Reasonable or Ridiculous? Engineering Intuition in Simulations

Dr. Elif Miskioglu, Bucknell University

Elif Miskioglu is currently a Visiting Assistant Professor of Chemical Engineering at Bucknell University. She graduated from Ohio State University in 2015 with a PhD in Chemical Engineering, and is interested in student learning in engineering.

Prof. Kaela Mae Martin, Embry-Riddle Aeronautical University, Prescott

Kaela Martin is an Assistant Professor of Aerospace and Mechanical Engineering at Embry-Riddle Aeronautical University, Prescott Campus. She graduated from Purdue University in 2015 with a PhD in Aeronautical and Astronautical Engineering and is interested in increasing classroom engagement and student learning.
Reasonable or Ridiculous? Engineering Intuition in Simulations

Introduction

A successful engineer must not only be proficient in complex calculations, or the simulation software that may perform these calculations, but must be able to evaluate whether a result is “reasonable or ridiculous.” This type of “engineering intuition” is essential, and teaching it is not always as straightforward as technical material.

Often described as a “gut feeling,” intuition is based on a set of rules applied subconsciously.¹⁻³ For complex situations, using intuition results in better decisions⁴ and is a crucial skill in becoming an expert according to the Dreyfus model.⁵ In the Dreyfus model, a novice (contrary to colloquial definitions as inexperienced or new to an area) is defined as someone who blindly follows the rules without examining the results. An expert, on the other hand, uses intuition instead of rules to obtain a solution.⁵ To move from a novice to an expert, one must gain tacit knowledge or knowledge that is difficult to explain or teach.

Technology in the classroom can aid the transition from novice to expert in a variety of ways. Clickers or automatic response systems have been shown to improve student learning through rapid feedback.⁶ Programs such as Learning Catalytics combine automatic response system technology with real time analytics and application, by allowing student response to be immediately tabulated and forming discussion groups among students with differing answers.⁷ It also offers the opportunity for students to ask questions real-time, without having to raise their hand or otherwise speak in class. Similar technologies include Poll Everywhere, Hotseat (Purdue University), and Top Hat.

One question that commonly arises with introducing technology into the classroom is whether the technology promotes deep or shallow learning. First described by Marton and Säljö,⁸ deep learning occurs when students try to understand the material while shallow learning occurs when students simply memorize the material. Some studies have found that including instructional software emphasized lower-level cognitive processes,⁹ but a larger number report learning gains when implementing technology in the classroom through virtual experiments or online instruction.¹⁰⁻¹³ Additionally, incorporating simulations into the classroom can increase visualization and problem-solving processes¹⁴,¹⁵ as well as show positive gains in student self-efficacy with respect to engineering skills.¹⁶

Virtual experiments offer an opportunity to provide students with valuable experience at a low cost (no laboratory space or consumables, only computer facilities, required), high flexibility (can be performed outside of class, does not require direct supervision, safety is not a direct concern), and great breadth (some disciplines may have experiences that are not feasible to provide to students directly, but can be simulated). For classes that traditionally do not include experiments, incorporating virtual experiments results in learning gains over equivalent courses that do not include experiments.¹⁷ For classes that include laboratory portions, virtual labs or simulations can replace actual labs.¹⁸ In some studies, students using virtual simulations outperform those who completed an actual lab.¹⁹
Simulations provide an opportunity to expand the reach of teaching, through engaging different learning styles. Diversifying course delivery methods can address a greater number of learning styles, including those typically neglected by traditional lectures. In one study, thermodynamics students who were slightly more active and visual had improvements in performance indicators. Some simulations have shown a learning style bias; however, often times in simulations, there are no significant differences with respect to learning styles.

As with any tool, simulations must be properly used to be effective. While studies have shown a tendency toward simulations to promote deep learning, there is an increasing cautiousness developing regarding technology’s effect on cognitive processes. In his provocative 2010 book *The Shallows*, Nicholas Carr suggests that the internet is resulting in a “hyperlink happy” society that is losing its ability to think deeply on material. In schools, the influence of our ever-connected society is easily visible in our digital device captivated students, and increasing dependence on search engines, calculators, and other software for answers and solution. It is thus critical that students matriculate from college not only adept in the prerequisite knowledge of their field, but also able to navigate and evaluate the variety of technological tools they are surrounded by. Machine must not only serve as a check for man, but man must also be able to evaluate the accuracy of the machine. Thus, we are interested in better understanding how engineering intuition manifests in undergraduate engineering students, through evaluation of student analyses on simulation problems. At this stage, we seek to understand if certain students, based on GPA, learning style, previous internship experience, or future career aspirations, demonstrate better intuition as measured by success on computer simulation homework problems.

**Methods**

Oftentimes students complete computer simulations without questioning if the assumptions or results are practical. The purpose of this research is to assess how students view assumptions and results when working with simulations of engineering problems and to identify any potential commonalities among students who more often demonstrate successful “engineering intuition” in response to simulated engineering problems. In our context, simulations refer to software-aided problem solution, where the software simulates a real-life process that is infeasible or inconvenient to produce in a classroom setting.

The studied classes come from two small private universities, one in the mid-atlantic and the other in the west. At both institutions, the courses of interest are required junior-level engineering classes that incorporate simulations into the classroom. The study began Fall 2015, and for that semester the class from the mid-atlantic university was chemical engineering course on separations with a total of 29 students. The course from the western university was two sections of a required aerospace engineering class on orbital mechanics. These two sections had a total of 67 students which were split into sections of 33 and 34. Required junior-level courses were chosen for this study since by this point in the curriculum students have had a few opportunities to work on open-ended questions possibly pushing them further along in their cognitive development. Further, the junior year is a key juncture in the transition from novice to expert, as students receive the last of their technical preparations before entering their senior capstone design coursework. As stated by Wankat and Oreovicz, simulations with commercial simulators increase course diversity and help push students into high levels of development such
as early or higher multiplicity in Perry’s theory of development where students admit that multiple answers can occur and instructors might not know the answer (a common premise of capstone design courses).

Students were given a survey to assess internship experience, ideal job, and the Felder-Soloman Index of Learning Styles (ILS). The learning styles assessment is an open source online tool with 44 questions that determines student’s learning style using the Felder-Silverman model of learning styles. The Felder-Silverman model is a multidimensional characterization of learning style preferences that attempts to qualify how individuals prefer to perceive, receive, manage, and interpret information. Each of the four dimensions contain two opposing preferences (active/reflective, sensing/intuitive, visual/verbal, sequential/global) and are evaluated on a scale from “strong” to “balanced” for a preference within each dimension. The Felder-Silverman model and ILS have undergone several validity and reliability studies. The ILS has been shown to have both good validity (construct and content) and reliability (measured as test-retest and comparison with self-evaluations). The ILS has ranked better than other popular models, including the Honey-Mumford and Kolb models, when looking at participant agreement with results. While the ILS has demonstrated good reliability and validity through the twenty years it has been used, there are precautions that must be taken when using any learning style model, including the Felder-Silverman, in the classroom. It is imperative to keep in mind that learning styles are preferences, not absolutes, and should not be used to brand students, predict their success, or prescribe career outcomes. Learning styles models offer a tool for better understanding our student populations, and developing improved teaching materials that allow us to develop more diverse thinkers.

The simulation tool used in the aerospace engineering course was General Mission Analysis Tool (GMAT), an open source trajectory optimization and mission analysis tool developed by NASA and private industry. Since its inception in 2007, GMAT has gone through multiple versions and been used by the research and industrial community. In addition to the survey in the aerospace engineering course, three of the ten required simulation homework assignments included an additional question on the assumptions/feasibility of results of the simulation problem. The three questions were:

1. If the goal of the mission is to do a lunar fly-by and someone proposed this mission, would you approve it? Why or why not?
2. If you were planning this trajectory, would you be worried about the lifetime of the spacecraft? Why or why not?
3. Would you fly this mission? Why or why not?

In the chemical engineering course, the AspenTech programs HYSYS and Aspen Properties were used for simulations. Aspen Properties is a chemical property database that allows users to look up thermodynamic information for chemicals and chemical systems. HYSYS is the simulation software that allows users to simulate a chemical plant or process. HYSYS is widely used in industry, and in chemical engineering senior design courses throughout academia. Because of technical difficulties throughout the term, only one question was asked in the chemical engineering course, near the end of term:
1. If your goal was to separate chemical A (specific chemical included) from chemical B, which column design would you recommend be built? Why?

The simulation questions were chosen to take a homework assignment and transform it into a potential real-life scenario. These open-ended questions required students to give an opinion about their solution and sometimes forced the students to search for information that was not presented in class. Solutions were not always obvious or straightforward, but in every case students did have the tools to arrive at the correct answer/recommendation. That is, no question was beyond the scope of their existing knowledge, but the questions did require an application of their knowledge that was likely less straightforward than they were accustomed to.

The questions were evaluated on the yes/no response or recommendation (for accuracy) and the validity of the reasoning behind the choice. The “correctness” of the questions (yes/no response) was marked as no answer, correct, or incorrect. If the question response was correct, the reasoning to that answer was determined to be either correct or incorrect. For the remainder of the paper, “correct answer” or “correctness” refers to the yes/no response while “correct reasoning” or “reasoning” refers to the rationalization of that choice.

The responses from students who gave consent were collected, compiled, de-identified and coded, and analyzed for common responses, misperceptions, and correlations with learning styles, internship experience, degree track, and overall homework and course grade. All statistical analyses were done using SPSS or JMP software. Categorical variables (e.g., success on problem) were analyzed using Fisher’s Exact Test as many contingency tables contained cells with counts of less than 5. Continuous numerical variables (such as final numerical grade percentage) were analyzed using non-parametric one-way ANOVA or t-test.

Results and Discussion

All three simulation problems in the aerospace engineering course showed statistical significance with respect to at least one study factor. Problems will be described as “problem 1,” “problem 2,” and “problem 3,” with the number corresponding to the additional questions used with each as described in Methods.

The data from the chemical engineering course was not as conclusive. No conclusions were able to be drawn due to the smaller sample size and the number of questions. Thus, we seek more information through continued study of the same population in Spring 2016 to be able to provide more conclusive results and compare the two disciplines.

Class Composition Comparison: Learning Styles

Even with some inconclusive data from the chemical engineering course, we can discuss learning style distributions across the two majors from the classes in question. Learning styles are one of our factors of interest in understanding student success on simulation problems, and thus, a comparison of learning styles profiles between the two classes is useful in better characterizing our populations and their potential differences. As shown in Figure 1, the aerospace engineering students (N = 67) show a balanced preference in most dimensions, except visual/verbal in which most students are visual. Previous work as shown similar distributions. On the other hand, the chemical engineers (N = 29) show a slight predominance of active, sensing, visual, and
sequential preferences in the respective dimensions. This result agrees with early work looking at learning style preferences across numerous large institutions.\textsuperscript{29} The difference between the aerospace and chemical engineering student populations with respect to learning styles profiles may be a side effect of curricular differences. The aerospace students are mixed (multiple years, multiple focuses) whereas the chemical engineers are all of the same cohort. The chemical engineering students have taken their core courses together for two years at this point, and this exposure to the same learning experiences may have allowed them to develop into a population with stronger preferences.

Factors: Grades (Homework and Course), GPA

With a significance level of 0.05, a variety of interesting results were observed for the aerospace engineering course (N = 54) with respect to overall homework grade (total of 10 homework assignments). The overall homework grade showed statistical significance (Fisher’s Exact Test) with respect to the reasoning for problems 1 (N = 39) and 2 (N = 23) but not problem 3 (N = 34). For problem 3, only one of the 34 participants who had the correct answer had incorrect reasoning, making a distinction between correct and incorrect reasoning impossible. However, for this problem, overall homework grade was a significant factor for the correct answer which was also seen in problem 2. Essentially, students who did well on the homework assignments often answered the problems correctly and had the correct reasoning.

In any educational study, it is important to address potential confounding factors. For this reason, we analyzed data for possible correlations between and beyond the factors of interest. Among these, we discovered that the aerospace engineering course had a strong correlation between homework grade (mean 80, standard deviation 12) and overall grade (mean 74, standard deviation 10) as seen in Figure 2 (matched pairs t-test, N = 54, \( \alpha = 0.05 \)). This result is not surprising, since a large portion of the final grade came from the homework assignments (40%). The same statistical significance trends (\( \alpha = 0.05 \)) between homework grade and problem reasoning or problem correctness for problems 2 and 3 discussed above also held for overall course grade. However, there was not a statistical significance between correct reasoning and overall homework grade for problem 1 (N = 35), which is unexpected given the relationship between homework and final grade. Because correct reasoning only considers the subset of the population that got the answer correct, this suggests that for problem 1, the subset of the population that had the correct answer does not show as strong a correlation between homework grade and overall course grade when considering the sample size.

GPA and final grade or final homework grade (N = 54) were not closely correlated (matched pairs t-test, \( \alpha = 0.05 \)), thus GPA was only a statistically significant factor in student responses to Problem 2 (Fisher’s Exact Test, N = 61). As shown in Figure 3 students with GPAs lower than 3.0 had a much lower frequency of answering the problem, and students with GPAs of greater than 3.5 largely provided the correct answer. Students with GPA between 3.0 and 3.5 left the problem blank, provided the correct answer, or answered incorrectly in about equal numbers. The GPA correlation did not hold as strongly for correct reasoning (N = 23) behind the answer on this problem, but still more than 50% of the students who had a GPA above 3.5 (and had answered the question) provided the correct answer, whereas less than 50% of the students with GPAs below 3.5 did.
Figure 1. Learning style profiles for Fall 2015 aerospace (a) and chemical engineering (b) courses studied. In most dimensions, the population exhibits a “balanced” preference for the aerospace students, except in visual/verbal where the large majority of students are visual. For the remaining dimensions, among students that have a preference, active, sensing, and sequential are seen in greater numbers. For the chemical engineers, active, sensing, visual, and sequential are the more common preferences in each dimension.
Figure 2. Scatterplot of final grade vs. homework score (both in percent) of aerospace engineering class showing largely linear relationship and indicating that factors may be confounded among population.

Figure 3. Student responses to problem 2 (correct answer) based on GPA. Students with GPA lower than 3.0 were more likely to leave the question unanswered, whereas students with GPA higher than 3.5 were more likely to answer and get the question correct.
It is highly interesting that problem 2 is the only one that showed this GPA correlation. This may be a result of the greater complexity in problem 2, as compared to 1 and 3. This result is evident even without looking at the problem itself, but just the question asked about the problem (see Methods). Problems 1 and 3 are direct inquiries (with justification) as to whether the student would recommend moving forward with the scenario (mission). Problem 2, on the other hand, is a more complex question that required the students to use additional resources and perform an additional calculation to properly assess. This additional step was not beyond the scope of their knowledge at this point in their education, however, it did require them to make connections that may not have been explicit in the problem or question statement.

The correlation between GPA and the responses to problem 2 suggest that when presented with a complex and integrated problem, students who are generally high achievers may be more inclined to take the extra steps necessary toward the solution. We may postulate that the students with GPA lower than 3.0 were not motivated to attempt the problem, or attempted the problem but were unsuccessful or dissatisfied with their solution, and thus did not turn in an answer. They did not even attempt to guess the first portion of the question by providing a simple yes or no, where they had a 50/50 probability of getting it correct. Students with GPA greater than 3.5 mostly got the question correct, showing that they either knew or were willing to seek help to figure out the additional steps needed to achieve a solution, and were willing to then do the subsequent work.

**Factors: Section, Internship, Military Experience**

Because there were two sections of the aerospace course, we were interested in whether there were any differences between the sections (one-way ANOVA, $\alpha = 0.05$). The only statistical significant difference (with attention given to possible confounding factors) between sections and answer correctness and reasoning occurs for the second problem ($N = 61$). We believe this statistical significance may have occurred for two potential reasons. The second section was taught later in the day, and a student in this section asked a question about how to answer problem 2 while there was no such question from the first section. While the instructor did not outright answer the question, students had a better understanding of how to answer the question if they were paying attention in class.

Problems 1 and 2 showed a statistical significance (Fisher’s Exact Test, $\alpha = 0.05, N = 61$) with respect to the factor of student internship experience. One third of the students had internship experience which were coded by discipline (Aero, Astro, Electrical, Mechanical). On both of these problems, we see students with internships in electrical engineering and aero as the top performers. However, it is important to note that the category of electrical engineering internships contained the top two performers in the class overall, and this may be a confounding variable in that respect. Furthermore, the students coded with aero internships were more likely to be seniors while the astro students were juniors. With the internship experience coded as described above, there was statistical significance between the correct reasoning for problem 1 and internship ($N = 39$). Interestingly, there was no correlation between internship experience and GPA.

If the significance level is raised to 0.10, problem 2 has a statistical significance between correct reasoning and internship experience (Fisher’s Exact Test, $N = 23$). Only one third of those with
Internship experience got this problem right, so if the students had the correct answer, the internship experience benefitted them in finding the correct reasoning. Continuing with a significance level of 0.10, the only problem to show a statistical significance between correct answer and internship experience (Fisher’s Exact Test, N = 61) was the first problem when the internship codes were not reduced to four categories but kept at an “expanded version” of seven. The third problem shows no link between internship time or experience and correct answer and reasoning, but the third problem was the only problem which showed a statistical significance at 0.10 between the correct answer and military experience (Fisher’s Exact Test, N = 61). Of the 5 students identified with prior military experience, however, only two answered the question (and correctly) compared to the 75% of the non-military students (N = 56) who answered (of which three-quarters answered correctly). While the Fisher’s Exact test ran for these is considered valid for low sample sizes, it is always important to look at the raw data and determine whether reported significances show meaningful differences. In the case of military experience and problem 3, we are not convinced what we see is a true correlation. Additionally, more caution must be used when applying these results due to the higher significance level.

**Factor: Learning Style Preferences**

In terms of learning style preference, there was a statistical significance ($\alpha = 0.05$) between correct reasoning in problem 1 (Fisher’s Exact Test, N = 39) and the active and reflective scale. Students were placed in active, balanced, or reflective categories. Active students were more likely to get the answer correct, reflective more likely to get wrong, and balanced were equally right and no answer. This result was the only one at a significance level of 0.05, but if the significance level was increased to 0.10, there was a statistical significance between the sequential and global strengths and correct reasoning in problem 3 (N = 34). Done by a Fisher’s-exact test on the 5 possible preferences when considering strength (strong sequential, moderate sequential, balanced, moderate global, strong global), strong preferences were less likely to answer or be correct (N = 59). Working through the problem required laid out, sequential, steps, but evaluating the scenario favored the global perspective. It is possible that the problem contained biases towards students with less strong preferences, who are potentially better able to engage with both learning styles simultaneously.

Thus, it appears that overall, there were little correlation between learning style and success on problems, as has been seen in others’ work. However, we did observe that intuitors were less likely to have any internship experience. The intuitors were half juniors and half seniors compared to the rest of the class which was one-third juniors and two-thirds seniors. Since juniors are less likely to have an internship experience than seniors, this relationship may confound the observed intuitor-internship trend.

**Conclusions and Future Work**

In better understanding student assessments of simulation problems, we see after one semester of study that potential important factors include internship and overall GPA. Students with internship experience were more likely to provide correct reasoning, suggesting that this “real world” experience may have aided them in the classroom. The GPA trend shows that students with low overall GPA were less likely to even answer the problem, and students with high GPA were much more likely to answer and do so correctly. This result is especially interesting
because the problem that showed statistical significance with respect to this factor was the only problem to require students to take extra measures by performing additional calculations, not just report on feasibility of a scenario after having done the initial problem calculation. It is possible that overall high achievers (students with high GPA, who show high success in classes across their curriculum) are more motivated to pursue more challenging problems. As there was no correlation between GPA and internship experience, we are not concerned about these observations being confounded.

With respect to learning styles, we see little correlation with success in assessing simulations. The differences between the student populations in aerospace and chemical engineering are interesting, especially as both populations agree with differing literature. While learning styles are an interesting tool for characterizing class profiles, they may not offer us much insight into the factors that promote development of engineering intuition, and specifically development of engineering intuition for simulation applications.

In future work, we intend to continue this study along all avenues to gather more data, especially on the chemical engineering population. This will allow for comparisons across institutions and disciplines, and better evaluate what factors to focus further, more concerted, efforts on. As of now, the GPA trend strikes us as most interesting, and we will pursue investigation of that further by incorporating more such “stretch” problems (problems that require additional steps to evaluate feasibility of a scenario, rather than a judgement on the results already computed) into our study and courses. In time, this will lead to more conclusive discussion on the factors that influence student success in assessing simulations, and improved strategies for teaching students “engineering intuition.”

References
2 Gigerenzer, G. Short cuts to better decision making. (Penguin, 2007).
3 Kahneman, D. Thinking, fast and slow. (Farrar, Strauss, and Giroux, 2011).
5 Dreyfus, S. E. & Dreyfus, H. L. A Five-Stage Model of the Mental Activities Involved in Directed Skill Acquisition (A155480). (1980).


Brown, A. O. *et al.* in *ASEE Annual Conference & Exposition*. (Indianapolis, IN, 2014).


Soloman, B. A. & Felder, R. M. Index of Learning Styles Questionnaire.


Miskioglu, E. in *ASEE Annual Conference and Exposition* (Seattle, WA, 2015).