USING HUMOR IN THE STEM CLASSROOM TO ENHANCE KNOWLEDGE TRANSFER

P. Ludovice, D. MacNair

Georgia Institute of Technology, Atlanta (UNITED STATES)

Abstract

We hypothesized that humor is most effective in science, technology, engineering and math (STEM) curricula when it is integrated into the educational lesson. This hypothesis is derived from previous work suggesting that humor is distracting when integrated in time (temporally), rather than integrated into the lesson content. We suggest this occurs because temporally integrated humor increases the extraneous cognitive load of the student, thereby reducing the germane cognitive load that facilitates learning. Here we also test the hypothesis that humorous examples similar in structure (isomorphic) to the relevant technical concept will facilitate knowledge transfer from relatable humorous examples to the more challenging technical concept. This was based on previous work that indicated knowledge transfer between isomorphic examples is more likely. Our results, from students in a second year introductory chemical engineering class, showed that germane cognitive load increased with the use of a humorous relatable example over a control class with a typical technical example. This humorous example was isomorphic with the chemical engineering lesson, and we observed a correlation between improved academic outcomes and increased germane cognitive load in these students. These results suggest that STEM educational outcomes may be improved by using relatable humorous examples, as long as the humorous example is integrated into the STEM lesson.

Keywords: humour, STEM, cognitive load, knowledge transfer, seductive details.

1 INTRODUCTION

Humor has been associated with various positive effects in education including an improvement in instructor immediacy, and educational satisfaction.[1-4] Humor has also been shown to increase creativity,[5] and has been shown to be effective in technical innovation.[6,7] It is difficult to see the connection between these improvements and educational outcomes. For example, humor has been associated with reduction in test anxiety and stress, but this does not regularly translate into improved performance.[8]

Previous work indicates that simply adding humor to technical education does not improve educational outcomes.[9,10] Added humor might distract from learning just as extraneous details can, as seen in the seductive detail effect.[11] Previously, we have observed that engineering students are not particularly aware of such temporally integrated humor.[12] We contend that temporally integrated humor increases the extraneous cognitive load of the student thereby reducing the germane cognitive load that contributes to learning as seen in Figure 1a.[12] Assuming the total cognitive load is fixed, extraneous cognitive load reduces the amount of germane cognitive load that may be focused on the educational challenge. Germane cognitive load is the mental activity associated with learning, and we contend that germane cognitive load will increase when a student draws an analogy between the concept to be learned, and a humorous/relatable context or example. By integrating the humor into the problem, we reduce the extraneous cognitive load and increase the germane cognitive load as seen in Figure 1b. Kalyuga contends that increased germane cognitive load can improve learning outcomes.[13]

By using a humorous/relatable example that is familiar and less intimidating to the student, we contend the lesson will be easier to learn. Many of the concepts in a STEM curriculum already have a high intrinsic cognitive load because the concepts are very complex. By replacing the technical lesson or concept with a humorous/relatable analogue, we add the humor to the intrinsic cognitive load and avoid adding extraneous cognitive load in the form of humor. This increases the important germane cognitive load fraction to improve learning outcomes. However, this approach is only useful if the knowledge derived from the humorous/relatable example can be transferred to the relevant technical example. We hypothesize that making the integrated humorous example similar in structure to the technical problem (isomorphic) will accomplish transfer of learning from the humorous/relatable problem to the actual technical problem.
Bassok and Holyoak suggest that knowledge transfer is more likely to occur between educational examples and technical problems if the example and problem are isomorphic, or similar in structure.[14] Gass and Priest found similar value in isomorphic examples for producing learning transfer in outdoor education.[15] Therefore, we will test a technical education example that uses integrated humor in an isomorphic example. The integration of the humor component into the educational example prevents the undesirable increase in extraneous cognitive load, and the isomorphic nature of the example helps facilitate knowledge transfer from the more relatable humorous example to a relevant technical one. Previously studied examples of integrated humor in the STEM curriculum were relatively isomorphic with the desired engineering example,[12,16] however, this is the first analysis of an integrated isomorphic example with a control.

![Figure 1. Schematic of the integrated humour approach to the total cognitive load. By choosing a problem topic that is already humorous, the humour is integrated into the intrinsic cognitive load of the problem (b.), rather than adding humour to the extraneous cognitive load (a.).](image)

We generate humor by juxtaposing a technical topic or procedure with a non-technical one. This juxtaposition produces the integrated humor that we hypothesize is critical to minimizing the extraneous cognitive load associated with temporally integrated humor.[12] Such integration can also potentially reduce the seductive detail effect. Juxtaposing these technical and non-technical ideas produces an incongruity, and truth or partial similarity results if the two topics are isomorphic. The combination of an incongruity with a partial truth or similarity is the essence of humor.[17,18] This is achieved two ways: (i) by applying a technical theory or analysis procedure to a common phenomenon, or (ii) by representing a complex technical principle with a non-technical analogy. The first approach was used for most of the interventions described elsewhere,[16] but the second approach was used here in a specific intervention to explain the equilibrium of two liquid phases used by chemical engineers to analyze the process of liquid-liquid extraction (LLE).

An additional advantage of using the aforementioned juxtaposition is to make it easier for technical instructors in the STEM fields to effectively produce humor for the educational interventions described here. Many instructors are intimidated about writing humor, but it is straightforward to combine a technical concept with a general non-technical concept. The non-technical concept should be common and well known so the students can relate to it. Therefore, we will often refer to this as a humorous/relatable concept to dispel the myth that this concept must be inherently humorous. It certainly can be humorous, but the humor is typically generated simply by the juxtaposition of the two concepts. Berk suggests choosing well-known examples of popular culture, and this can serve as this non-technical element.[19]
2 METHODOLOGY

2.1 Educational Venue

Both an experimental and control lesson were taught in a chemical engineering mass and energy balance class (CHBE2100 in Fall semester 2016) at the Georgia Institute of Technology in Atlanta, GA in the southeastern United States. The Georgia Institute of Technology’s most popular programs of study are in engineering, and each of its engineering programs are ranked in the top 10 in the 2019 U.S. News & World Report rankings for graduate engineering. CHBE2100 is required for all chemical engineering students, and is typically taken in the first semester of their second year. The experimental and control sections were applied in two different sections of the class taught by the same instructor with over ten years of experience teaching this particular class (P. Ludovice).

This research was carried out under Institute Review Board Protocol H13322 for human subjects research. The cognitive load instrument was adapted for this and other classes, and could be taken in class on a laptop computer or a smart phone. Cognitive load surveys and their associated informed consent were handled by D. MacNair to avoid any conflicts with P. Ludovice who taught the class. In addition to the cognitive load measurements, class exams were used analyze learning outcomes relative to the cognitive load results. This work focused on a single example to teach students the phase behavior of Liquid Liquid Extraction (LLC). LLC is typically taught in this class, but involves some challenging concepts. It consists of adding a third liquid solvent to a mixture of two or more solvents to extract one of the components of the mixture. As such, it involves a complex triangular diagram that describes the phases of three liquid components. This diagram, called a ternary phase diagram, is trianulargly shaped in seen in Figure 3.

This class was taught in the Fall Semester of 2016, and the two sections were offered in the same lecture hall at 9:05-9:55am, and 10:05-10:55am on Monday, Wednesday and Friday for a 15-week semester. The early class was the experimental section, and the control section was the later class. This particular lecture hall did have theater style seating which hampered the students’ ability to engage in active learning while working the examples in class. In this class, students read the relevant material in the text before engaging in group activities to apply mass and energy balance solution techniques in class. After class, they applied these techniques to weekly on-line homework assignments. In both classes, they took unannounced quizzes on the reading material to motivate them to prepare for class. In addition to a final exam, the students also took three exams during class. A question from the second exam were used to analyze educational outcomes for both control and experimental sections.

2.2 Cognitive Load Measurement

Cognitive load theory separates the cognitive load required for learning into three components: (i) intrinsic, (ii) germane, and (iii) extraneous.[20] The intrinsic component describes the cognitive load required to understand the problem at hand. This should be constant regardless of the educational intervention used. The germane cognitive load is that required to process and understand the educational lesson resulting in a positive learning outcome. Germane load formulates patterns, connections and generalizations that bring about learning. The extraneous cognitive load is any additional cognitive processing related to understanding any extraneous details of the problem. We will utilize the original three-component version of cognitive load theory despite Kalyuga and co-workers’ claim that germane cognitive load is not independent, and is simply a form of intrinsic cognitive load.[21] We contend that, for complex technical lessons, the mental effort required to understand the complex problem framework, interactions and variables (intrinsic load) is independent of the germane load. This may not be the case for simpler non-technical problems.

The aforementioned cognitive load test instrument was originally based on a generalized instrument developed previously.[22] This instrument is a ten-question survey that measures the three types of cognitive load. It was tested on Ph.D. psychology students after a statistics lecture. Principal component analysis of the results showed three distinct components that were consistent with the three types of cognitive load. Additional confirmation was obtained from a confirmatory factor analysis of statistics lectures for three larger undergraduate classes.

This instrument was specifically designed for a statistics lecture. However, it was adapted for a computer science class by simply changing a few words from those that referred to statistics to wording that referred to computer programming in question eight of the ten-question test instrument.[23] Application of this survey instrument to large undergraduate programming classes reproduced the same three factor results from a confirmatory factor analysis. We simply made minor changes to a few words to reflect our
interest in specific lessons mass & energy balances in contrast to statistics or computer programming. Most of the ten questions made general reference to the topics and concepts (intrinsic load), the instructions and explanations (extraneous load), and the knowledge and understanding achieved (germane load). Therefore, only a minor change to question eight was required for these adaptations. The reference to programming was changed to “mass & energy balances” for application to a basic chemical engineering mass and energy balance class, as see in the bold text in Figure 2.

1 The activity covered concepts that I perceived as very complex.
2 The activity covered concepts that I perceived as very complex.
3 The activity covered concepts and definitions that I perceived as very complex.
4 The instructions and explanations during this activity were very unclear.
5 The instructions and explanations were, in terms of learning, very ineffective.
6 The instructions and explanations during this activity were full of unclear language.
7 The activity really enhanced my understanding of the topic(s) covered.
8 The activity really enhanced my knowledge and understanding of mass & energy balances.
9 The activity really enhanced my understanding of the concepts covered.
10 The activity really enhanced my understanding of the concepts and definitions.

Figure 2. Questions in the cognitive load measurement survey used in this study. Answers were collected on an eleven point Likert scale from 0 to 10. The intrinsic, extraneous, and germane loads were measured by questions 1-3, 4-6, 7-10 respectively, and the bold text was modified to fit our mass and energy balance class.

2.3 Educational Intervention

The engineering lesson involved the understanding and use of a ternary phase diagram. This is a triangular diagram that describes the phases that result when mixing three liquid components together (see the left side of Figure 3). To examine the educational outcomes we used the 9:35am class as the control group, and the 8am class as the experimental group. As discussed above, the control group was taught this concept using the traditional ternary phase diagram whose three corners represent pure components of liquid A, B, and C as seen on the left side of Figure 3. After each of these different interventions, the students went on to solve related homework and test problems involving LLE using a real ternary phase diagram of water, acetone and methyl isobutyl ketone from the text book.

The two-phase region on the left side of Figure 3 lies at the bottom of the triangular diagram indicating that component B is the compatibilizing solvent. Adding sufficient amounts of component B to a mixture of A and C will cause the A-rich and B-rich phases to dissolve in each other. The tie lines (bold diagonal lines) in Figure 3 indicate the composition of the phases that separate in the two-phase region. If the overall composition of the mixture resides in the two-phase (grey shaded) region the liquids will separate into two phases located at the ends of these tie lines. As indicated in Figure 3, the C-rich phase has a higher concentration of component B than the A-rich phase based on the end of a particular tie line. This composition distribution is determined by specific molecular interaction between the solvents, the details of which are difficult to grasp even for chemical engineering students.

To address the intellectually challenging nature of these ternary phase diagrams and the associated LLE process, we apply integrated humor to help make this concept easier to understand. We use the analogy of a middle school dance to make this more relatable (right side of Figure 3). In the U.S., middle school typically contains students between 11 and 14 years or age. The fact that the two-phase region (grey shaded region in Figure 3) occurs when few adult chaperones are present represent the common observation that young people tend to segregate into dominant gender groups because they find mixed gender groups awkward at this age. Here, adult chaperones represent the effective compatibilizing
solvent because they are aware of this awkward phase and encourage boys and girls to talk with each other. The tie lines represent the common belief that boys are more likely to engage in mischief than girls, so more adult chaperones cluster with the group that is predominantly boys. This generalization is based on common experiences that are more relatable than the specific molecular interactions of chemicals that determine this phase diagram.

To make the analogy even more relatable to current university students, it was mentioned in class that the slope of the tie lines on the right side of Figure 3 might change if the dance was changed to on-line social media interaction. The tie lines might reverse their slope because adults are more concerned about girls than boys on social media. This was based on recent media reports of increased involvement of girls in cyberbullying, but there is data to suggest this trend. The class found this analogy very funny based on a pronounced laughter response. No claim of whether or not this trend in cyberbullying is actually true, but the analogy appeared to be understood by students in the classroom. While not a part of this class, this analogy is potentially relevant to more advanced LLE topics. Such changes in the two-phase region and the slope of the tie lines can occur due to changes in system temperature or pressure.

Figure 3. Ternary phase diagrams used to introduce the concept of Liquid Liquid Extraction (LLE) in the control (left) and experimental (right) classes. The apex corresponds to 100% of the component labelled at that apex. The shaded area indicates two phases with the composition at the ends of the lines (tie line) in the shaded area.

The figures above were used in class to explain how traversing this ternary phase diagram produces different phases. However, a ternary phase diagram from the text, where the A, B and C components were labeled with specific chemicals, was used in both the homework assignment, and the second exam that covered this material. Following the introduction to this concept, a sample problem was done in class the involved the determination of the conditions under which a single phase is obtained, and a mass balance on a LLE process. Question two from exam two was used to determine the educational outcome from this educational intervention. The conditions under which a single phase is obtained and an LLE process mass balance were addressed in part B and part A of question two respectively.

3 RESULTS

3.1 Cognitive Load Results

Figure 4 illustrates the results of the cognitive load survey given the week following the in-class lesson using the middle school chaperon analogy (experimental intervention) in the early class, and the standard chemical component approach (control) in the later class. As expected, the intrinsic cognitive load was essentially identical for the two cases because the lesson and the problems used to practice the lesson where identical except for the use of the analogy in the experimental intervention. Because the control case had no additional jokes we expect the observed result that extraneous cognitive load did not change between the control and experimental cases. However, the total cognitive load was not
conserved. This could occur because our survey instrument did not properly measure these cognitive load components, or that these components are not truly independent.[21]

This value of approximately 5 for the intrinsic load average is relatively high based on a comparison with other examples from a chemical engineering numerical methods class.[16] When the intrinsic cognitive load is slightly higher, humorous interventions appear to fail because they cannot overcome such a high intrinsic cognitive load. Hopefully, this clarification of this rather challenging concept produced by the integrated humorous example will translate into improved educational outcomes. We examine this in the section below using the outcome of class assessments from this mass and energy balance class from 2016.

![Cognitive Load Survey](image)

Figure 4. Mean values cognitive load survey for the three cognitive load components for the interventions used in the mass and energy balance class, LLE (N=42), school (N=30). The error bars are the 90% confidence intervals about the mean. The * symbol indicates germane load is statistically different from intrinsic load (p < 0.05).

3.2 Learning Outcome

Exam two in this mass and energy balance class covered the LLE material from the class. Question two on this exam had two parts. Part A (worth 20 out of 35 total points for question two) was a mass balance on a LLE process, and part B (worth 15 out of 35 total points) used the ternary phase diagram to determine the phase of a resulting mixture of two different solutions. It is this phase determination that was isomorphic with the middle school analogy used in the classroom lesson on LLE. This isomorphism can be seen in Figure 3, as the diagrams are literally the same shape. Part A depended more on generic mass balance skills that were not addressed by the educational intervention. Therefore, we hypothesize that part B of this question might have a better average score for the experimental group. Figure 5 shows that there was not a statistically different average score any component of question two between the experimental and control cases. In Figure 5, question two is simply the sum of the scores from the two parts of the question. Because part A was a general mass balance on the LLE process, and was not addressed by the experimental intervention, we do not expect to see a difference for this part of the question. However, we did hypothesize the part B would have a higher score for the class with the experimental educational intervention that used the middle school chaperone analogy.

Despite the lack of statistical difference, it is discouraging that the control group showed higher scores in all cases. It may be that this lack of difference is due to other confounding variables. We hypothesized that an increase in germane cognitive load should contribute to improved educational outcomes, and
calculated the correlation between the components of the cognitive load test results and educational outcomes. This Pearson correlation between the scores on question two and the cognitive load scores, along with their associated statistical significance are seen in Table 1.

**Table 1.** Pearson correlation (r) between the cognitive load results in Figure 4 and the performance in question 2 from exam 2 on ternary phase diagrams and their statistical significance (p). The grey-shaded element is the only statistically significant correlation (p<0.025).

<table>
<thead>
<tr>
<th>Cognitive Load</th>
<th>Experimental (“school” intervention)</th>
<th>Control (“LLE” intervention)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>question 2A</td>
<td>question 2B</td>
</tr>
<tr>
<td>intrinsic</td>
<td>r=0.1554</td>
<td>r=0.1641</td>
</tr>
<tr>
<td></td>
<td>p=0.4068</td>
<td>p=0.3778</td>
</tr>
<tr>
<td>extraneous</td>
<td>r=0.1399</td>
<td>r=0.2447</td>
</tr>
<tr>
<td></td>
<td>p=0.4527</td>
<td>p=0.1791</td>
</tr>
<tr>
<td>germane</td>
<td>r=0.0123</td>
<td>r=0.4115</td>
</tr>
<tr>
<td></td>
<td>p=0.9487</td>
<td>p=0.0224</td>
</tr>
</tbody>
</table>

Among the results in Table 1, only one correlation is statistically significant (grey-shaded part of Table 1). That is the correlation between the performance of the students in the class with the experimental educational intervention on part B of question two, and the germane cognitive load result for those same students. This correlation is positive which indicates that students with a higher germane cognitive load score performed better on part B of this question that tested their ability to understand how phases change in this ternary phase diagram. While the same correlation for the control group was also positive, it was not statistically significant. Not surprisingly, the correlations for both control and experimental groups were negative between that test score and the intrinsic and extraneous cognitive load.

The results from Figure 4 and Table 1 provide support for our hypothesis that humour integrated into an educational intervention increases germane cognitive load, without increasing extraneous cognitive load, and this is associated with improved educational outcomes. However, these results are inconsistent with the data seen in Figure 5. We suggest that this inconsistency is due to confounding variables. Initially we hypothesized that the early class was not as academically qualified as later class (control group). We compared various averages of academic metrics including the grade in the class, university grade point average (GPA), pre-university GPA, and scores for all three sections of the Scholastic Aptitude Test (SAT) taken by U.S. students for college admissions. While all of these metrics...
were higher for the control class, none of the differences were statistically significant. This suggests that there may be a slight academic difference between the two classes, but other variables also play an important factor. These other variables include that fact that the students may be more attentive in the control class which as later in the day, and the fact that the instructor (P. Ludovice) may be more effective presenting the lesson for the second time that day. The students studied here are assigned class registration times based on seniority, and students with more pre-university class credit typically choose later classes. This additional class credit may correlate with academic performance. While it is often assumed that university students perform better in later classes, this is a complicated issue. One study of class start times showed that alcohol consumption was correlated with later classes and can actually decrease academic performance.[25]

4 CONCLUSIONS
The results from this study on chemical engineering students in an introductory mass & energy balance suggest that humour can positively impact knowledge transfer for a challenging technical problem. Application of humour in an integrated fashion did increase germane cognitive load, and students who reported an increase in germane cognitive load also showed improved educational outcomes. Since the educational outcomes were measured using standard chemical engineering test problems, we assume that knowledge transfer occurred between our engaging humorous example and chemical engineering fundamentals. This was a focused investigation using a humorous example that was isotropic with the control educational lesson. Consistent with the literature, these results indicate that our isotropic example did facilitate knowledge transfer. However, a test with a non-isotropic example should be made to confirm that the isotropic example is necessary to produce knowledge transfer.

ACKNOWLEDGEMENTS
We gratefully acknowledge financial support from Grant # EEC-1340480 of the NSF RIGEE Program, and assistance from S. Bramblett of the Georgia Tech office of Institutional Research & Enterprise Data Management. We also appreciate the contribution of the Georgia Tech students involved in this study and the graduate and undergraduate teaching assistants associated with these classes. The contributions of Dr. Jessica Plumley, and Wendy Newsstetter is also gratefully acknowledged.

REFERENCES


