as the need for greater magnet strength, which caused them to iterate from their first prototype. Early in the group’s design process, when prompted by the researcher, Ross, a member of the group, suggested testing with
students or users who knew nothing about the prototype. This suggestion was along the lines of the concept testing \[^{18}\] that professional engineers perform with potential users to identify any shortcomings of the product. Unfortunately, any human-behavior evaluation of this kind never occurred. In their final mechanical evaluation, another group member, Jacob, posed a hypothetical situation to his group questioning the effectiveness of the prototype. Hunter defended the prototype, looking for confirmation from his group that it was unlikely to happen. In this group, physical evaluation using the prototype informed design decisions. One group member had an understanding that verification of its functionality by users unfamiliar with the prototype would give them evidence that their solution solved their problem.

The Stair Group tackled a problem that was human-behavior oriented with multiple end-users, a factor which gave them more difficulty in determining how to test their prototypes. The group had a number of ideas on how to determine if their solution worked. Steven suggested mini-interviews, a concept testing method also commonly used in professional engineering \[^{18}\]. Mike hypothesized that the sign could become ineffective but did not suggest any way of testing that. The group decided to use personal observations made throughout the week between STOMP classes, which spurred them to create an entirely different final prototype.

These three groups show that students, when given an open-ended, ill-defined problem, exhibit the beginnings of the ability to develop criteria and evaluate their ideas and prototypes. The students’ evaluation of their prototypes seemed to be related to the type of problem and solution they pursued. Groups with more device-based prototypes, such as the Gum Group and the Backpack Group, were able to try out and test within the group and therefore able to conduct more evaluations and gain more feedback than groups that took on problems that were human-behavior oriented multiple end-user groups, such as the Stair Group. These groups were challenged with understanding what kind of evidence they needed and how to obtain that quantity of data. All three groups showed the beginnings of performing evaluations like professional engineers.

**Implications**

In both K-12 and undergraduate engineering curricula, students learn to solve well-structured problems \[^{2, 24}\] and less often tackle the open-ended, ill-defined problems that are similar to those found in professional engineering environments. This pilot provides preliminary evidence that with minimal scaffolding 5\(^{th}\) grade students are able to tackle ill-defined, open-ended engineering problems and show productive beginnings for developing criteria and evaluating their solutions. Most existing design-based curricula for elementary-age students does not give them opportunities to engage in the problem scoping, criteria definition and evaluation that this curricula provided. Moreover, the existing research did not give any insight into whether elementary students could participate meaningfully in all the practices needed to design a solution type of problem. While more work needs to be done to understand the dynamics of students participation in this type of problem and to articulate best practices for supporting and engaging students in ill-defined, open-ended engineering problems, this work gives initial support for engaging elementary students in open-ended, ill-defined engineering design problems.
One of the particular dynamics that should be explored is the relationship between the type of problems and the strategies for testing and evaluation. In this study, we saw that students were easily able to engage in evaluating the functionality of device-type prototypes as evidenced by the Guminator Group repeated testing of forks and liquids for removing gum. However, with the Backpack Group and the Stair Group, we saw students suggest ways of evaluating their design (have a group of students use the new backpack straps or observe groups of students on the stairs) but not actually follow through with the planned evaluation. The suggested evaluations for these products had components of psychology or human factors research that required the coordination of multiple test subjects and larger scale data collection. This difference warrants further exploration of how students engage in different types of problems and how they can be supported in evaluating more human-centered. A deeper understanding of this dynamic would have implications on how students select problem, scaffolding during brainstorming, and support for different kinds of evaluation and data collection techniques.

References


