AC 2008-2922: THE ROLE OF DIAGNOSTIC REASONING IN ENGINEERING DESIGN: CASE STUDIES

David Crismond, The City College of New York

Dr. David Crismond is an Associate Professor of Science Education at the City College of New York. He received his masters degree in 1992 from MIT's mechanical engineering department, and earned his doctorate in Human Development and Psychology from the Harvard Graduate School of Education in 1997. His career in education has included public school teaching, developing engineering design-related interactive multimedia materials at MIT, and design-oriented science curricula at TERC and Georgia Tech. He has been Principal Investigator for the NSF-funded curriculum development project, Technology for Science, and an NSF-funded teacher professional development project, Design in the Classroom. Dr. Crismond's main research interests revolve around the issues of K-12 design cognition and pedagogy, and teacher professional development in science and pre-engineering.

Case Studies on the
Role of Diagnostic Reasoning in Engineering Design

Introduction

Design activities have been used in K-12 classes to contextualize student learning of STEM ideas, to raise interest in difficult-to-teach topics, and as transfer tasks to test student understanding. One of the enduring conundrums in engineering design is that designers, regardless of level of experience, can end up with final products that look remarkably similar to their first sketches or prototypes. A number of explanations for this problem, which has been dubbed "functional fixedness" (Cross, 2000) and "idea fixation" (Sachs, 1999), have been proposed for this phenomenon where little seems to get learned or gained through cycles of design iterations. One hypothesis that this study investigates is the notion that idea fixation, especially when done by beginning designers, is simply due to the novices not noticing weaknesses in their current plan or prototype. If all looks well with a prototype or product that performs poorly during testing, there then would be no driving force to change the current design. The lack of capability to notice problems in a sketch, prototype or product may be related to the little studied role of diagnostic reasoning in engineering design.

A review of the existing literature suggests that to diagnose and troubleshoot the systems they are planning, designers need a number of ideas and cognitive tools to do the job. Such a conceptual model may include (1) an understanding of systems; (2) relevant device knowledge and (3) relevant domain knowledge, such as how the device or system works; (4) topographic knowledge (Rasmussen, 1984) that amounts to a mental map of the product or system's physical layout, and (5) and an understanding of procedures for doing troubleshooting and testing hypotheses about possible faults (Jonassen & Hung, 2006). For science teachers using design tasks, domain knowledge would include the science and engineering principles that explain basic product functions (e.g., Hooke's Law for return springs; Newton's Laws of Motion for vehicle or projectile motion). While such understandings by themselves are thought not to be sufficient for doing effective fault diagnosis (Jonassen & Hung; Morris & Rouse, 1985), they may enable practitioners more effectively to transfer their diagnostic and troubleshooting skills to new situations (MacPherson, 1998).

In the context doing engineering design, diagnostic reasoning involves in part the zooming in and out of attention in order to investigate the various levels of system performance. Such attentional focusing can help the practitioner isolate faults, which can reduce the complexity of the system being considered, lessens the load on working memory, and in turn improve troubleshooting performance (Axton, Doverspike, Park & Barrett, 1997). When designing, flaws that get detected via diagnostic reasoning can also inspire ideas for new features or as-yet unthought of systems. As in scientific discovery, noticing unexpected properties or behaviors during testing can be a powerful impetus for conceptual change (Kolodner & Wills, 1996).

This paper's working hypothesis is that diagnostic reasoning is critical for students to notice flaws in their designs, change them and so improve the quality of their final products through the process of iterative design. Teaching students diagnostic reasoning in the context of doing technological or engineering design can become an authentic context in K-12 settings for:

(a) Teaching science and engineering science concepts related to how the device works;
(b) Using inquiry-like observational skills and the capability to zoom in/zoom out attention when analyzing products;
(c) Developing a description of the desired behavior of the planned device; and,
(d) Using systems thinking when predicting or explaining the device.

This paper reports on a study that the author conducted in March-June 2007 and which involved 41 ninth-grade students from Columbia, Missouri, who were taking a year-long course that combined physics and engineering science. Over a three-month period, students' diagnostic reasoning and troubleshooting capabilities were evaluated within the context of the work students were doing on a sequence of 3 design activities and units of study. This included students first designing-and-building model parachutes, then bottle rockets with passively deployed parachutes, and finally small- and then large-scale trebuchets. As a culminating activity and product test that took place in the school parking lot, the 11 teams tested their devices by shooting a projectile (tethered tennis ball) 3 times at a given target from a known distance – 10 meters – and then 3 times each from two distances that were announced only during the day of testing -- 18 meters and 6 meters.

This paper reports on the work of two teams that were randomly selected among the 11 in these classes. Diagnostic-reasoning interviews were conducted at various points in students' work on these three design challenges, where students answered variations on the following 4-item protocol regarding their plans, prototypes and final parachute and trebuchet projects:

(1) Describe observations -- how the device behaved during testing.
(2) Name any problems with a product's performance, if any.
(3) Tell why the device behaved as it did.
(4) Suggest one or more remedies to fix the problem.

This research reports on the reasoning of the student from each of 2 of the 11 teams who was randomly selected for closer study. The subject's thinking was revealed via questions that focused on five aspects related to diagnostic thinking. These include: (a) students' diagnosis of what is wrong with a design based on observations of the prototype's behavior; (b) suggested remedies to fix the design problem; (c) explanations of how the devices being designed work; (d) analysis of the products from a systems perspective; and (e) descriptions of the "ideal" version and preferred behaviors of the devices they are fashioning. Analysis of these interviews that explored in-depth diagnostic reasoning associated with the last design activity (trebuchets) will be combined with commentary regarding the reasoning of the other members of the two teams selected for study.

Trends noted regarding the diagnostic reasoning of individuals in these two teams will then be linked to ratings of the overall product quality of the teams' final large-scale trebuchets, which were given by two blind external evaluators – a mechanical engineering professor and tech ed teacher, both who had experience with these design challenges. These raters reviewed photographs of the trebuchets at various stages of development, as well as videotapes of the devices and final testing done in the school parking lot. They used 14 previously agreed-upon criteria related to the final product's function, structure and behavior. A maximum Quality Score of 10 points for these trebuchets was based on the following scoring rubric:

- ✓ Behavior (3 points) – Accuracy and Repeatability
- ✓ Function (3 points) – Control of: projectile motion during launch, arm friction, arm at triggering, undesired motion in arm, undesired motion in base, undesired motion in superstructure, tradeoff between lever arm and counterweight, easy of loading, ease of triggering, and safety.
- ✓ Structure (2 points) – Durability
- ✓ Economy (1 point) – Minimization of materials
- ✓ Aesthetics (1 point) – Appearance and Craftmanship

Data Analysis

The interviews described below were conducted while students worked on a unit of instruction on trebuchets that was done over the last 7 weeks of the school year. The interviews aimed to get a picture of students' ability to do focused observations, describe their devices as systems and how they work, talk about design decisions and changes they made to their plans and prototypes, and why they made them. The interviews also included a second set of items where students described the model parachutes that they had developed previously. Students were also shown QuickTime videos of 6 parachutes and asked to analyze each according to the 4-item diagnostic reasoning protocol: what do you notice; what name do you give to the problem; why does the problem (if any) occur; and how would you remedy it? These data were used to describe the diagnostic capability of each student, and rank order that individual's capability with others within and across teams. The team's work on their trebuchet is noted by describing changes to the trebuchet prototypes made by the team over the course of the unit. A breakdown of the Quality Score for the final trebuchet is then noted and its relationship to the team's diagnostic reasoning capability discussed.

Team N

With the first of the two randomly selected teams, called Team N in this study, the questions from the protocol were asked of 3 of the 4 team members who were present during the day of the trebuchet interview. Team N's developed 3 separate prototypes in the course of their work (see Figure 1), and finished their final project on time and during the last day that construction on projects was permitted for the class. There was evidence of purposeful and significant design changes being made to a number of the model trebuchet features.

Figure 1. Team N's three versions of their model trebuchet

Team N's final project, shown in Figure 2, was given a Quality Score of 6 out of 10, which made them tied for 2nd place for Overall Quality with two other teams in the class. The class average QS was 5.74. The breakdown of Team N's score (and compared to the averages for the other 10 teams) was: Behavior 1.3 (0.96); Function 2.1 (2.24); Structure 1 (0.86); Economy 1 (0.9); and Aesthetics 0.6 (0.77). Team N's accuracy score for hitting the target at the three distances was third highest in the class.

Figure 2. Team N's solidly constructed final trebuchet used a bushing system to reduce friction and a long structural base to increase stability by reducing recoil

Overall, Team N's answers to the diagnostic interview questions were among the most articulate of all teams tested. All 3 members of Team N's team were able to focus their attention on the performance of their devices, note problems and give explanations for design decisions to correct those problems. The randomly selected individual from Team N, Nathan (a pseudonym), although not the premier diagnostic reasoner in his team, was first to answer the question that asked students to compare the first and latest version of the model trebuchets that the team developed. He mentioned three changes in particular: to the triggering mechanism, the frame and structure that holds the axle and throw-arm subsystem, and the spacers that were used by some teams to keep the throw-arm centered on its axle. Nathan's explanations for these changes to the structure were well-reasoned: "The brace that was going to attach the base to the tower, it doesn't really matter what the angles are. It's a bracket and all it has to do is form a triangle." The causal reasoning he used to justify design decisions was not always complete, but this was corrected by another teammate, Rob, when the details in the latter's explanations for certain choices fell short.

Nathan was subject to certain misconceptions about the mechanics of trebuchets. One that was quite common among the 11 teams involved the preferred angle of release for an ideal trebuchet. In their physics class, students had been taught that a 45° angle was preferred for their system. Most, however, interpreted this to mean that the throw arm should be at that angle, not that the projectile should depart the trebuchet at a 45° angle, which typically was when the arm was near its vertical position and undergoing in most treb designs maximum velocity.

One change made to Team N's early prototyope involved using a nut-and-screw to hold the counterweights in place instead of a bent copper wire. This alteration was instituted in order to avoid a flaw that Nathan diagnosed, "It won't fall apart like the last time", and allowed the team to test different weights more efficiently. When team member Rob was asked to tell the pros/cons for using a free-swinging versus fixed weight, he noted that the counterweight with a linkage would " smoothly drop like that… and then you'll get a rocking back and forth." Nathan's more precise observational skills were in evidence when he correct his colleague and said, "It'll jerk, though. … it's just a theory for me, but if it's a [free] swinging counterweight, it's going to have more momentum, than just with the fixed counterweight."

Nathan was ranked second in his squad for diagnostic capability. Top ranked in Team N was a student with a real diagnostic talent who was able to (1) take note of a problem, (2) *see that it was possible to correct the problem*, and (3) suggest one or more remedies for the problem. Mitch was Team N's talented loner-designer whose keen skills at observing mechanical devices were complemented by mental models of how mechanisms such as trebuchets work as well as by excellent fabrications skills. The last probably arose from Mitch's experience on the family farm, where he was involved in building and repairing mechanical devices.

Mitch was the first person of all subjects in this study to propose a bushing system to reduce friction, and one of the first to notice the recoil and propose ways to eliminate the rocking of the trebuchet structure. Other students did noticed this motion, both on their own and when asked to describe how their devices worked. When asked to predict in which way the trebuchet first moved during recoil, students in the study were more often incorrect than correct. Mitch not only noticed the problem of recoil ("Really, they all do that, they all throw themselves"), but also correctly identified that the trebuchet's base initially moves forward in response to the counterweight swinging during descent towards the rear of the device, and suggested ways to remedy the problem.

Diagnostic reasoning prowess also involves a focusing and "zooming in" of attention to problem areas in a design. Mitch demonstrated this when he described the performance of his team's first trebuchet prototype:

> "If you watch, this thing [freeweight] comes down with so much force. You see, right there, right there, it hit, and the whole thing shook, so like it more than likely will break."

This performance was in notable contrast to that of another Team N member. Rob was a capable physics student and accurately applied physics concepts when describing how the trebuchet. Specifically, Rob most capably incorporated the notions of kinetic and potential energy when describing the trebuchet in action, compared to the other Team N members. However, Rob was least capable to use the diagnostic vision needed to notice problems with the assurance that would lead to actions that can solve the problem. Rob's performance in reviewing QuickTime movies of his team's devices and discussing the problem of recoil, showed that he was capable of analyzing the details contained in these videos:

> "If you notice, the weight hasn't even left, if you watch this, right now… it throws itself up before, the weight hasn't even left and its, the back end is leaving the ground and throwing itself forward. Which could be a problem."

Second-tier diagnosticians, like Rob and Nathan of Team N, were capable of noticing anomalies in product performance, especially when asked to do so in the context of reviewing video playbacks at regular speed or frame-by-frame. As can be noted in the last statement quoted above, they are less certain about their judgment that what they have noticed is a genuine problem, or that a problem they have tentatively diagnosed might be fixed.

Team Z

A member of Team Z was the second randomly selected subject for case analysis this study. Three of Team Z's four team members were present for the interview reported below. Team Z was notable that it developed but did not have adequate time to test its 2 prototypes (see Figure 3). Construction of these prototypes was of marginal quality; the earlier device was not operable at the time of the interview. The number of critical subsystems mentioned in Team Z's discussion of early and late prototypes was far fewer than those described by Team N.

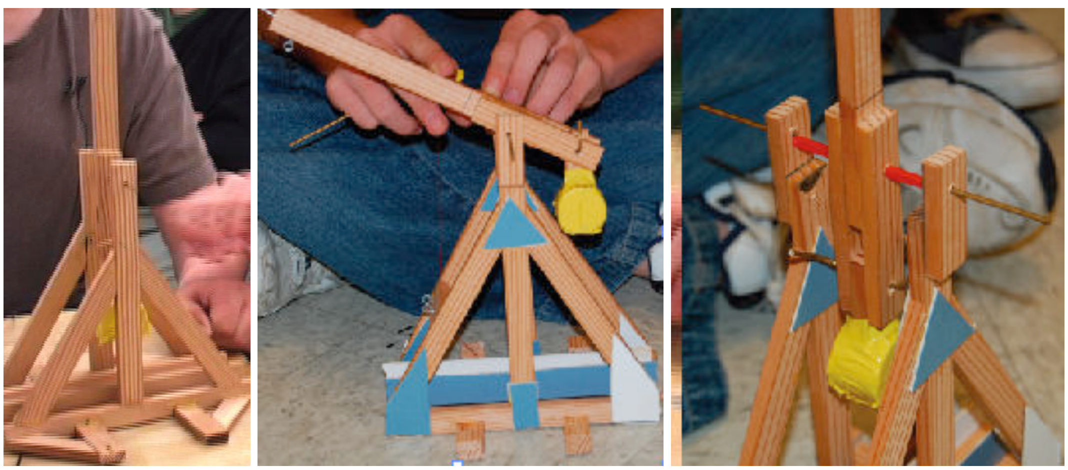

Figure 3. Team Z's two model trebuchets went largely untested

Team Z finished construction on its final trebuchet during the first day of actual testing. The Overall Quality score given by the two external reviewers for the device shown in Figure 4 was 4.4 out of 10. This meant that Team Z's final project was ranked behind that of all other teams in the class, where the class average for quality was 5.74. The breakdown of the quality score for this team (and compared to the averages for the other 10 teams) was: Behavior 0.5 (0.96); Function 1.7 (2.24); Structure 0.4 (0.86); Economy 0.8 (0.9); and Aesthetics 1 (0.77).

Figure 4. Team Z's trebuchet lacked a mechanical trigger and felt apart during final testing

Team Z's answers to the diagnostic interview questions were less detailed those most teams in this study. This may be due in part because the team as a whole was poor at time management and spent little time in the proving grounds of testing prototypes before it needed to build the next version or final product. However, whether the speed of production is in based in part on skills like diagnostic reasoning and other cognitive strategies related to engineering design cannot be determined in this study.

The randomly selected subject selected for case study, Forrest (a pseudonym), was ranked by the author as the mid-level diagnostic reasoner of his team. Forrest was the spokesperson for his team, and was first to answer the question involving the comparison of the first with the latest model trebuchet. He mentioned three changes related to his team's prototypes: changes to the base structure and materials used to make the device more durable, and changes in the bracing supports. "We added these things [angled floor supports] to help support it, keep it in place so that it was not moving. … We have them on all 4 corners -- so that it can have the most support it can get."

The least skillful diagnostician in Team Z was Ralph, also discussed the materials used in fabricating the model trebs: "The part that I thought was cool as hell – this was made out of paper while this was made of wood." *DC: Why did you think that was interesting?* "I just thought it looked cool."

Ralph and to some degree Forrest, were noted as commenting upon superficial features of their designs, and were not focused on features that would significantly contribute to better performance in their devices. Forrest also tended to overstate claims he made about the performance of the brace, and his team's work in general.

When asked questions that involved the analysis of the model trebuchet's subsystems, Team Z was quite articulate and could divide up the entire device in a sensible way. They spontaneously identified a crucial component among these subsystems: the hook which holds the tethered projectile. Tiger, the best diagnostician in the team, was quick to noted the following:

"The [most important] is the hook, if it is off just a little bit, it can mess up your catapult, no matter how good it is." Forrest added, "It's one thing to get right." However, when asked what the plan was for the hook's design in the large-scale trebuchet, Team Z's answer showed that little thought had gone into the design of this component of the entire device:

> DC: What are you going to make your hook out of?
> Forrest: Bigger metals.
> Ralph: It's just a bent bar.
> Forrest: Actually, we don't really know yet, because we haven't gotten that far.

Team Z was subject to the same misconception mentioned earlier about the preferred angle of release, and the prediction direction of recoil was incorrect. However, when shown digital movies of a release (not a test using a projectile) of their two prototypes, both Forrest and Ralph noted general rather than specific comments of trebuchet behavior ("It tends to wobble here."), and made conclusions that led to little action "It rocked forward and rocked back, that was good." It was interesting to note that a mental model of what are preferred behaviors was only modestly developed in Team Z members.

Team Z's weaker diagnostic capability was notable throughout this interview, and was most prominent in the less-than-focused, generalized way that products were reviewed during the day-to-day process of design work on an extended project. Far fewer concepts related to the underlying science of how trebuchets work were noted in Team Z than N, but systems thinking seems to be equivalent among the groups. In all, Team N had 3 of 4 members who were competent at this form of thinking, while only one member of Team Z was skillful at observing, naming, explaining and remedying problems in prototypes and final products.

Conclusions

The role that diagnostic vision play in designing is an understudied area of engineering design research that relates in some design model to the latter phase of reflective practice (Schön, 1983). Studying the relationship between product quality and diagnostic reasoning is a fruitful area of future study. The results of this case-study comparison preliminarily suggest that effective diagnostic reasoning may be related to better product design, since such a capability enables designers to learning more from product testing. When effective diagnostic reasonsers perform more design iterations, they may achieve greater insights into their devices and discover more features to improve while still working within the time-constrained setting in which most design work takes place.

Disparities between intended function and the product's behavior during testing cannot shape iterative design efforts if they go unnoticed. What students learn as they iteratively plan, build and test their designs depends in part on whether their attention is focused or diffused. Helping teachers to get students to notice critical and problematic features in their designs is a challenge that may be helped with the simple application of the 4-item diagnostic reasoning protocol used in this study, though the efficacy of such actions will require future testing in laboratory and classroom settings.

REFERENCES

Axton, T. R., Doverspike, D., Park, S. R., & Barrett, G. V. (1997). A model of the information-processing and cognitive ability requirements for mechanical troubleshooting. *International Journal of Cognitive Ergonomics, 1*(3), 245–266.

Cross, N. (2000). *Engineering design methods: Strategies for product design, 3rd Edition.* New York: John Wiley & Sons.

Jonassen, D.H., & Hung, W. (2006). Learning to troubleshoot: A new theory-based design architecture. *Educational Psychology Review, 18*(1), 77-114.

Kolodner, J.L., & Wills, L.M. (1996). Powers of observation in creative design. *Design Studies, 17*, 385-416.

MacPherson, R.T. (1998). Factors affecting technological troubleshooting skills. *Journal of Industrial Teacher Education, 35*(4), 1-18.

Rasmussen, J. (1984). Strategies for state identification and diagnosis in supervisory control tasks, and design of computer-based support systems. *Advances in Man-Machine Systems Research, 1*, 139–193.

Sachs, A. (1999). Stuckness in the design studio. *Design Studies, 20*(2), 195-209.