



## Investigating the Impact of Model Eliciting Activities on Development of Critical Thinking

### Dr. James A. Kaupp, Queen's University

Researcher and Adjunct Professor (Msc '06, PhD '12) at Queen's University, Kingston, Ontario, Canada in the Faculty of Engineering and Applied Science. Educational research interests include engineering education development, critical thinking & problem solving, outcomes based assessment and interactive learning through technology. Scientific interests include regenerative medicine, tissue and biomedical engineering and human biomechanics.

### Dr. Brian M Frank P.Eng., Queen's University

Brian Frank is an associate professor in Electrical and Computer Engineering, where he has taught courses in electronics and wireless systems. He is the DuPont Canada Chair in Engineering Education Research and Development, and the Director of Program Development in the Faculty of Engineering and Applied Science where he works on engineering curriculum development, program assessment, and developing educational technology. He is a co-founder of the Canadian Engineering Education Association and is currently coordinating the Engineering Graduate Attribute Development Project, working with National Council of Deans of Engineering and Applied Science and the Canadian Engineering Accreditation Board, to develop national guidelines and resources for outcomes assessment in engineering education.

### Dr. Ann Shih-yi Chen, Queen's University

Ann Chen is an Adjunct Assistant Professor at Queen's University. Before coming to Queen's, she completed her PhD in second language education at the University of Cambridge and worked for a major testing company in the UK, where she helped to develop a nation-wide language proficiency test. More recently, she has been involved in a variety of educational studies employing mixed-method research designs.

# Impact of Model Eliciting Activities on Development of Critical Thinking

## Abstract

This paper presents a portion of a study on how model eliciting activities (MEAs) impact critical thinking development in first year engineering. MEAs were integrated into a first year engineering course, alongside instruction in critical thinking skills (CTS). CTS were benchmarked and assessed by pre- and post-testing using the Cornell Level-Z and the International Critical Thinking Essay Test. In addition to these tests, think-aloud protocols were conducted and coded to assess CTS development between experimental and control group volunteers. Preliminary analysis of the coding illustrates a potential difference in CTS between the experimental and control groups, which the authors believe may be attributed to MEA instruction. Further analysis of the pre- and post-tests, student experience and background surveys, and the remaining units of analysis in the think-aloud protocols is required before any further conclusions can be drawn.

## Introduction

Model eliciting activities (MEAs) are realistic problems used in the classroom that require learners to document not only their solution to the problems, but also their processes for solving them (Shuman, Besterfield-Sacre, Bursic, Vidic, & Sieworiek, 2012). MEAs have been developed and used in a variety of subject areas, including mathematics, economics, and environmental engineering. Studies have shown MEAs to be valuable in helping students to develop conceptual understanding, knowledge transfer, and generalizable problem-solving skills.

MEAs have been integrated into a first-year undergraduate engineering course at Queen's University, a medium-sized Canadian university. Students in this course are asked to work collaboratively on three different MEAs, introduced in a three-week cycle. While each MEA requires students to employ different areas of subject knowledge, students are taught to approach all three MEAs using critical thinking skills. For example, students are guided to draw concept maps, question the credibility of information sources, incorporate a range of factors into their decision-making, and consider the implications of their conclusions. These skills are what Paul calls "elements" of critical thinking—invaluable thinking processes involved in any complex problem-solving activity (Paul & Elder, 2006).

A research team has been formed at the university to investigate the impact of the MEA-integrated course on students' development of critical thinking skills. Ultimately, the team aims to determine whether the MEA-integrated course facilitates students' critical thinking. During the fall semester of the 2012/2013 academic year three instruments will be used to evaluate the critical thinking skills (CTS) of first year engineering students. These instruments will be used

as both a pre- and post-test in order to benchmark CTS of the incoming first year students, and determine the effectiveness of MEA instruction on developing student critical thinking ability. These instruments are the Cornell Critical Thinking Test Level Z (Cornell Z) (“Cornell Critical Thinking Test,” n.d.), International Critical Thinking Essay Test (ICTET) (“International Critical Thinking Essay Test,” n.d.), and think-aloud exercises.

Queen’s University is collaborating with engineering programs at three other universities to evaluate critical thinking using the Cornell Level Z at the start and end of the 2012/2013 fall semester. First year student volunteers from the physics department at Queen’s were solicited for participation in the study to serve as a control group, due to similar curriculum. The physics department and other participating institutions are serving as comparison and control groups due to similarities in incoming student characteristics and/or learning environment.

Institution	# of Participants	Control/Compare Grouping		Intervention	Pre/Post Tests			
		Offers 1st Year Design Course	Critical Thinking Development	Intervention Activity	CLA	Think Aloud	Cornell Z	International Critical Thinking Essay Test
Queen's Engineering	722	✓	✓	✓ (MEA Instruction)	✓	✓	✓	✓
Queen's Physics	8	✗	✗ (Control)	✗	✗	✓	✗	✗
University of Toronto	32	✓	✓ (Compare)	✓ (Design Course Elements)	✗	✗	✓	✗
Dalhousie University	N/A	✗	✗ (Control)	✗	✗	✗	✓	✗
Memorial University	N/A	✗	✗ (Control)	✗	✗	✗	✓	✗

In this study, the MEA-integrated curriculum is an independent variable (or intervention); students’ critical thinking ability is a dependent variable.

The entire first year student body (approximately 650 students) participated in an MEA-integrated curriculum, and were invited to participate in the broad study, which allowed the researchers to use the scores from their MEAs and critical thinking tests. Stratified sampling was used to assign various pre and post instruments. These assessments were part of the course requirements, so the participation rate was close to 100%. Additionally, students were invited to participate in think-aloud activities that are the focus of this paper. Such protocols originate from cognitive science, but have been widely used problem solving and critical thinking research (Douglas, Koro-Ljungberg, McNeill, Malcolm, & Therriault, 2012; Ku & Ho, 2010; Norris, 1990; Steif, Lobue, Kara, & Fay, 2010). A total of six students were part of this study.

### Instructional context

The study was set in the first year of an undergraduate engineering program at [institution name masked], a medium-sized university in Canada with a first year engineering class of approximately 650 students. The program has a common first year - all students take the same courses, and are allowed free choice between ten engineering programs in their second year. In the first semester of the program students take courses in statics, chemistry, earth systems,

engineering graphics, and calculus. They also take a full-year course focused on engineering design and practice.

APSC-100 is a team-based, project-based course to promote a sense of curiosity about engineering, and promote creative thought. The course is divided into three modules: Module 1. Problem analysis and modeling; Module 2. Experimentation and measurement; Module 3: Engineering design. Each of these is one semester long and equivalent in weight to a standard one-semester engineering course. This study was embedded into the delivery of the problem analysis and modeling module (module 1) (Frank, Strong, Sellens, & Clapham, 2012).

The problem analysis and modeling module (module 1) is a semester-long integrative experience that uses concepts from engineering sciences, natural sciences, and mathematics courses to solve complex open-ended problems. The course is structured around three complex problems known as model-eliciting activities (MEAs) that were addressed sequentially in three-week blocks over the semester. The situations described in the MEAs require students to create and use a mathematical model of a physical system using MATLAB, and deal with professional issues including ethical dilemmas, conflicting information, and incorrect/missing information. MEAs have been used in the course for the past three years (Frank & Kaupp, 2012).

The module learning outcomes were:

1. Apply a prescribed process for solving complex contextualized client-driven problems (ill-defined, multiple constraints, problems, unknown information)
2. Create and apply appropriate quantitative model and analysis to solve problems.
3. Effectively communicate technical information following a prescribed format and using standard grammar and mechanics.
4. Apply concepts including occupational health and safety principles, economics, law, and equity to engineering problems.
5. Identify and resolve a simple ethical dilemma by applying professional codes of ethics and engineering standards.
6. Apply critical and creative thinking principles to solve contextualized problems.
7. Apply numerical modeling tool (MATLAB) to create model used for solving complex problems.

The module was structured help students develop confidence and skills in solving complex engineering problems – problems for which all information is not known, in which there is ambiguity, where the goals are not necessarily clearly defined.

In most weeks the one-hour-per-week lecture followed a structure like:

- The instructor presented a recent problem or news article related to the lecture objective; in some cases the students respond to a problem posed using a web-based audience response system (ARS) in teams
- The instructor presented or reviewed the problem being solved during the three-week session
- The instructor led a short discussion on a topic related to the problem being studied

- Students work on some component of the problem in small groups which will contribute to the solution; in many cases this included an open-text question answered using the web-based ARS.

The students also attended a 2 hour studio each week. The studio opened with a short quiz on prior reading and/or online videos on MATLAB, followed by a short discussion of some MATLAB concept. The majority of the studio focused on a problem that contributed to the current MEA. Students received a small mark each week for completing the task, encouraging students to keep up with the course material.

The three MEAs were:

- MEA 1: Cable ferry failure (Weeks 2-4): This problem focused on the failure of a cable ferry.
- MEA 2: Wind turbine design (Weeks 6-8): This problem will focus on the analysis and design of a wind turbine.
- MEA 3: Building heat loss (Weeks 9,11,12): This problem focused on the design of the insulation for a net zero home.

Each MEA required students to develop a model of a physical system that was used to solve a problem presented by a fictitious client, write MATLAB code to implement the model, and self-evaluate their report against 3-5 of nine critical thinking elements identified in the Paul and Elder critical thinking model (clarity, accuracy, relevance, logicalness, breadth, precision, significance, completeness, fairness, and depth). Table 1 show the elements embedded into each MEA.

Critical thinking elements were explicitly targeted in all three by discussing principles in class, using in-class activities, and embedding into the MEA requirements. In during one of the lectures focused on MEA1 students created lists of the kinds of questions that should be asked when investigating an accident, which led to a discussion about asking questions. In their final deliverable they were required to identify the kinds of questions they would upon arriving at an accident investigation site. In MEA2 students were required to summarize relevant information including an assessment of its source and credibility, uncertainties, and biases.

**Table 1 - characteristics of the three MEAs**

Category	MEA1	MEA2	MEA3
<b>Technical</b>	Stress and strain, drag	Fluid flow, lift	Heat transfer
<b>Design</b>	Problem definition, concept mapping	Decision making (e.g. weighted evaluation matrices)	Decision making (e.g. weighted evaluation matrices)
<b>Professional</b>	Safety, risk assessment, concept maps	Associations, codes and standards	Economics, codes of ethics, equity

<b>Critical thinking</b>	Asking questions, uncertainty in information, identifying erroneous or conflicting information	Assessing information credibility, argumentation, assumptions, inferences	Bias, inferences
<b>Communications</b>	Report format, English usage, argumentation	Report format, English usage, argumentation, concision	Report format, English usage, argumentation, concision

The conceptual framework for the module is shown in Figure 1. Course activities were designed to introduce teaming skills, and encourage students to their learning, and processes used to solve these problems, and to continue improving them (self-regulation). These are shown in the centre of the circles below. Students were encouraged to apply the elements of reasoning to the problems they solved (shown in the middle ring below), some of which were discussed in class. Student self-evaluated their submission against the work standards shown in the outermost ring.

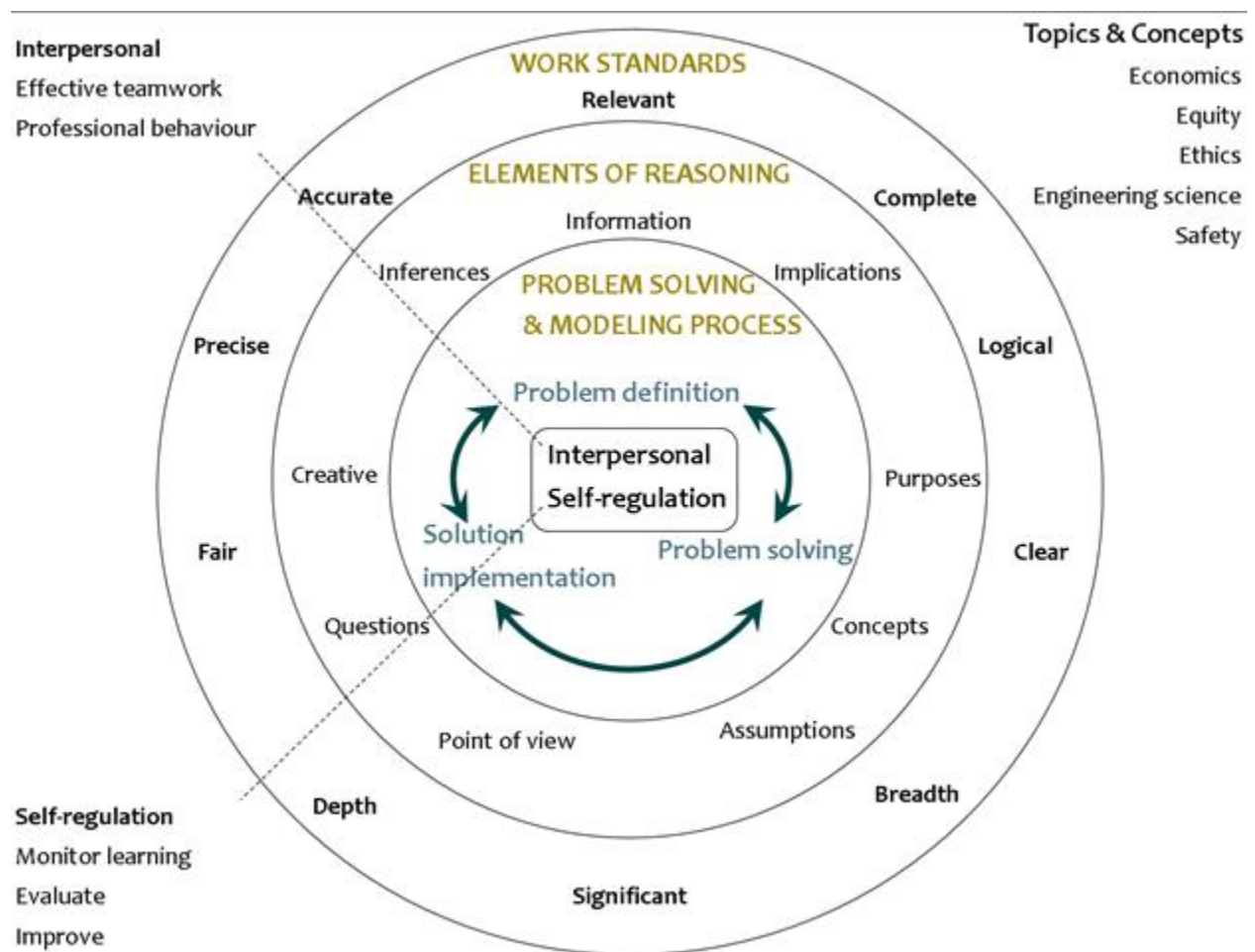


Figure 1- Conceptual framework used for course.

## **Think aloud study**

A control group of participants was recruited from students enrolled in first year physics course in the Faculty of Arts and Science. These students do not see MEAs in the program, and generally these students do not take courses explicitly target open-ended problem solving.

The control group was composed of three physics students, all of whom participated in both the pre and post study. The engineering (experimental) group originally consisted of three students, but one failed to attend for the pre-test, so the pre-test involved two students. The engineering group post-test also had one no-show, so in the end only one of the two students participated in both the pre and post-test.

The think aloud pre-tests for both groups were conducted in the initial weeks of the fall semester (week 1 or 2) in order to minimize potential effects from other courses. It was assumed that the educational backgrounds for both groups would be similar, as both groups are comprised of high school graduates accepted into their respective programs. However, no more information on subjects was collected regarding other academic and non-academic factors prior to the fall semester, which have been shown to affect CTS development (Terenzini, 1995).

The post tests were conducted in week 11. In the think aloud sessions, groups of 2 to 3 participants were first introduced to the facilitator and their background and introduced to the expectations of the think aloud exercises and that the session would be videotaped. The subjects were then taken through a warm up exercise in which they had to provide improvements for a common appliance or simple machine (e.g. bicycle, washing machine), and they would present those improvements to the facilitator. Following the conclusion of the warm up activity, the actual think aloud began.

Participants were presented with a loosely defined, open-ended scenario with a system based on a fundamental physics problem. The scenario asked the subjects to evaluate the system and some request, and provide solutions for the requests, along with any additional safety recommendations they see fit. Subjects were provided with supplemental information of varying pedigree to help them with their recommendations. Subjects had one hour in which to solve the problem while ‘thinking aloud’ their solution. At the end of the hour the subjects were asked to present their recommendations.

### **Pre-test**

The pre-test think aloud tasked subjects to provide safety recommendations to a city council regarding their proposed toboggan hill for a winter festival. An email to the team from the city council was provided, outlining their request and the details of the problem. Subjects were provided with supplemental material to help address the problem.

- An independent opinion on toboggan safety
- A newspaper article on tobogganing safety

- A student-created list of friction coefficients
- A textbook excerpt of friction coefficients
- Information about average mass of American children and adults
- A scientific article on human tolerance and crash survivability
- A physics equations sheet

The primary areas of contention for this scenario are regarding the pedigree of the supplemental information and that the initial parameters provided result in very unsafe slope conditions. The initial email from the city council includes an unsolicited reference from a councillor with a military background regarding human impact tolerance. This reference, alongside the independent opinion on toboggan safety, parts of the newspaper article and the student-created list of friction coefficients should be identified in some way by the subjects as potentially unreliable. Ultimately the subjects should realize that the hill, under initial conditions is very unsafe and should proposed informed recommendations resulting from informed analysis.

### **Post-test**

The post-test think-aloud tasked subjects to provide safety recommendations to an amusement park, FunZone Amusements, regarding their proposed prototype rollercoaster. An email to the team from the company was provided, outlining their request and the details of the problem. Subjects were provided with supplemental material to help address the problem.

- Summary of American Society of Testing and Materials (ASTM) Standards on Amusement Park Device Design
- Scholarly articles on Rollercoasters and G-Forces
- Reports & Articles on roller coaster safety
- A physics equation sheet

The primary areas of contention for this scenario are regarding the conflicting information of the supplemental material, the interpretation of a professional standard and the rollercoaster parameters. The initial email from the company includes some unsolicited guidelines regarding track dimensions and average velocity. These reference dimensions lead to potentially unsafe conditions for passengers. Upon further analysis of supplemental materials, subjects will have to integrate multiple references to form a cogent argument. The question regarding re-using existing carts and restraints also challenges the subjects' ability to interpret information on a technical chart and classify their solution by a professional standard.

### **Think Aloud Framework**

The videotapes of the pre and post think aloud interview sessions were transcribed and annotated. The research team divided students' think aloud protocols into five segments with each segment consisted of a particular issue with which the participants addressed. For this paper, the research



team selected the safety recommendations unit of analysis, as the team felt that this unit would potentially display the greatest amount of elements of critical thinking. The safety recommendations unit was then analysed through open coding (Leydens, Moskal, & Pavelich, 2004; Wiersma & Jurs, 2005), identifying the elements of critical thinking according to Paul's theoretical framework (Paul & Elder, 2006), as illustrated in Table 2. Once coded, the safety recommendations were reviewed for common themes and notable differences between pre and post-tests in the experimental and control group, as well as comparing and contrasting the two groups.

**Table 2- Paul's theoretical framework for elements of critical thought.**

Elements	Indicators
Purposes	<ul style="list-style-type: none"> <li>· Did the participants clarify the purpose of the task given?</li> <li>· Did the participants ask about the purpose of supplemental material given?</li> </ul>
Questions	<ul style="list-style-type: none"> <li>· Did the participants clarify what questions they were supposed to answer?</li> <li>· Did the participants have a plan of action to answer the questions that they identified?</li> </ul>
Points of View	<ul style="list-style-type: none"> <li>· Did the participants ask whether there were other relevant viewpoints that should be considered?</li> <li>· Did the participants ask about the viewpoints expressed in supplemental material given?</li> </ul>
Assumptions	<ul style="list-style-type: none"> <li>· Did the participants identify or question their own assumptions?</li> <li>· Did the participants ask the extent to which their own assumptions were valid?</li> <li>· Did the participants identify assumptions made by authors of supplemental material provided?</li> </ul>
Data/information	<ul style="list-style-type: none"> <li>· Did the participants identify what information they were lacking?</li> <li>· Did the participants ask how they could get the information that they needed?</li> <li>· Did the participants question the source of supporting information?</li> </ul>
Concepts	<ul style="list-style-type: none"> <li>· Did the participants ask whether the concept or theory considered applicable to the given situation?</li> <li>· Did the participants ask whether there was another theory or principle that would better explain the given situation?</li> </ul>
Inferences/conclusions	<ul style="list-style-type: none"> <li>· Did the participants ask whether their conclusions were supported by their analysis or supplemental material provided?</li> <li>· Did the participants ask whether there were alternative conclusions that would also fit the data?</li> </ul>

## Analysis of Engineering Pre-test Safety Recommendations Unit

In this pre-test, the subjects balanced their conclusions between assumptions, either drawing from personal or anecdotal experience or the subjects' previous knowledge base of physics concepts, supplemented occasionally by data and literature. The subjects' conclusions did not question validity or authenticity of supporting literature and were quickly adopted and used alongside assumptive reasoning to form their conclusions. Subjects did not question the validity and accuracy of their own assumptions, and only briefly considered additional implications of the problem

Conceptually, the subjects displayed a comprehensive understanding of the problem and the different solution methodologies, integrating force concepts, kinematics, energy and impact and momentum into their solution. They utilized these concepts in formulating their conclusions, particularly by understanding the relationships between friction, sled velocity and impact.

However, the conclusions reached by the subjects that resulted from the use of concepts were not expanded into a well-sourced argument supported by data from their solutions, but instead drew upon personal experience assumptions in making safety recommendations. For instance, the subjects' recommendation for a strip of rubber to increase friction to reduce speed was based upon conceptual knowledge that an increased coefficient of friction will ultimately reduce the velocity, but the implications of such a transition was only briefly raised and wasn't a focus of further analysis. In a following conclusion, the subjects recommend a minimum depth of snow for slope conditions, and raise the issue of sudden stopping being hazardous but don't connect this with their rubber transition.

### Recommendation regarding increasing friction

[01:02:05]

>> Subject 2: Yeah, [multiple speakers] so lean on the side of caution, definitely half the hill for dry conditions. Do not allow people onto icy parts of the hill. That is because there is essentially a zero kinetic friction on ice. Also, if the hill is not covered by at least twelve inches I believe it was of snow, recommend caution going down. And -- oh, if there is grass being shown, don't allow them to use that section of the hill.

[01:02:41]

>> Subject 1: Definitely.

[01:02:43]

>> Subject 2: And -- so that is because stopping and starting again result in a safety hazard and result in injuries.

Similar reasoning was utilized regarding recommendations for the slope conditions of the hill, in which the subjects made conclusions that originated through their conceptual background supported by assumptions and personal experience rather than detailed analysis or supported by literature, such as safety fences, depth of snow on the hill, and the use of a wetting agent on the hill. While some of these conclusions are very valid, the implications of the use of a wetting agent on a winter sliding hill are not considered, which may potentially lead to ice formation and unsafe conditions.

Subjects did utilize supplemental data and information to form conclusions for safety recommendations, both in hill conditions and in the use of helmets. However, the authority sources of the information were not questioned, as these were independent opinions, or a newspaper article with mention of scientific study. The information presented was simply accessed and assimilated into their solutions without any concerns for accuracy or credibility. The subjects did not also address the limitations in the source that reported data only to a speed of 8 m/s for helmets, but relied on concept based assumptive relationships between the height of the hill, co-efficient of friction and velocity, of which the accuracy and implications remained unquestioned.

### **Analysis of Engineering Post-test Safety Recommendations Unit**

In this post-test, the subjects based their conclusions off of their initial analysis supplemented by concepts and supplemental information. There were minimal conclusions based off of assumptions formed by first person or anecdotal evidence or experience. Interestingly, the assumptions that were made were quickly addressed by the subjects as being potentially untrue or invalid and recommended that they be retracted.

Conceptually, the subjects displayed an understanding of the underlying physics within the problem, quickly selecting energy as their solution methodology. The subjects self-assessed their methodology and formed an equality describing the physical parameters of the system. However, with further analysis the subjects performed some mathematical errors, which ultimately led them to some inaccurate conclusions.

Subjects evaluated the information that was presented to them, placing trust in scientific publications from journals and incorporated this information into their conclusions. They also utilized the information in a supporting role to evaluate conceptual accuracy and refine analyses that formed the basis for some of their conclusions. Subjects displayed the ability to quickly read, identify and summarize key information in supplemental materials and incorporate them into their conclusions rather than become frustrated with the density of the text and disregard the supplemental material entirely.

[00:33:09]

>> Okay, so essentially this entire article summarized there's a few numbers like 4, 6 G's-- yea, this article essentially says that you can't use just absolute value with the G force to say-- to determine how safe something is. You also have to look at the time integral over which it's applied. So, I think for the most part roller coasters cannot fall into a category where you can generate too many G forces to fast. That's kind of the summarization of this article.

Interestingly, the subjects considered potential implications in their conclusions and highlighted addressed their own limitations. The subjects were concerned that what they considered to be the height would not be capable of producing the recommended speed. They recommended that the height be increased, but should obey additional requirements such as building codes and

zoning regulations. Regarding their own limitations, subjects were hesitant to comment or make comprehensive conclusions on areas that they were uncertain, but expressed the need for further investigation in their recommendations.

### **Comparison of Engineering Pre-Post**

The engineering subjects illustrated improvements in elements of critical thinking, specifically in areas of concepts, information and implications. In the pre-test, the subjects balanced their conclusions between assumptions, either drawing from personal or anecdotal experience or the subjects' previous knowledge base of physics concepts, supplemented occasionally by data and literature. The supporting literature was adopted and used alongside assumptive reasoning to form their conclusions, without any further investigation. Subjects did not question the validity and accuracy of their own assumptions, gave the briefest consideration to potential implications of their solutions.

In the post-test, the subjects based their conclusions off of their initial analysis supplemented by concepts and supplemental information. There were minimal conclusions based off of assumptions formed by first person or anecdotal evidence or experience. Concepts of physics and forces were used in a progressive manner, with subjects developing mathematical relationships between variables rather than simply using the equations. The validity of supplemental information was questioned, with the subjects considering the authority of the provided information. Subjects also considered the implications of their conclusions; highlighting limitations and possible areas of concern within their conclusions to be addressed in the future and even highlighted a possible conflict with building codes with the track being an object of considerable height.

The subjects demonstrated improvement from the pre-test, particularly in the elements of concepts, assumptions, information and implications. Subjects questioned the validity and source of supplemental information as well as questioning the accuracy and validity of their own assumptions and conclusions, in contrast to the pre-test where this was poorly demonstrated. Concepts were used beyond simple application, as the subjects used concepts along with mathematical formulae to produce a mathematical model relating the design parameters that required recommendations. Lastly, the subjects vastly improved in implications, considering additional codes and requirements outside of the scope of the recommendations and readily highlighted potential gaps in their knowledge that should be improved upon.

### **Analysis of Physics Pre-test Safety Recommendations Unit**

In this pre-test, subjects experienced a great deal of difficulty with the problem. Initially the subjects discussed potential safety recommendations prior to completing any sort of analysis, forming a list of needed safety recommendations. Subjects then proceeded to survey the supplemental material for information. Subjects surveyed the information and selected what they considered appropriate within the context of their own personal experience. They had selected the coefficient of friction, slope conditions, helmet selection, weight restrictions and general safety concerns prior to conducting analysis to determine or solve for the actual velocity and impact parameters.

The majority of these recommendations were based solely on assumptions, primarily resulting from personal or anecdotal experience, rather than stemming from analysis of the problem. Supplemental materials were not evaluated for validity or credibility and were used to enforce or strengthen the subjects' assumptions, although one subject did raise the need to determine the final velocity.

Subjects spent roughly half the allotted time discussing safety recommendations, and then attempted to solve the problem to determine a recommendation for the allowable weight limit for the hill. This ultimately led to the subjects' inability to solve the problem, as they were confused with how to set up the problem. Subjects used kinematics as a solution methodology and used force concepts to determine frictional forces. They had difficulty determining how to include the effects of friction in the analysis, which led to increased frustration due to being unable to solve the problem. The subjects eventually could not solve the problem in the remaining time allotted and did not provide any safety recommendations or conclusions from their analysis of the problem.

### **Analysis of Physics Post-test Safety Recommendations Unit**

In this post test subjects based their conclusions primarily on assumptions with minor use of concepts and supplemental information. Subjects used personal and anecdotal experience to formulate conclusions rather than support conclusions with supplemental information. There was little validation or questioning of their assumptive conclusions, and recommendations were given without any reference to the limitations of their knowledge, expertise or any societal implications. The subjects spent the majority on the minutiae of the problem and primarily focused on achieving the proper mathematical solution rather than utilizing the solution methodology to formulate actionable conclusions.

The primary conclusion supported by concepts was the subjects' initial recommendations on the height of the primary and secondary launch hills. The subjects utilized a variety of methodologies to progressively formulate their conclusions, applying kinematic analyses, free body diagrams and force balances, and energy methods. Ultimately, the subjects reached a conclusion determining the height of the launch hills through energy methods and force balances. Subjects did not make any assumptions to simplify the problem, regarding the shape of the loop, and included friction in their solution that appeared to have compounded the difficulty of solving the problem for the subjects.

Subjects used the supplemental information in a limited fashion. Subjects surveyed the information provided to determine the exposure limits of humans to G-forces, and the ASTM standard for restraints. However, there was no questioning the validity or credibility of the information or reconciling of the multiple reported values for G-force tolerance. Subjects simply adopt the information presented in the supplemental information, and do not question its accuracy or contrast it with other available information. They attempt to use the information to form a conclusion but miss several key points that were illustrated in the additional supplemental information and end up basing their conclusion on assumptions from personal experience.

Regarding the restraints, subjects experience difficulty interpreting the graph depicting the standards, which may illustrate an incomplete understanding of the underlying concepts involved in forming a solution.

The majority of the conclusions were primarily based on assumptions that did not have any corroboration from supplemental information or concepts. Subjects formulated conclusions by common sense, intuition or personal experience. A primary example of this was the recommendation for a braking system on the launch hills in order to slow the coaster down due to friction on the track. This conclusion could have been supported by further analysis and solution of the problem. Conversely an argument can be made regarding this conclusion. It can be argued as a conclusion supported by understanding of concepts. Subjects speak about the braking system being able to ensure the coaster has travelled to a sufficient height to gain enough potential energy to complete the return trip.

### **Comparison of Physics Pre-Post**

The physics subjects, serving as the control group displayed considerable improvement between the pre and post-tests. In the pre-tests participants approached the problem in reverse, reviewing the supplemental information and forming recommendations at the beginning of the session, prior to any analysis. These conclusions, as safety recommendations, were primarily formed on assumptions based on their own personal experiences, and not a comprehensive analysis and solution of the key elements of the problem. The subjects eventually attempted an analysis of the scenario, but became confused and lacked a clear plan or method to solve the physics aspect

of the problem. The subjects continued to remain stuck on the analysis past the allotted time, and failed to provide any conclusions or safety recommendations resulting from that analysis.

In this post-test subjects approached the problem similarly to the way displayed by the engineering group in their pre and post-tests. In the posts tests, subjects based their conclusions primarily on assumptions with minor use of concepts and supplemental information. Subjects used personal and anecdotal experience to formulate conclusions rather than support conclusions with supplemental information. There was little validation of questioning of their assumptive conclusions, and recommendations were given without any reference to the limitations of their knowledge, expertise or any societal implications. However, these elements of critical thought were present in their conclusions, which can be viewed as an improvement from their performance on the pre-test despite the relative quality of these elements.

## **Analysis**

During the pre-test, both groups presented recommendations conclusions derived primarily from assumptions based on their personal experience. In the control group pre-test, subjects did not progress to the point where they could provide recommendations from analysis. Interestingly, the control group began by forming potential recommendations at the start of the think aloud exercise. These recommendations were formed solely from assumptions based on personal experience, and did not display any additional elements of critical thinking resulting from analysis. The experimental group did progress through the exercise providing safety recommendations at the end of the think aloud session. The experimental group displayed a reliance on assumptions based on personal experience, similar to that of the control group. However, the experimental group did form some conclusions originating from forming relationships from multiple concepts as well as incorporating supplemental data and information to partially support a recommendation, but ultimately integrated those conclusions into an assumption based on personal experience. Both groups did not question the credibility or validity of the supplemental material, the implications of their recommendations, or display any reflection on their assumptions.

The post-tests for both groups were conducted in the final week of the fall semester, after classes concluded, in order to fully capture any changes that may have occurred due to the subjects fall semester academic and non-academic experiences. Upon analysis, both the control and experimental group displayed an increased use of CTS, when compared to their pre-tests from the beginning of the semester. The control group continued to primarily rely on assumptive personal experiences to form conclusions, but did begin to incorporate conceptual elements into recommendations, utilizing combinations of methodologies and extensions of theory to inform their conclusions. The control group also showed the beginnings of using supplemental material

to support their conclusions, similar to how the engineering group utilized supplemental material in their pre-test.

The experimental group illustrated a reduction in personal experience based assumptive reasoning. The conclusions presented by the engineering group were supported by concepts, information, implications, and evaluation of assumptions. The experimental group utilized a number of concepts in to create a mathematical relationship between the design parameters in order to solve the problem, which could ultimately lead to a descriptive model. Compared to the pre-test, the experimental subjects considered the credibility of supplemental information and used multiple sources to assist their solution and form the basis for safety recommendations. The experimental group also displayed use of new elements of critical thinking not demonstrated during the pre-test, considering the implications of their solutions with respect to potential zoning regulations and questioning their own assumptions and limits of knowledge.

These differences between the groups may be attributed to the MEA and critical thinking instruction present in APSC 100. The use of concepts and information, considering implications and reflective evaluation of assumptions displayed by the experimental group in the post-test was similar to the methodology covered by instruction and model eliciting activities the subjects experienced in APSC 100. The control group, having no explicit critical thinking instruction, displayed increased use of concepts and the beginnings of using supplemental information to inform their conclusions. But, similar to the experimental group pre-test, did not begin to consider the credibility or quality of the supplemental information.

These observed differences may also be attributed to the varying educational backgrounds the different groups may possess, or the differences in individual experiences during the semester. As a means to determine potential reasons, a small exit interview was conducted at the end of the post-test. Subjects were asked if there was anything in their university or life experience that helped them with the thinking in this type of problem.

The control group responded that the pre-test, alongside a similar roller-coaster based question in their classes helped them during the post-test.

[00:52:28]

>> Subject 1: I mean I think really having done as a more problem like the last time doing this, helped us do better this time. Because even if this had been our problem the first time, we wouldn't have been done.

[00:52:39]

>> Subject 2: It's just the thought process.

[00:52:40]

>> Subject 1: It was more than the thought process.

[00:52:41]

>> Subject 2: We did this type of question in class, like last week--

[00:52:44]

>> Subject 3: It was a little bit different, like we had friction instead of a loop, but it was still kind of similar.



Interestingly, in additional follow-up questions regarding their previous experience, the subjects expressed that the end goals of the scenario, to provide safety recommendations were of lesser importance to them than actually solving the problem, and that the supplemental material provided explicit guidelines.

[00:47:05]

>> Facilitator: Did you do that because of an experience you had before, the first time and found that that worked better?

[00:47:11]

>> Subject 2: We did read the material at the beginning but like we picked out the information that we needed. We decided that safety was a concern at the end so we left all the safety material out. Whereas we picked out the specific information like the data of the mass, the height required actuating the physics problem and last time since we found out that the physics problem took longer, sorry, the safety took a lot longer, it was less important. We said that that was for the end.

[00:47:38]

>> Subject 3: Yeah and usually the safety would just say like here's some things you should add into consideration. So it wasn't stuff that we really had to think about. It was just like oh yeah mention this when you're talking about safety. So that's what the materials will tell you so you don't have to spend that much time on them.

[00:47:51]

>> Subject 1: And last time we split our attention into like, you were doing quite a bit of reading, and we were doing trying to figure out the calculations and it just didn't work. It was better to have 3 minds working on it all at once and then doing safety stuff later because that's not as important, but it's the like keep in mind you should add padding and a different [inaudible] than incorporate that into the question.

There was also a question about the supplemental material, about potential recommendations that weren't supported by the materials. The physics group felt that friction was not well represented, which was a particular area they included in their analysis and was a source of uncertainty, and further elaborated on how they provided their conclusions based on personal experience.

[00:48:19]

>> Facilitator: Because I noticed you had a flow of activity. It's not like when you were all focused on the same thing. Were some safety recommendations not supported by the supplemental material ?

[00:48:42]

>> Subject 1: Friction. There wasn't friction in any of this and now it's a little bit difficult to know, like if we had known that like the force between the-- like the friction force between the wheels of the cart and the rail, then that would have changed our second height and it would have changed like the stopping, like the brake power which would have effected like the safety of the necks of the people who are going in whiplash and a bunch of other issues so.

[00:49:16]

>> Subject 3: We actually took a lot of the stuff that we knew about rollercoasters, just like being on them. Just like oh yeah we've experienced brief blackouts so they're fine for a healthy person because they build rollercoasters like that because it didn't explicitly say that in any of these papers. Things like that.

Similar questions were posed to the experimental group at the end of their post-test. Their answers to the questions, illustrate that there was a difference in how the groups viewed supplemental information and how personal experience helped with their solutions.

[01:06:44]

>> So, and I notice that you really used this supplemental material on the-- how did you use it exactly?

[01:06:55]

>> Well, when stuff that we thought we needed worked we had to kind of rely on it entirely.

[01:06:59]

>> Yea, that's true.

[01:07:00]

>> That was-- I remember I didn't use the supplemental information essentially at all with the first one that I did partly because I understood what was going on as far as the science was concerned. But, in this one we pulled essentially everything out of the supplemental information. The restrictions on the height and stuff, the guidelines to G forces--

[01:07:25]

>> Umm hmm.

[01:07:25]

>> Age limits, restraint types.

[01:07:27]

>> Oh, when a lot of those things we wouldn't know kind of--

[01:07:31]

>> Yeah.

[01:07:32]

>> Just out of our-- out of the top of our heads. We would need to research that.

[01:07:37]

>> So, were there some safety recommendations not supported by the supplemental material that you had thought of?

[01:07:51]

>> Not really.

[01:07:52]

>> Yeah.

[01:07:52]

>> I think--

[01:07:53]

>> Pretty much some of it was common sense, but mostly what we kind of pulled from--

[01:07:58]

>> Well, some of it was past experience as well. Like when we realized that we got something that said the initial height should be half the radius or--

[01:08:06]

>> That just didn't make sense to us.

[01:08:09]

>> Yeah.

[01:08:09]

>> Just from like knowledge of a roller coaster.

[01:08:11]

>> It-- yeah generally understanding physics that doesn't make any sense either--

[01:08:15]

>> Yeah.

[01:08:15]

>> Because you're gaining a lot of energy. So--

[01:08:17]

>> There is some common sense--

[01:08:19]

>> There is some background knowledge--

[01:08:20]

>> Just experience with this type of thing, knowing what's going in. So, like roller coasters that go upside down tend to have really beefy restraints and generally the hills are always higher than the loops.

[01:08:38]

>> Have you been on roller coasters?

[01:08:40]

>> Yeah.

[01:08:41]

>> A few.

[01:08:43]

>> Definitely.

[01:08:45]

>> When you were talking about solving this problem could you feel yourself on one or trying to remember.

[01:08:52]

>> Yea, I actually-- I went recently actually, so I was trying to remember what it felt like and kind of what it looked like. And that's kind of like when I was thinking about the age limit restraint and I was thinking I was just on like the newest roller coaster at Wonderland and I was sitting with a bunch of eight year olds. So, I knew that it must not depend greatly on the age limit because you would think that roller coaster-- I don't know how high it is, but it went pretty fast and it was pretty high and had a very steep slope. So, in order for there to be an age limit I think-- how would they even check that too? Would someone be standing there--

[01:09:34]

>> Yeah.

[01:09:35]

>> Where they would have to like going into the roller coast-- going into the park that day you would have to have like a bracelet on that says your age. So, it must not depend that much on age so much.

[01:09:45]

>> I think they use height as just--

[01:09:47]

>> Yeah height as a general.

[01:09:48]

>> It's a lot easier to tell height than age because you can't fake a height.

[01:09:51]

>> Yeah.

[01:09:51]

>> Well you could be, but it would be a lot more difficult.

[01:09:56]

>> So, there's just like past experiences like that that kind of went into our thought process I think.

Subjects were asked if there was anything in their university or life experience that helped them with the thinking in this type of problem. The engineering group responded that they felt the experiences in [course name masked] helped them with this type of thinking.

[01:11:07]

>> And one last question. Is there anything in your university education or life experience at university that helped you with the thinking of this task?

[01:11:18]

>> Well, even, I heard you saying that these sources are credible. And I never really thought about that kind of thing until like [course name masked], sort of evaluating the credibility of our sources. And so, that was definitely in there. Some-- a lot of the physics behind it we kind of learned in high school I think.

[01:11:40]

>> Mmm hmm Yeah.

[01:11:41]

>> But, definitely being able to read a report and take out the necessary information because a lot of the time in high school you're just kind of handed that information. And here we had to find it on our own. And kind of just like figuring out a problem with minimal instructions too I think. Something I know I've learned so far.

[01:12:04]

>> Yea, one of the things I've definitely learned at university is it's when you approach open ended problems like this it's as important to understand what you know and as well as understand what you don't know. And in fact, the stuff you don't know is almost more important than the stuff you know. And you have to realize that and address that and be able to deal with not knowing everything and being able to understand what that means for your results. So, here not being able to know the-- how age relates to dangers from G forces knowing that we know-- well, we can't actually say a number with any credibility at all. So, its more-- it's better to just say there should be an age limit, but--

[01:12:45]

>> But we don't know that.

[01:12:46]

>> But we have no idea of what it should be and same thing with the height. I mean we can specify an exact height and exact radius and say it would work. But it's more useful to them to get ratios that say if your height is this your radius should only be this, which is what we going for ideally with the indirect form, I think that was a little silly, but things aren't often so black and white and that the inequalities we set up are probably more useful than an actual ratio.

[01:13:16]

>> But, it was good to kind of set out that goal at the start because when we're just given a problem like that it's good to think okay how are we going to approach this problem? And that was a great thing that you said at the start was okay let's try to find a ratio--

[01:13:37]

>> Yeah.

[01:13:37]

>> That kind of got us started so because, again in high school they would probably just help you start the problem--

[01:13:46]

>> Yeah.

[01:13:46]

>> Then you would do something--

[01:13:47]

>> The problem would say here's a roller coaster, find the ratio between R and H--

[01:13:51]

>> Yeah.

[01:13:52]

>> As opposed to just here's a roller coaster design it.

[01:13:55]

>> So, yeah that's definitely a big difference between high school and university so far.

[01:14:01]

>> Being able to find your own subsets to problems.

[01:14:04]

>> Yeah.

## Conclusions

The goal of this study is to determine the effectiveness of using Model Eliciting Activities in the curriculum of a first year engineering course on developing critical thinking skills (CTS). The preliminary results of this study, drawn from the coding analysis of the recommendations unit of the think aloud protocol may suggest a difference in the CTS between the experimental and control group. However, due to the limited response of volunteers for the think aloud protocols it is difficult to draw more substantial conclusions.

We believe that from this initial analysis of the think aloud exercises that MEA instruction, alongside critical thinking training in the APSC 100 curriculum, may have the potential to increase student critical thinking ability. The experimental group showed enhanced CTS through increased application of elements of critical thinking in two viewpoints. First, the experimental group illustrated gains in applying elements of critical thinking between their pre and post-test. Second, the experimental group showed an increased application of the elements of critical thinking when compared to the control group. It is these differences that lead us to believe that the specialized curriculum, in the form of MEA instruction in [course name masked], may be responsible for this increase. However, further investigation and analysis of the standardized critical thinking tests (Cornell Level Z & International Critical Thinking Test), analysis of student experience surveys, exploring student background, and further analysis of the remaining units of analysis of the think aloud exercises will be required to draw further conclusions.

## Acknowledgements

The authors acknowledge the support for this research from the Higher Education Quality Council of Ontario (HEQCO), and from the DuPont Canada Chair in Engineering Education Research and Development. This research was approved by the General Research Ethics Board.

## References

- Cornell Critical Thinking Test Level Z. (n.d.). *The Critical Thinking Co - Cornell Critical Thinking Test Level Z*. Retrieved from <http://www.criticalthinking.com/getProductDetails.do?id=05502&code=c>
- Douglas, E. P., Koro-Ljungberg, M., McNeill, N. J., Malcolm, Z. T., & Therriault, D. J. (2012). Moving beyond formulas and fixations: solving open-ended engineering problems. *European Journal of Engineering Education, 37*(6), 627–651. doi:10.1080/03043797.2012.738358
- Frank, B., & Kaupp, J. (2012). Evaluating Integrative Model Eliciting Activities in First Year Engineering. In *Proceedings of the 2012 Canadian Engineering Education Association (CEEA2012) Conference*. Winnipeg, MB.
- Frank, B., Strong, D., Sellens, R., & Clapham, L. (2012). Progress With the Professional Spine: A Four-year Engineering Design and Practice Sequence. Presented at the 2012 International CDIO Conference, Brisbane, Australia.

- International Critical Thinking Essay Test. (n.d.). *International Critical Thinking Essay Test / Critical Thinking*. Retrieved January 7, 2013, from <http://www.criticalthinking.org/store/products/international-critical-thinking-essay-test/185>
- Ku, K. Y. L., & Ho, I. T. (2010). Metacognitive strategies that enhance critical thinking. *Metacognition and Learning, 5*(3), 251–267.
- Leydens, J. A., Moskal, B. M., & Pavelich, M. J. (2004). Qualitative methods used in the assessment of engineering education. *JOURNAL OF ENGINEERING EDUCATION-WASHINGTON-*, *93*(1), 65–72.
- Norris, S. P. (1990). Effect of Eliciting Verbal Reports of Thinking on Critical Thinking Test Performance. *Journal of Educational Measurement, 27*(1), 41–58. doi:10.1111/j.1745-3984.1990.tb00733.x
- Paul, R., & Elder, L. (2006). *A guide for educators to critical thinking competency standards: Standards, principles, performance indicators, and outcomes with a critical thinking master rubric* (Vol. 8). Foundation Critical Thinking.
- Shuman, L. J., Besterfield-Sacre, M., Bursic, K. M., Vidic, N., & Sieworiek, N. (2012). CCLI: MODEL ELICITING ACTIVITIES. In *CD Proceedings: 2012 ASEE Annual Conference and Exposition*. San Antonio, TX.
- Steif, P. S., Lobue, J. M., Kara, L. B., & Fay, A. L. (2010). Improving problem solving performance by inducing talk about salient problem features. *Journal of Engineering Education, 99*(2), 135–142.
- Wiersma, W., & Jurs, S. G. (2005). *Research methods in education*. Allyn and Bacon.