

The Marshmallow Metaphor—Iterative Design Tailored to 6th Graders

David Reeping, Dr. Kenneth Reid
Ohio Northern University, Ada, OH 45810
Email: d-reeping@onu.edu

Abstract:

In an effort to integrate engineering concepts into a middle school environment, 6th graders were tasked with a design lab popularly known as “the marshmallow challenge.” Instead of beginning the challenge right away, a brief 10 minute lesson was included to give perspective to the lab. In all, the objectives of this workshop were the following:

Student Centered

- Introduce topics of engineering using a joint lecture and hands-on approach
- Generate interest in engineering
- Encourage creative thought and problem solving when presented with constraints
- Help students understand the role of failure in design and the value of prototyping

Research Centered

- Observe the teamwork dynamics and processes of each individual group
- Identify trends specific to teams and similarities in designs across the three sections.

During the lab, each team was observed to some degree for certain qualities and actions. A general synthesis of the section-wide performance reveals a trend where later classes produced taller and overall better structures than earlier sections. Notes taken during the competition and results of the competitions are presented and compared to each other.

The goal of this paper is to detail the lab for any educator looking to implement an engineering activity and supply learning objectives that leave open ended research capabilities—in addition to the presentation of notes gathered from the lab held. Those who favor the integration of engineering in the K-12 curriculum may also find this paper of interest.

Introduction and Background:

In design, iteration is a common centerpiece that aids in producing the best possible result for an experiment. A consistent cycle of designing, prototyping, and testing is the logical process taken that has birthed the buildings, products, and vehicles of today. Few designs have an initial success so great that they could be deemed perfect, so certain steps needed to be taken for those preliminary designs to sprout their wings.

For example, the design process of the airplane can be traced back to a large scrapyard of failed prototypes that claimed their fair share of lives. Although one eye-catching prototype was expected to catch a tailwind and soar through the air, the clunky design ravaged the pilot with

brutal gusts and forced him into an uncorrectable downward spiral. Today's testing methods have ideally removed almost all possibility of unintended causalities, so learning from failure has become much more of a spectator sport. While an unfortunate reality, failure will always remain an important component of design. Of course, it is not the faulty design to be thankful for, it is the learning opportunity that failure presents.

Scaling down a project to a design lab for students can express this lesson: automotive engineers did not start with a roaring sports car just like mechanical engineers tinkering with firearms did not begin with chattering machine guns. At some point, something went wrong, but a new idea rose from the remaining bits and pieces to produce today's design.

In the K-12 environment, a lab that features the elements of iterative design and the inclusion of multiple engineering concepts is feasible, but is often dependent on outreach from a local university when a school lacks an engineering program. As of now, some efforts have been made to slip engineering into the curriculum through the back door. In fact, notable additions to curricula like "Engineering is Elementary" and "Project Lead the Way" have made impacts across numerous states ¹. Another subtle way of introducing important concepts into the K-12 setting is through Model-Eliciting Activities (MEAs). These activities are math problems that are applicable to real world scenarios and involve students collaborating to create an appropriate model that describes the situation ². Since engineers use modeling constantly during the design process, the use of MEAs to educate K-12 students in the field would be a fairly easy addition. "The Marshmallow Challenge" provides an opportunity for student to gain hands on experience with engineering and also yields interesting results that can provide a framework for future research in STEM education.

Engineering Crash Course—Giving Perspective to the Lab

Before beginning the lab, all three sections were opened with a short lecture on basic concepts of engineering applicable to the project. The lesson revolved around the following ideas:

- Design Qualities
- Criteria and Constraints
- Role of Failure

To illustrate design qualities, the students in the 10:10 and 12:45 sections were posed with the following scenario: *suppose we are going to design a phone to compete with the major companies, what qualities should our phone possess to be competitive?* The answers were the expected array of eye catching and standard features that current day phones advertise: touch screen, voice commands, small size, and apps. The purpose of this exercise was to help students distinguish between good and bad criteria. Any criteria that could have been assigned a numerical value or upper and lower limits were circled and the students were asked questions concerning the ambiguity of their design qualities. The students learned that good criteria should be quantifiable and these qualities should not be entirely subjective. In the 2:45 class, the following scenario was used instead: *suppose we are going to design a car to compete with the major auto manufacturers, what qualities should our car have to be competitive?* This question was significantly less effective in producing answers to make the "quantifiable speech."

The students tended to name parts of the car rather than design features like “a powerful engine” where a specific horsepower or engine could have been assigned.

To complement criteria, the students learned about the significance of constraints. While it was difficult to coax a definition out of them in their own words, the 6th graders seemed receptive to this definition: *a constraint is a limitation on a design.*

Finally, a short discussion on the role of failure in design was given. The message of this topic can be summarized with the words of Henry Petroski, “tragic accidents and failures always bring renewed focus on the nature and reliability of engineering structures and technological systems of all kinds”³. While the students were experiencing failure on a smaller scale, the principle remained the same. So understandably, every design we have today has experienced failure at some point in its development. After the introduction, the class transitioned into the lab.

The Marshmallow Challenge

The 6th graders were tasked to do a lab that was originally a design challenge introduced by Peter Skillman and popularized by Tom Wujec at a TED conference in 2010⁴. Today, it is commonly used as a team building exercise in major businesses and universities. Yet, both the engineering applications and team dynamic components were examined for this lab activity.

First, the goal of “the marshmallow challenge” is to build the tallest free standing structure using the following items in a premade kit in under 18 minutes:

- 20 sticks of spaghetti
- 1 yard of tape
- 1 foot of string (originally 1 yard in Wujec’s instructions)
- 1 marshmallow

However, there is a clever twist that elevates this challenge to a different level. For the structure to qualify, *the marshmallow must be at the highest point.* Also, the following rules apply:

- Students can use as much or as little of the materials as they wish.
- Items in the kit can be manipulated in any way (i.e. breaking spaghetti sticks in half) except the marshmallow must remain whole.

The unusual aspect surrounding this design lab is that kindergarteners (and engineers, of course) are typically the best performers. Business students, on the other hand, perform the worst—below average. The reason for this is fairly simple; kindergarteners do not fight for power in their groups and, most importantly, they create prototypes for their design. Wujec explains, “business students are trained to find the single right plan...and then they execute on it,” but therein lies the problem². These students wait until the last possible opportunity to test the marshmallow and end up hoping their one design will hold the weight.

Another interesting twist is the involvement of an incentive—a reward of sorts. In the first section, there was no prize for winning beyond bragging privileges. However, the second class was treated to a different scenario. Before starting the challenge, the section was presented with a

plump bag of candy with the assuring phrase, “oh, and winner gets this bag of candy.” Wujec’s experience revealed that offering a prize drastically decreases the number of successful teams⁴. The 12:45 section was the only class that was offered the reward officially, but the third class may have been aware or inferred there was a prize. Note that this lab was conducted during a normal school day, and students were not monitored outside of the classroom, since this was not an intention of the lab.

Results:

Section I of the Appendix holds the results of the competition; the heights and winners are listed and divided by class. Along with recording the heights, a picture was taken of each structure (for the complete collection of structures broken down by class, please refer to Section III of the Appendix). In the case of the first class, the students had already begun to disassemble their structures, so the pictures may look misleading and sloppy due to photographer error. The results of the design lab have been broken down into two categories for the sake of organization: designs and observations.

Designs

Eventually, the final designs across the three sections began to exhibit some similarities. The idea of the “best” design will be discussed further in “Observations,” but the trend of designs could be easily seen through the graphic flowchart depicted below (Figure 1).



10:10 ----- 12:45 ----- 2:45
Figure 1: Progression of Similar Designs

As expected, the last class did the best at demonstrating the value of self-reinforcing geometric shapes like triangles. However, earlier classes seemed to understand this concept to a point, but fell short in delivery. The students were not taught this explicitly during the opening lesson, but were most likely clued in through the latent curriculum that an iterative design project creates. By sharing experiences with the later sections, a sort of “communal design” project is created—where the group of students working separately eventually converges to one design pattern as time goes by. When comparing the average height of the “communal design” structures to the average height of those that do not follow the pattern, the average for “communal” structures is 18.65 inches while the “non-communal” average is 13.83 inches—a spread of almost 5 inches (actual value is 4.82 in.). Group B3’s design was disqualified for being a supported structure, but was averaged into the “non-communal” values since it was standing and measured anyway—unlike A2, B1, and B2. Perhaps the structures would have grown in height and become more similar if more classes were tested.

Observations

Instead of sitting back and watching 18 minutes slowly dwindle down on the stopwatch, each team was observed and their performance was noted in addition to recording the height of their structure. To do so effectively, there were three specific questions that could be answered with a simple checkmark or number:

- How are students utilizing their supplies? How much or how little are they using?
- Are students sketching out designs? When are they doing this work?
- When are the students involving the marshmallow in their design? Early on, or just before time is called?

These questions produced an interesting trend data-wise when the answers are compared across the three sections. Refer to the Appendix Section II for the complete table.

How are students utilizing their supplies? How much or how little are they using?

10:10 Class

Unusually, the first group of students was frugal in the use of supplies and was practical in design approach. Group A1 used the least supplies and had the most visibly and theoretically stable design. Group A2 and A3 had similar approaches in management of supplies, taping and tying spaghetti together and seeing what would stand upright. Interestingly, A2 had two simultaneous designs being constructed. While this may sound appealing, splitting up a team with limited resources to build different designs under a time limit was not a wise decision. Perhaps if teams had more supplies, time, and members, tasking sub-teams to separate duties would have been effective.

12:45 Class

The second section used more supplies than the first, but followed the same theory of bundling spaghetti in Groups B1 and B3. Group B4 was the only team in all three sections to not use string, but had the tallest structure in the class. Both B1 and B2 used everything in their kit, and were disqualified since their structures could not stand upright.

2:45 Class

The last class completely emptied their kits and incorporated all of the supplies into their designs. Also, this was the first class that asked for more supplies during the challenge. Of all the sections, this class used their kit in the most effective way—except for Group C1, which seemed to be having some trouble throughout the activity.

Are students sketching out designs? When are they doing this work?

Students were encouraged to do pre-design work and sketch possible structures. Yet, “encouragement” did not seem to motivate the participants in the later section. In terms of pre-design, a visual representation of the trend is most likely the best route to explain. Figure 2 shows a graph based on three answers from the students: yes (before beginning to build), yes (after beginning to build), and no.

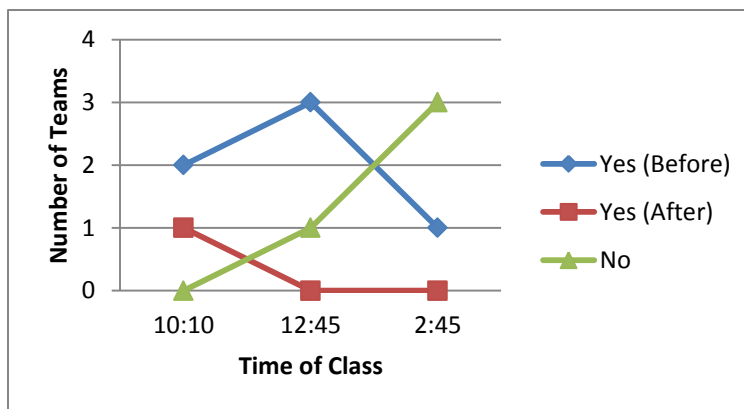


Figure 2: Trend in Sketching Designs

As the day progressed, pre-lab assumptions predicted that pre-design work would increase. However, the students proved that assumption wrong when the third class almost completely abandoned any notion of modeling. At first, it was shocking that even after discussing design qualities before beginning the challenge, the students still did little to prepare. Yet, the third class’s designs were vast improvements over the previous two sections. By comparing the success of teams versus pre-build work in Figure 3 below, the unusual trend can be seen.

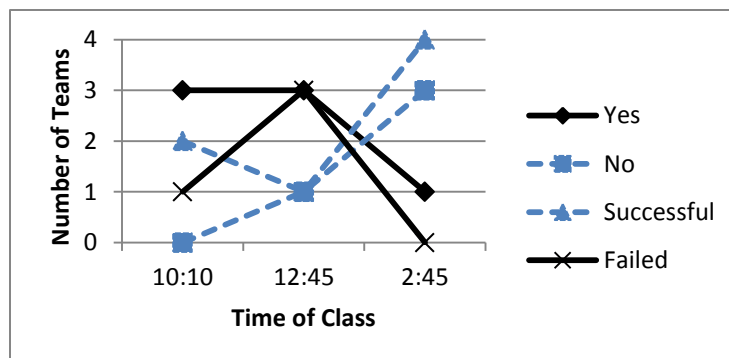


Figure 3: Success versus Pre-Build Designing

This may look like an engineer's nightmare: "higher success rates with less design? That can't be true." Since the activity was conducted within the normal school setting, it is possible that students discussed the activity outside of the classroom. While it is not certain, it would explain the lack of pre-design work and the increased number of successful structures.

***When are the students involving the marshmallow in their design?
(Early on, or just before time is called?)***

10:10 Class

After the activity had concluded, the students were asked a simple question, "how many of you just pushed the marshmallow aside and ignored it?"—*every* student raised his or her hand. The students also commented that the marshmallow seemed light and would not be too much trouble. The students were then asked a follow up question, "what do you think the marshmallow represents metaphorically?" After a few good answers that did not quite hit the mark exactly, the true nature of the marshmallow was revealed to be hidden assumptions in a project—something that appears to be minor, but can be significant enough to render their design an utter failure⁴.

12:45 Class

This class had a unique situation unlike the other two sections. Since this class was offered a prize, the results are understandably skewed. However, the promise of a plump bag of candy as reward did have an effect on the student's behavior—especially concerning the marshmallow. Only one team included the marshmallow early in their design and was the first group to build around the marshmallow, not the base or structure. Other students concentrated on the height rather than the marshmallow. Post-activity answers to the two questions were similar.

2:45 Class

By the final class, the students had picked up on the tricks behind the lab. Every single group included the marshmallow in their design early—even though it was never explicitly made out to be a good idea. Understandably, the students were able to give answers to the questions quicker than previous classes.

Secondary Observations

Over the course of the three sections, a couple of commonalities in questions and small observations were noted:

- Students commonly asked for more supplies.

In the first two classes, this was asked before the challenge had begun. However, nearly every team in the third class asked for more supplies in some form.

- Students would ask if they could tape the base to the table.

The classes were not at fault here. It was never made clear that the structures could be taped to the table, and students assumed they couldn't until they were told. Yet, this would not have affected the results since this discrepancy was dealt with early in the challenge.

- Students asked multiple times about the qualities of the “best” design.

In some way or another, students alluded to the existence of a perfect design in their questions. While it is hard to label a design as the *best*, some designs are far more effective than others. The most common design that was also the most efficient is best represented by Group C3’s design in Figure 4 below.

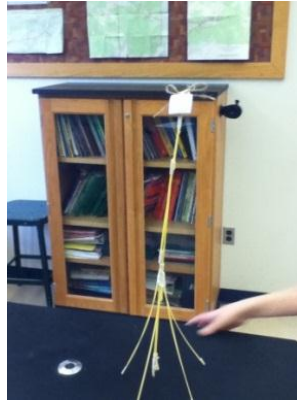


Figure 4: Design C3

The features of this design may be summarized as *a structure consisting of a pyramid / triangular prism base (with “legs” jointed or separate) with spaghetti bundled together built up from the triangle’s / pyramid’s apex*. Structures that followed this design philosophy proved to yield the best results. Students utilizing this design improved over the course of the day with a net gain of 3.5 inches. In fact, the tallest structure built out of the three sections was C3 measured at 21.5 inches.

- The two groups of three performed exceedingly well when compared to the other groups of four in the same class.

In the last class, two groups of three had to be made, C2 and C3. While they were expected to keep up with the other groups of four, C1 and C4, the smaller teams outperformed the larger groups instead—also trumping efforts of the previous classes. Their success could be attributed to more focused team dynamics. With fewer people, C2 and C3 had more cohesion as a team and could focus on tasks better than C1 and C4. This assumption is reassured by observations of C1; there was little relevant communication and the group was rarely on task.

Did Competition Negatively Influence Students?

In a study on how competition could be a barrier to learning, researchers tested to see if students who were concerned about ability relative to their opponents would lower their effort to win by not losing in a competition⁵. In other words, the students came to terms with the fact that they would not be winning the challenge, so the team decided to settle for second best. This study could be applied to the 12:45 section, where a reward was offered. Although there were more failed designs that ended up being disqualified, the 12:45 section did the most pre-design work when compared to the other two classes. While it is difficult to determine whether or not this

phenomena was present in the design lab, tighter monitoring and a set of related questions about their motivation could bring sufficient data to draw conclusions. This question remains to be investigated more fully.

Limitations

Through the lab, there were a few unavoidable limitations. First, the timing of the activity limited the amount of students that could have participated. Since this lab was held close to the school's Thanksgiving break, quite a number of students had already left to go on vacation, which shrunk the pool of competing teams. If this lab experiment had been conducted during a normal week, the amount of teams would have been higher, and there would be more relevant data to evaluate. Second, having equal teams across the three sections would have been ideal; yet, the timing also made this fairly difficult to achieve. Also, the first class began to disassemble their structures before a picture was taken of each; that was simply a miscommunication.

On the other hand, the smaller class sizes did provide a few advantages. Since only three or four teams were competing at once, it allowed for closer observation. The teams could be followed through nearly every step in the design process, and it could be determined approximately where team dynamics tended to crumble.

Concluding Thoughts for Future Research:

K-12 educators often strive to incorporate projects into the science curriculum. Students conduct the activity, fill out the corresponding lab sheet, discuss the results, and then move on to a new topic the next day. With such an abrupt change in pace, this raises a valid question: do these activities have a lasting impact on student learning or are these labs extraneous additions to the general curriculum?

A future paper would examine the students' retention of concepts through the use of a short post-activity test given a certain amount of time after the lab. Scores would be averaged by class and compared while written responses would be evaluated to find commonalities.

Finally, reflection on the criteria, constraints, goals and outcomes of a project is rarely incorporated. A reflection on the project could be administered and assessed.

References:

- [1] Carr, Ronald L., Lynch D. Bennett IV, and Johannes Strobel. "Engineering In The K-12 STEM Standards Of The 50 U.S. States: An Analysis Of Presence And Extent." *Journal Of Engineering Education* 101.3 (2012): 539-564. *Education Research Complete*. Web. 27 Feb. 2013.
- [2] Mann, Eric L., et al. "Integrating Engineering Into K-6 Curriculum: Developing Talent In The STEM Disciplines." *Journal Of Advanced Academics* 22.4 (2011): 639-658. *Education Research Complete*. Web. 27 Feb. 2013.
- [3] Petroski, Henry. *To Forgive Design: Understanding Failure*. Cambridge, Massachusetts: The Belknap and Harvard University, 2012. Print.
- [4] Wujec, Tom. *The Marshmallow Challenge*. Tom Wujec. n. d. URL: <http://marshmallowchallenge.com/>. Accessed on: January 26, 2013.
- [5] Wang, X. H., & Yang, B. Z. (2003). Why competition may discourage students from learning? A behavioral economic analysis. *Education Economics* 11(2), 117–128.

David Reeping is a first year student majoring in Engineering Education with a minor in Mathematics. He is a Choose Ohio First scholar inducted during the 2012-2013 school year.

Dr. Kenneth Reid is the Director of First-Year Engineering, Director of Engineering Education and an Associate Professor in Electrical and Computer Engineering and Computer Science at Ohio Northern University. He was the seventh person in the U.S. to receive a Ph.D. in Engineering Education from Purdue University. He is active in engineering within K-12, serving on the JETS and TSA Boards of Directors and 10 years on the IEEE-USA Precollege Education Committee. He was named the Herbert F. Alter Chair of Engineering in 2010. His research interests include success in first-year engineering, introducing entrepreneurship into engineering and engineering in K-12.

**Appendix
Section I:**

Table 1: Height of Structures

Teams	Height (inches)	Winning Team (Marked as X)
Results of 10:10 Class		
A1	18	X
A2	DQ (2)	
A3	14	
Results of 12:45 Class		
B1	DQ (1)	
B2	DQ (1)	
B3	17, DQ (3)	
B4	21.25	X
Results of 2:45 Class		
C1	10.5	
C2 (Group of 3)	18	
C3 (Group of 3)	21.5	X
C4	14.5	
KEY DQ(1) Structure fell DQ(2) Marshmallow was not at the highest point, but the structure was still standing DQ(3) Structure was not free standing		

Section II: Observations

Table 2: Primary Observations

Team	Supplies Used	Did you sketch out a design?	Tried Marshmallow First (Marked as X)
Results of 10:10 Class			
A1	Spaghetti (16)	Yes (B)	
A2	Spaghetti (18)	Yes (A)	X
A3	Spaghetti (17)	Yes (B)	
Results of 12:45 Class			
B1	Spaghetti (20)	No	
B2	Spaghetti (20)	Yes (B)	
B3	Spaghetti (18)	Yes (B)	
B4†	Spaghetti (19)	Yes (B)	X
Results of 2:45 Class			
C1	Spaghetti (20)	No	X
C2††	Spaghetti (20)	No	
C3††	Spaghetti (20)	Yes (B)	
C4	Spaghetti (20)	No	
<p>KEY (A) drawn after beginning to build (B) drawn before beginning to build † B4 was the only group not to include string in their structure †† Team of 3 students</p>			

Section III: Pictures of Structures

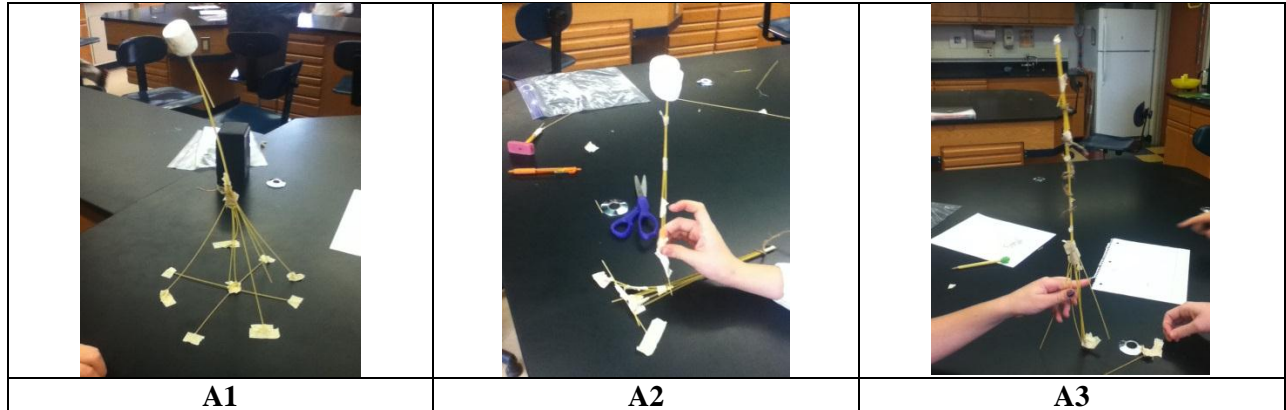


Figure 5: Structures from 10:10 Class

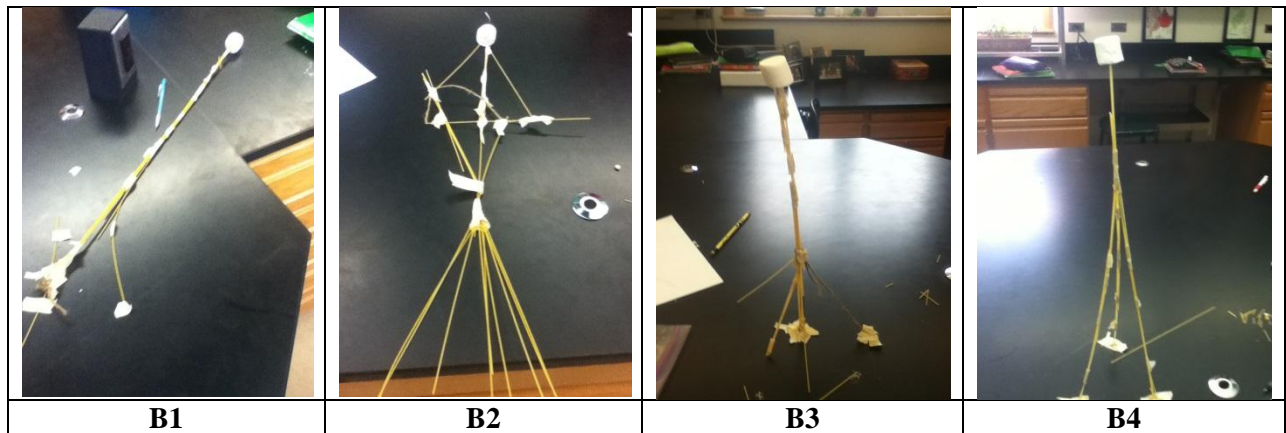


Figure 6: Structures from 12:45 Class

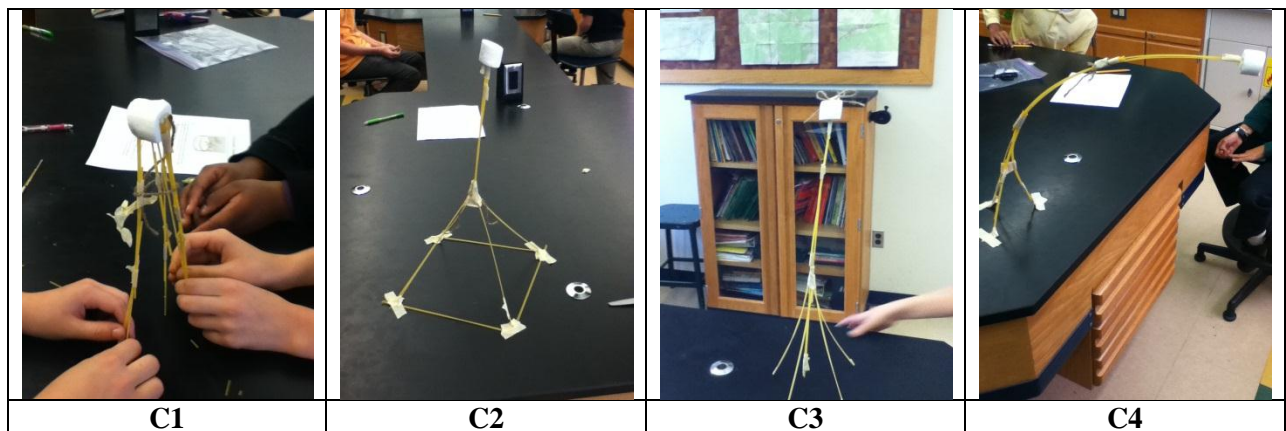


Figure 7: Structures from 2:45 Class