Integrated Humor in the Engineering Classroom Using Isomorphic Mapping

Jessica Plumley,[1] David MacNair,[2] Wendy Newstetter,[3] Peter J. Ludovice[1]

[1]School of Chemical & Biomolecular Engineering, [2]School of Mechanical Engineering, [3]College of Engineering

Georgia Institute of Technology, Atlanta, GA 30332, U.S.A.

Abstract

Numerous claims that humor can improve education have been made in various areas. However, there is precious little direct evidence of this, particularly in technical areas such as engineering. A method to formulate such humorous/relatable examples is described here that is simple for technical educators to use in educational interventions. Using a cognitive load survey, adapted from previous survey instruments, we characterized the cognitive load components of students in response to traditional and humorous/relatable educational interventions in two chemical engineering classes at a large southeastern engineering school. This survey produced self-consistent results that were used to determine what type of humorous/relatable intervention could improve the germane cognitive load associated with improved educational outcomes. If the intrinsic cognitive load is relatively high, a humorous/relatable example integrated into the educational lesson can improve germane cognitive load. However, the example must be isomorphic in structure to the associated educational lesson to achieve such results. If the intrinsic cognitive load is sufficiently low, then increased germane cognitive load can be achieved without such an intervention. Humor itself is not actually required as long as the humorous/relatable example isomorphic and integrated into the educational lesson. However, humor often results from the incongruity of mixing a technical and non-technical concept that are isomorphic. Humor can add the additional benefits of increased student engagement and instructor immediacy.

Introduction

Humor has been associated with various positive effects in education including an improvement in instructor immediacy, and educational satisfaction (Wanzer, 1999; Freitas, 1998; Arbaugh, 2001; Aragon, 2003). However, it is difficult to see the connection between these improvements and educational outcomes. For example, McMorris et al. (1997) associates humor with reduction in test anxiety and stress, but this does not regularly translate into improved performance. Given the challenging nature of science technology engineering and math (STEM) curricula, we previously suggested that humorous analogues to challenging technical problems may help learning, but the humor must be integrated with the educational intervention to be effective (Hu et al., 2017). Humor has also been shown to increase creativity (Ziv, 1983), and has been shown to be effective in technical innovation (Ludovice et al. 2010, 2013).

Here we hypothesize that humor will improve engineering education by increasing germane cognitive load. Kalyuga (2009) contends that increased germane cognitive load can improve learning outcomes. 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. However, the addition of humor can distract from learning by increasing the extraneous cognitive load. Previous work indicates that simply adding humor to technical education does not improve educational outcomes (Vance, 1987 & Fisher, 1997). Added humor might distract from learning just as extraneous details can, as seen in the seductive detail effect described by Harp (1998). Previously, we argued humor must be integrated into the educational lesson to avoid increasing the extraneous cognitive load (Hu et. al., 2017). Here we integrate humor with engineering lessons and test the effect on components of the cognitive load. We hypothesize that humor will increase cognitive load as long as it is sufficiently integrated to avoid undue increase in the extraneous cognitive load. This is tested by measuring cognitive load associated with humorous and traditional educational interventions in an engineering classroom. An isomorphic pair of these interventions (one traditional and the other humorous) will also be used to check if such an increased cognitive load results in improved learning outcomes.

The educational intervention here applies a very specific type of humor based on our generalized theory of humor. There are numerous theories of humor that are sometimes confusing and redundant. These range from superiority theory (Duncan, 1985), to incongruity theory (Attardo, 1997), and relief theory (Shurcliff, 1968). Theories such as the generalized theory of verbal humor (Attardo, 2017) described humor based on topical categories. Superiority theory is essentially a topical-based theory that addresses humor based on one group assuming a superior position relative to another group. Examples include "Lawyer Jokes" in which society assumes a morally superior position relative to lawyers whose lack of ethics is comically exaggerated. Such topical characterization is a secondary effect to what actually produces humor. Topics are relevant to the humor audience in that the audience may not be familiar with the humor topic, or may be offended by it. The fundamental production of humor is derived from the relationship between two narratives inside each joke. If there is an incongruity between these two narratives, a humorous response often results as the audience attempts to reconcile this incongruity. However, an increased level of humorous response may result when the audience finds some truth in the incongruity as they attempt to reconcile it. Thus, humor results from two narratives that are simultaneously partially true and partially false at the same time, and our attempts to reconcile this fact. Puns are a simple illustration of this as the multiple meanings of a homophone create two simultaneous narratives in which one meaning is true and the other meaning is false or incongruous. This theory is the essence of previous work by Raskin (2012) and Hurley and co-workers (2011), however they state the theory in slightly different terms. Raskin states that humor requires two narratives that are incongruous, but also contain a partial similarity. Hurley and co-workers state that humor occurs when an element is originally taken to be true, and is later diagnosed as false consistent with the aforementioned theory. Further, Hurley and co-workers suggest that reconciling this true/false dichotomy produces laughter as a reward for this reconciliation that clears up conflicting ideas in the mind (equivalent to garbage collection in disk drives or dynamically allocated memory space in computers). Laughter has evolved as a reward for this important garbage collection behavior of the mind. It is also possible that laughter has evolved as a reward for the reconciling a similar reconciliation of ideas that produces innovation (Ludovice, et. al. 2010, 2013).

Employing the aforementioned theory of humor. We generate humor by juxtaposing a technical topic or procedure with a non-technical one. This juxtaposition produces the integrated humor described previously (Hu et. al., 2017) that we hypothesize is critical to minimizing the extraneous cognitive load associated with the limitations of isolated humor. Such integration can also potentially reduce the seductive detail effect. Juxtaposing these technical and non-technical ideas produces an incongruity, and the truth or partial similarity results if the two topics are isomorphic. 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 here, but the second approach was used 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).

Bassok and Holyoak (1989) 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. Gass and Priest (1993) found similar value in isomorphic examples for producing learning transfer in outdoor education. Therefore, in addition to testing the relative cognitive loads for our examples we will attempt to quantify knowledge transfer using the isomorphic LLE example described above.

An additional advantage of using the aforementioned juxtaposition is to make it easier for technical instructors in the science, technology, engineering and math (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 humorous. It certainly can be humorous, but the humor is typically generated simply by the juxtaposition of the two concepts. Berk (2002) suggests choosing well-known examples of popular culture for this non-technical element.

Methods

In this work, we will use a modified instrument for measuring the various cognitive loads after both traditional (control) and humorous/relatable (experimental) educational interventions. This instrument was applied in class one class period later than the application of the educational intervention to both a chemical engineering applied numerical methods class (Spring semester 2016) and a chemical engineering mass and energy balance class (Fall semester 2016) at a large southeastern university. These are required courses in the chemical engineering curriculum, and all students were chemical engineering majors. Typically, students take the mass and energy balance class in the first semester of their sophomore year, and the numerical methods class in the second semester of their sophomore year. Both classes were comprised of two temporally adjacent sections, and all classes were taught by the same instructor with over two decades of experience teaching these classes. This research was carried out under Institute Review Board Protocol H13322 for human subjects research. The cognitive load instrument was adapted for these two classes, and the instrument format was identical, except it was applied in the numerical methods class as a paper survey, and in the mass and energy balance class as an on-line survey students could take with their laptop or smartphone. In addition to the cognitive load measurements, class exams were used analyze learning outcomes relative to the cognitive load results.

Cognitive load theory separates the cognitive load required for learning into three components: (i) intrinsic, (ii) germane, and (iii) extraneous (Sweller, 1988). 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 at hand. We will utilize the original three-component version of cognitive load theory despite Kalyuga and co-workers (2011) claim that germane cognitive load is not independent, and is simply a form of intrinsic cognitive load. 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.

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.

} intrinsic

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.

} extraneous

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 (or numerical methods).**
9. The activity really enhanced my understanding of the concepts covered.
10. The activity really enhanced my understanding of the concepts and definitions.

} germane

Figure 1. 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.

We previously suggested that humorous and relatable examples could increase the germane cognitive load by providing an initial reference or framework for learning that is less intimidating (Hu et. al., 2017). By integrating the humorous/relatable example, the extraneous cognitive load can also be minimized. To test whether an integrated humorous example can improve the germane cognitive load while simultaneously not increasing the extraneous load we utilize a test instrument developed by Morrison et al. (2014).

The aforementioned cognitive load test instrument was originally based on an generalized instrument for cognitive load measurement developed by Leppink et al. (2013). This instrument is a ten-question test 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 by Morrison et al. (2014) 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. 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 numerical methods or 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 as seen in Figure 1. The reference to programming was changed to “mass & energy balances“ or “numerical methods“ for application to a basic chemical engineering mass and energy balance class or an engineering numerical methods class respectively, as see in the bold text in Figure 1.

TABLE I. Educational Interventions for Fall 2016 chemical engineering numerical methods classes. Types include HR=humorous/relatable example, and T=a traditional chemical engineering example.

Intervention	**Type**	**Description**
coffee plant	HR	Solve a linear mass balance on an industrial plant that makes instant coffee
VLE	T	Solve a set of non-linear equations to calculate the vapor-liquid equilibrium (VLE) of two organic solvents using Raoult’s Law.
dating service	HR	Carry out a statistical covariance analysis on a set of surveys typical of dating service questionnaires to determine correlations between personality variables
plant economics	T	Calculate the net present value and rate of return on the costs and income associated with a typical chemical plant
regression	T	Carry out linear regression to fit the constant pressure heat capacity (C_p) as a function of temperature using a functional form familiar from the previous mass & energy balance class

The numerical methods class was taught in two different classrooms for 80-minute periods twice weekly on Tuesdays and Thursdays. Section A was taught from 8:05 to 9:25am in a room with moveable tables that was well suited to the active group activities in a typical blended classroom. Section B was taught from 9:35 to 10:55 in a classroom with fixed terraced counters, but with moveable chairs. This room configuration was not as ideal as the venue for section A, but it was much better than the fixed theater-style seating in both sections of the mass and energy balance class. This particular class was a blended classroom in which the students watched a video on the fundamentals of a particular numerical method before class, and then engaged in group activities to program that numerical method in class and apply it to an engineering example that was either traditional or humorous/relatable. After class, the students completed the MATLAB program they began in class, and carried out a more extensive version of the project begun in class. Table I below shows the educational interventions examined in this class. They are drawn from two different types: a more traditional example based on typical chemical engineering processes (T), and an example that is non-traditional that we think is less intimidating because it is humorous or more relatable (HR).

TABLE II. Educational Interventions for Fall 2016 chemical engineering mass & energy balance classes. Types include HR=humorous/relatable example, and T=a traditional chemical engineering example.

Intervention	Type	Description
hair dryer	HR	Students analyzed the change in mass, molar and volumetric flowrates between the input and output of a handheld hair dryer to understand conservation of mass and the manner in which volumetric flowrates vary with temperature and pressure
separation	T	Aromatic hydrocarbon separation using two distillation columns
liquor	HR	Calculate the vapor liquid equilibrium of ethanol and water to analyze the boiling points of various distilled spirits
LLE *	T	Analyze the liquid phase behavior of three partially miscible liquids used in liquid liquid extraction (LLE) with a ternary phase diagram
school *	HR	Carry out the same analysis as above, but the topic was introduced in class using a relatable analogy of the blending of boys and girls at a middle school dance
tank drain	T	The dynamic draining rate determined from the mechanical energy balance (MEB) was integrated to determine the time required to drain a tank
hot tub	HR	A transient energy balance on water recirculating between a hot tub and the adjacent heater to determine the temperature transient

* 9:05am class used a humorous/relatable example of the middle school dance, while the 10:05am class used a traditional example of liquid liquid extraction (LLE)

The mass and energy balance class is typically taken in the second semester of the first year, or the first semester of the second year. This is a prerequisite for the numerical methods class discussed below. Since we tested the numerical methods class before the mass and energy balance prerequisite there was no overlap of student subjects in the study of these two classes. 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. 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 video or reading material to motivate them to prepare for class. Table II describes the interventions used in this class.

Results and Discussion

The results from the cognitive load surveys for the numerical methods class interventions described in Table I are seen in Figure 2, which contains the mean values for both classes of the three cognitive load components. The variation in sample size is due to the variation of students enrolled in the study who were in class that particular day when the cognitive load survey was administered. The interventions are listed in the order they were applied, and the slight decrease in sample size may also be indicative of survey fatigue as the semester progressed.

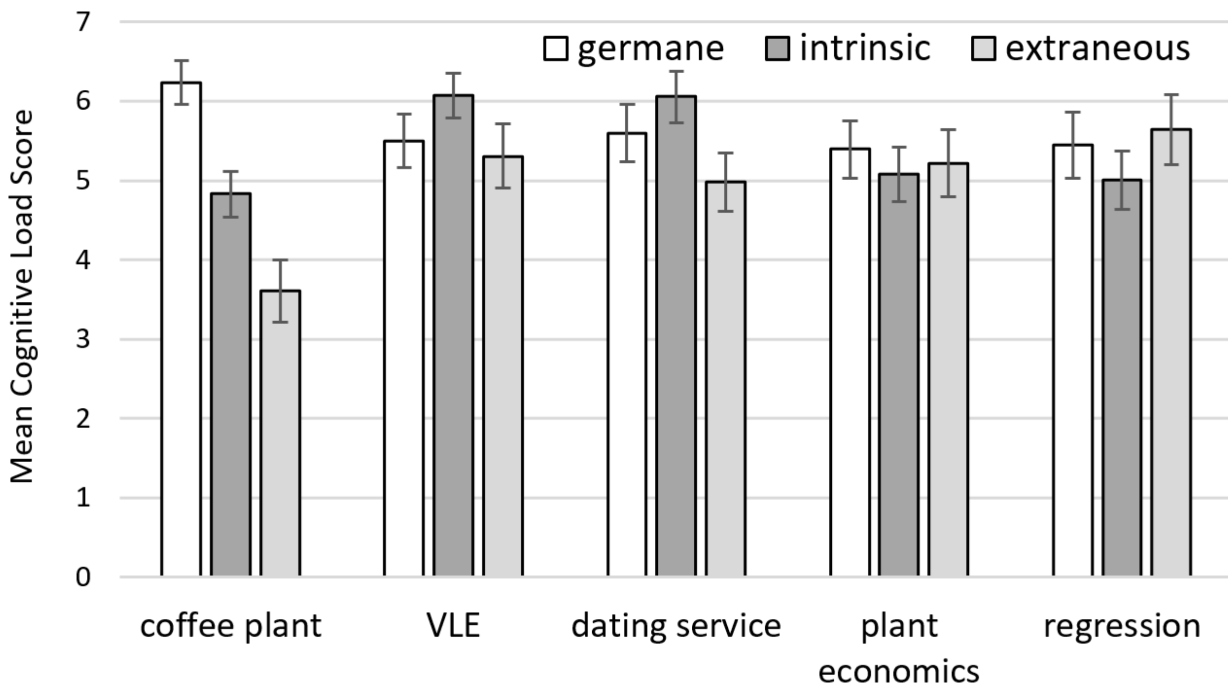

FIGURE 2. Mean values of the cognitive load survey for the three cognitive load components for the interventions used in the numerical methods class: coffee plant (N=92), VLE (N=93), dating service (N=88), plant economics (N=79), and regression (N=77). The error bars are the 90% confidence intervals about the mean.

According to the proposed hypothesis, the three traditional interventions from Table I should not see a large germane cognitive load. In contrast, the two humorous/relatable interventions should see a large germane load without increases in the extraneous cognitive load. The results of the cognitive load survey in Figure 2 are consistent with this hypothesis, with the exception of the “dating service” intervention. For all the traditional interventions, the difference between the mean cognitive load score for the germane load was not statistically different than the intrinsic load. However, the germane load was statistically different from the intrinsic load for the “coffee plant” intervention ($p < 0.01$). This also occurred with a small extraneous load component. In Figure 2, this intervention shows the desired pattern of the germane load above the intrinsic load, and the extraneous load lower than the intrinsic load. However, this pattern was not observed for the “dating service” intervention.

This lack of a high germane cognitive load in one of the two humorous/relatable interventions likely suggests that additional variables are important in the prediction of a successful educational intervention. Additional relevant variables might include: (i) the nature of the intervention, and (ii) the nature of the educational lesson.

In examining the difference between the two humorous/relatable examples, we note that the coffee plant example was relatable, but not very humorous. While water, coffee grounds, and soluble coffee components are easier for a student to relate to than obscure industrial chemicals, there is nothing funny about them. In contrast, the dating service intervention was relatable and humorous via the juxtaposition of the dating service data with covariance analysis. We noted a humorous response when this approach was introduced in class by briefly discussing the patent assigned to eharmony.com (U.S. Patent 6,735,568 B1, 2004) that discusses such analysis methods in an on-line dating service. This suggests that a relatable example or context is sufficient to increase germane cognitive load, and humor may interfere with effective learning.

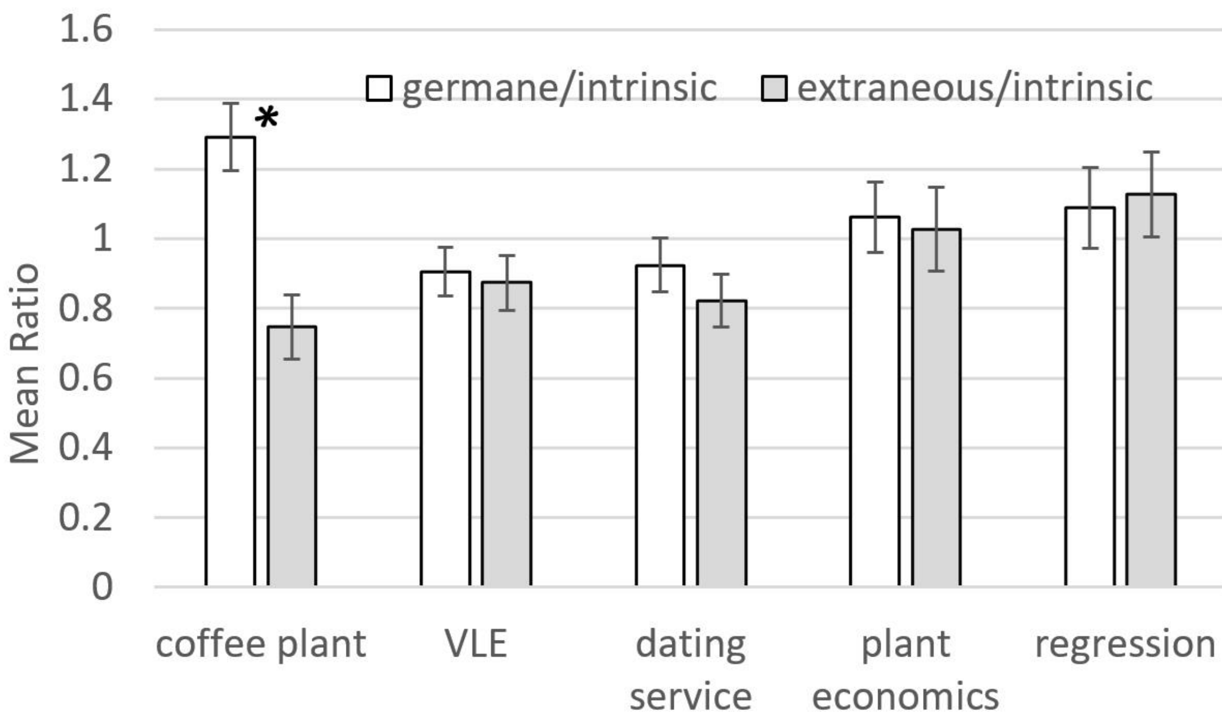

FIGURE 3. Mean values of the ratio of the germane and extraneous cognitive load to the intrinsic cognitive load for the interventions used in the numerical methods class. The error bars are the 90% confidence intervals about the mean. *statistically different means ($p < 0.01$).

Another possibility is that the results may depend on the nature of the educational goal. The variation in intrinsic cognitive load among the interventions in Figure 2 might also be an important variable. In Figure 3 we normalize the results based on the variable intrinsic cognitive load by dividing the respective cognitive loads by the intrinsic load, which should not vary with changing educational approaches. Figure 3 shows similar results to Figure 2. The “coffee plant” intervention should improve germane cognitive load without increasing the extraneous cognitive load, while the "dating service" intervention produced similar results to the traditional interventions. However, the intrinsic cognitive load for the "dating service" interventions was statistically higher than the "coffee plant" intervention ($p < 0.01$). It may be that the application of a humorous/relatable example is only effective if the intrinsic cognitive load is below a certain threshold.

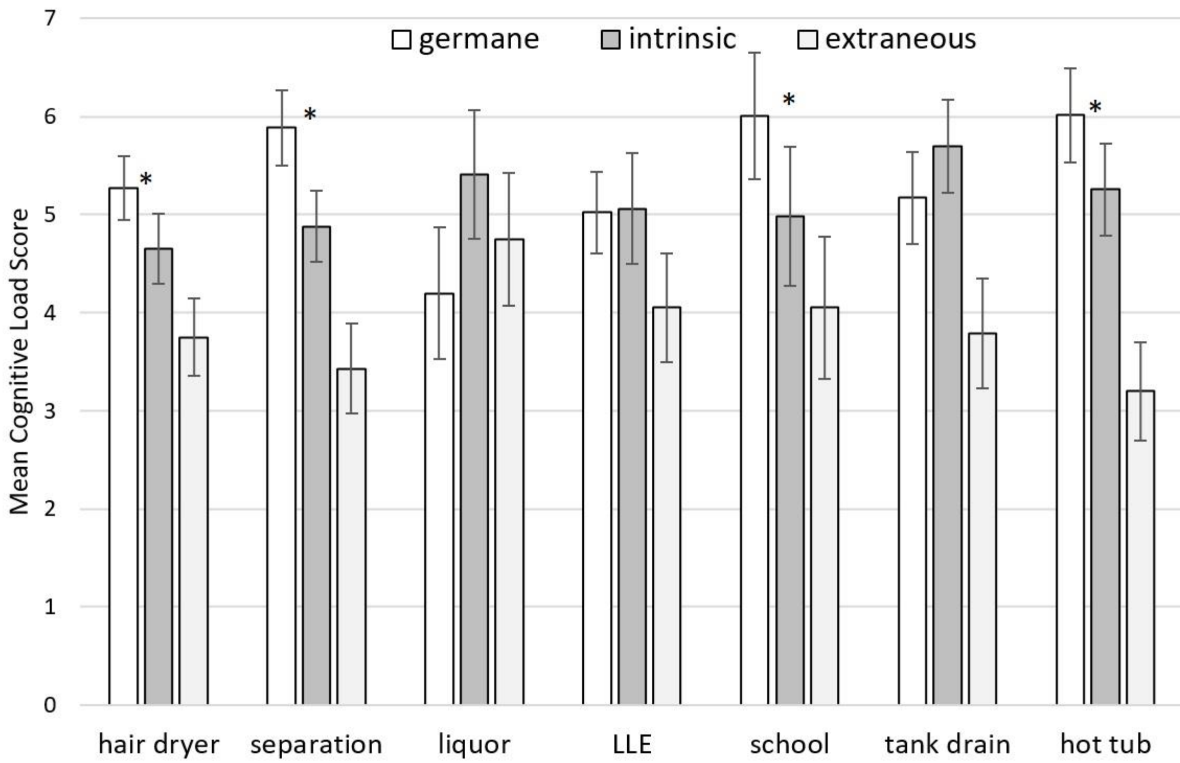

FIGURE 4. Mean values cognitive load survey for the three cognitive load components for the interventions used in the numerical methods class: hair dryer (N=94), distillation (N=75), liquor (N=38), LLE (N=42), school (N=30), tank drain (N=44), and hot tub (N=44). Both sections were included in the means except for the LLE and school intervention which were applied in different sections. The error bars are the 90% confidence intervals about the mean. * indicates germane load is statistically different from intrinsic load ($p < 0.05$).

The significantly higher intrinsic cognitive load of the "dating service" intervention relative to the "coffee plant" intervention is due to the complexity of the applied covariance matrices. Producing and visualizing a normalized covariance matrix is an advanced topic that is new to the students. The "coffee plant" intervention required the solution of a set of linear equations to solve a mass balance. Students had done this previously in the mass and energy balance course that is a prerequisite for this numerical methods class. The students had solved such problems before with one to three linear equations, but this problem involved the solution of eight to ten (depending on how the mass balance was written) linear equations. Students were sufficiently familiar with this type of problem such that a humorous/relatable context was not even required to produce increased germane cognitive loads. In contrast, the "dating service" example was sufficiently complex that no modification of the educational approach would have increased germane cognitive load because the intrinsic cognitive load was too high. We will continue to analyze this possibility using examples from the mass and energy balance below.

The mean cognitive load scores for the mass and energy balance class interventions from Table II are seen in Figure 4. As with the numerical methods class, the results are mixed. We observe that the interventions showed a value for the germane cognitive load that is greater than the intrinsic cognitive load for three of the four humorous/relatable interventions. Two of the three traditional interventions showed a low germane cognitive load. These results are similar even when they are normalized by the intrinsic cognitive load seen in Figure 5.

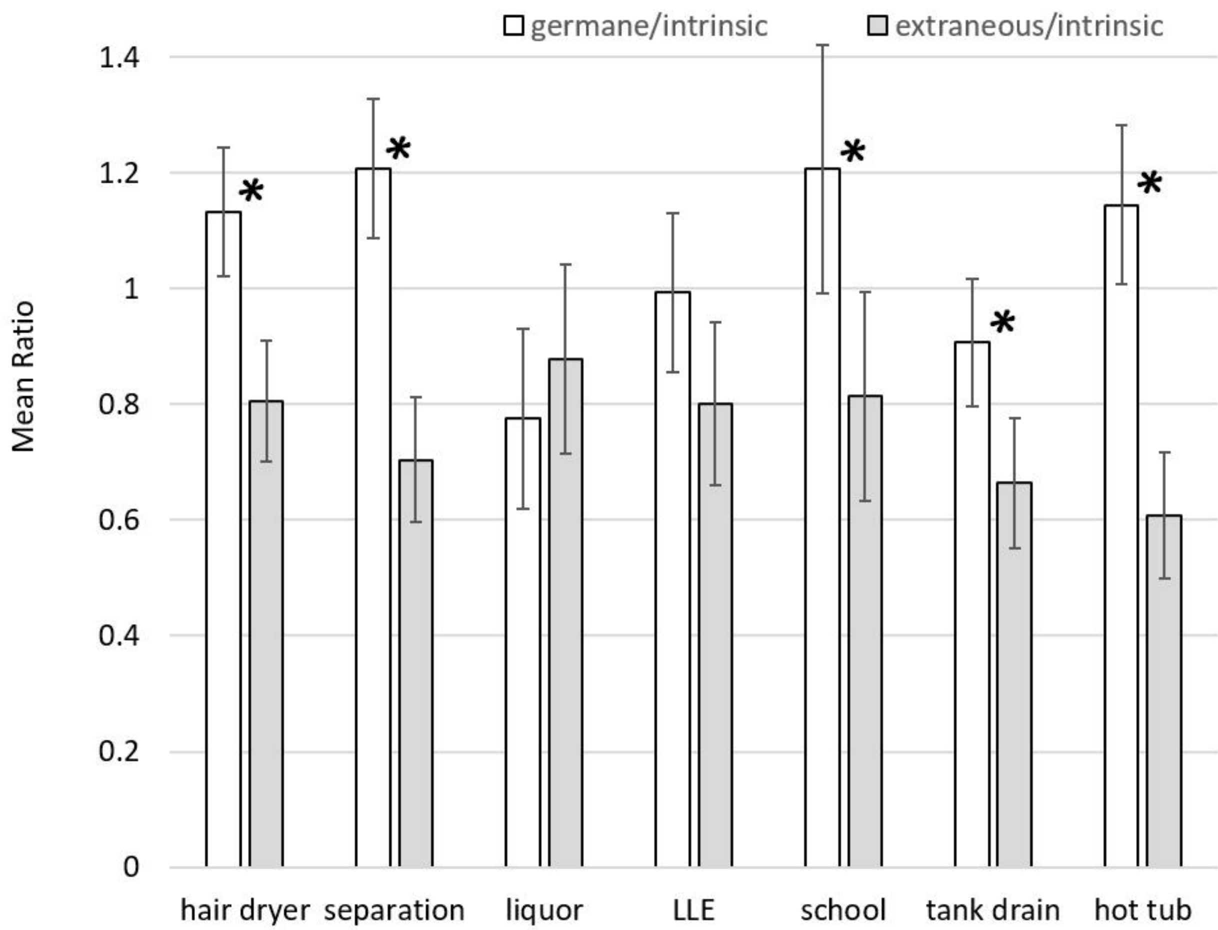

FIGURE 5. Mean values of the ratio of the germane and extraneous cognitive load to the intrinsic cognitive load for the interventions used in the numerical methods class. The error bars are the 90% confidence intervals about the mean. *statistically different means ($p < 0.01$).

The failure of the "liquor" humorous/relatable intervention may be due to the high intrinsic cognitive load associated with this educational lesson similar to the "dating service" intervention results from Figures 2 and 3. It appears that intrinsic cognitive load scores greater than approximately five may be problematic for this approach. The "separation" intervention used the traditional context of distillation columns separating aromatic organic chemicals, but it still produced a high germane cognitive load. In contrast, the other traditional interventions ("tank draining" and "LLE") did not produce a high germane load. Perhaps this was due to their value of the intrinsic cognitive load greater than 5. While both traditional interventions with intrinsic cognitive load scores above five ("tank draining and "LLE") did not produce a high germane load, two of the three humorous/relatable interventions with intrinsic load scores about five did.

We contend that the failure of the "liquor" intervention to produce a high germane load is due to the structure of this intervention. The "hot tub" and "school" interventions are of similar structure, or isomorphic, with the related chemical engineering problem. The "hot tub" intervention is a recirculating loop between a vessel and a heater that is identical to the same loop in the reboiler of a distillation column. However, the problem produces a temperature profile that students can relate to their own experience waiting for a hot tub to warm up. Similarly, the "school" intervention is a completely isomorphic variation of the "LLE" intervention. This example uses the exact same ternary phase diagram for liquid liquid extraction (LLE) but labels it differently with numbers of middle school boys and girls and chaperones. Adult chaperones are the common solvent that motivate boys and girls to mix at a middle school dance. This is much more relatable than using a common solvent to induce two immiscible solvents to mix. In contrast, the "liquor" example simply changed the context of ethanol distillation to that of understanding the process of distilled spirits. This additional context simply added additional information that is not directly assisting in understanding the problem. The very high value of the normalized extraneous load for the "liquor" intervention (seen Figure 5) indicates this context is extraneous and detracts for the germane cognitive load. This is the only intervention whose normalized extraneous cognitive load is higher than the germane load (see Figure 5). This appears to be an example of the seductive detail effect discussed above.

The results from the mean cognitive load scores in Figures 2 through 5 indicate that educational lessons with low intrinsic cognitive load can produce high germane cognitive load even with traditional educational approaches. While the addition of a humorous/relatable context or analogy to the lesson may improve global variables such as instructor immediacy, they may not be necessary to achieve high cognitive load associate with improved learning. This is the fundamental approach of cognitive load theory applied to education: to reduce intrinsic cognitive load such that there are sufficient mental resources to devote to germane cognitive load. This reduction is often achieved by breaking the educational lesson into sufficiently small segments. However, some technical concepts are sufficiently advanced that even this reductionist approach will result in minimal sized segments that still retain significant intrinsic cognitive load. For such lessons, it does appear that using a humorous/relatable example as context or an analogy can increase germane load. However, the integration of this example suggested previously (Hu et. al. 2017) is not sufficient. The humorous/relatable example must be isomorphic in structure to the technical problem and not just add extraneous details to reproduce the seductive detail effect.

The cognitive load scores appear to be consistent, exhibiting a high extraneous load for a non-isomorphic intervention. Further, the "LLE" and "school" interventions are traditional and humorous/relatable approaches to the identical educational lesson, and they have essentially identical intrinsic cognitive load scores. This is consistent with the fact that intrinsic cognitive load should not be a function of the specific intervention used to teach the particular lesson. The significant change in the germane load with no change in the intrinsic load suggests the germane load is independent of intrinsic load (at least for these complex technical problems). While increased germane cognitive load has been associated with improved learning outcomes (Kalyuga, 2009), this is not direct evidence that our approach results in such improvements. Thus, we will use the "LLE" and "school" interventions to exam educational outcomes.

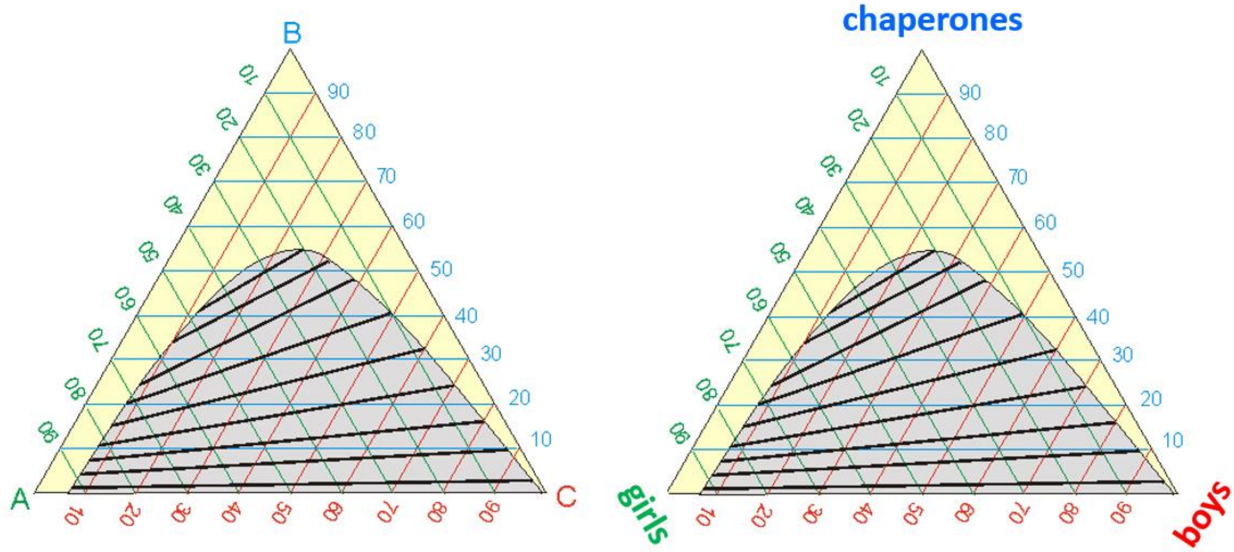

FIGURE 6. Ternary phase diagrams used in the "LLE" intervention (left), and the "school" intervention (right). The diagram depicts the two-phase region in gray of three generic liquid solvents and how a single phase is achieved by adding solvent B (left), and the analogous example of chaperones being used to get middle school boys and girls to mix at a school dance (right). The lines (tie lines) in the two-phase region indicate the amount of solvent B in the A-rich and C-rich phases and the number of chaperones in the girl-rich and boy-rich phases.

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 6. After each of these different interventions, the students went on to solve identical homework and test problems involving liquid liquid extraction using a real ternary phase diagram of water, acetone and methyl isobutyl ketone from the text book. The tie lines in Figure 6 indicate the composition of the phases that separate in the two-phase region. These indicate that the C-rich phase has a higher concentration of component B than the A-rich phase. This composition distribution is determined by molecular interaction between the solvents. Because this molecular interaction is an advanced concept (see the intrinsic cognitive load for the "LLE" and "school" interventions in Figure 4), we use the analogy of a middle school dance to make this more relatable. In the right side of Figure 6, when the boys and girls are in separate groups there are more chaperones with the boys. This assumes that chaperones believe the boys are more likely to cause trouble at a dance, and require more direct supervision. Further, we showed students the same ternary phase diagram with the slant of the tie lines reversed such that there were more chaperones in the groups that contained mainly girls. This was associated with middle school girls on the internet given the stereotype that girls more aggressively exclude their peers on social media (although research does not confirm this). Poking fun at an undeserved stereotype is a more relatable example then explaining that changing the fluid pressure can sometimes change the molecular interactions to reverse the tie lines. Using the identically shaped ternary phase diagrams with different variable labels makes these two interventions isomorphic.

The increased germane cognitive load associated with the “school” intervention should be associated with improved learning outcomes. Woo et. al. (2014) found a strong correlation between performance and germane cognitive load. However, we also used the two sections of the same mass and energy balance class to examine actual educational performance. Questions on the ternary phase diagram and the use of this diagram to carry out a mass balance on a liquid liquid extraction process were included on both the second of three exams during the semester, and the final exam. Unfortunately, the relevant questions on the final exam were different because the final exam questions were different. This was required because, while the classes were temporally adjacent, the final exams were scheduled on different days. To prevent one class section from talking with students in the other section about the final exam content, different exams were constructed. However, the second in-class exam was given in adjacent time periods so the identical questions were given. Specifically, question 2 on the second exam asked students about ternary phase diagrams. Part B of question 2 asked students to determine the amount of common solvent that would allow a mixture of the two immiscible solvents to blend into a single phase. This is exactly what the “school” intervention addressed. Part A of the question asked students to take information from the ternary phase diagram and carry out a mass balance on liquid liquid extraction. Question 2A acts as a generic control in this topic as this was not addressed specifically by the “school” intervention.

Using two sections of the same class as a control and experimental sample is only useful to the extent that these two sections represent similar cohorts. While the instructor, text, classroom, and curriculum were identical for the control and experimental group they were held at different times. The experimental group was the earlier class and it is possible both the students and instructor were more alert later in the day. Also, the sampling could be biased because students with more credit hours are allowed to register first, and they tend to register for later classes. To examine this, we found that the later class (control group) had taken more total hours in college, had a higher H.S. and college GPA as well as a better GPA in this particular class. The control group also exhibited better scores on the math, verbal and writing portions of the SAT exam. While none of these differences were statistically significant (the largest difference was in SAT Math scores, $p<0.13$) it does suggest that the control group may contain students with better academic performance. This result is consistent with the observation by Crisp et. al. (2009) that the SAT-Math scores are correlated with performance of STEM majors at Hispanic Serving Institutions. The control group contained 5% students admitted as transfer students, and 95% students admitted as freshmen. The experimental group contained 24% transfer students, 2% students re-admitted after being dropped for insufficient GPA, and the remainder freshman admit students. D’Amico (2014) showed that community college transfer students in the STEM fields typically completed less initial hours and had a lower initial GPA after transferring. Wang (2015) showed that community college transfer students do not perform as well when pursuing STEM degrees

The results of question 2 on exam 2 are seen in Figure 7, and show no statistical difference between the control section ("LLE" intervention) and the experimental section ("school" intervention). If the "school" intervention improved learning outcomes then the experimental section should have performed better on question 2B in Figure 7. Not only was there no statistical difference between the performance, the control group actually performed better. The fact that the control group performed better on both parts of the question may suggest that the control group is academically superior to the experimental group. This is qualitatively consistent with the superior performance of the control group in the aforementioned metrics like GPA and standardized tests.

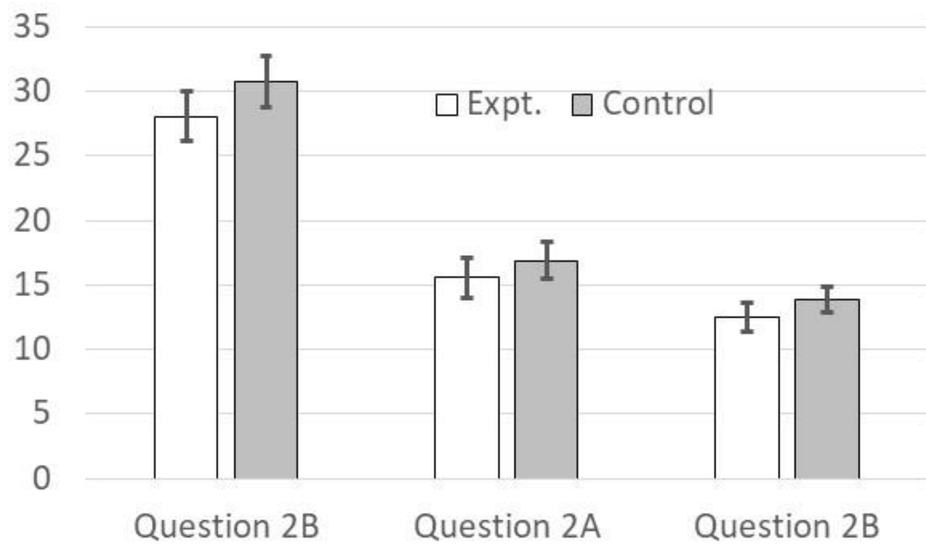

FIGURE 7. Average scores for question 2 from exam 2 in the mass and energy balance class for the control section for the "LLE" intervention (N=31), and the experimental section for the "school" intervention (N=25). The error bars are the 90% confidence intervals about the mean. None of these differences were statistically significant ($p < 0.135$ for question 2B).

TABLE III. Pearson correlation (r) between the cognitive load results and the performance in question 2 from exam 2 on ternary phase diagrams and their statistical significance (p).

	Experimental ("school" intervention)			Control ("LLE" intervention)		
cognitive load	**question 2A**	**question 2B**	**question 2 (total)**	**question 2A**	**question 2B**	**question 2 (total)**
intrinsic	r=0.1554 p=0.4068	r=–0.1641 p=0.3778	r=0.0367 p=0.8448	r=–0.1187 p=0.5720	r=–0.0379 p=0.8574	r=–0.1033 p=0.6231
extraneous	r=0.1399 p=0.4527	r=–0.2447 p=0.1791	r=–0.2430 p=0.8068	r=0.0939 p=0.6552	r=–0.0123 p=0.1557	r=–0.0738 p=0.7621
germane	r=–0.0123 p=0.9487	r=0.4115 p=0.0224	r=0.2128 p=0.2590	r=–0.1258 p=0.5490	r=0.1472 p=0.4826	r=–0.0192 p=0.9274

A more focused analysis of this intervention is provided by the Pearson correlation of the cognitive load results and the performance on question 2 of exam 2 in Table III. The only statistically significant correlation in Table IV is the positive correlation between germane cognitive load and performance on question 2A for the experimental group ($p < 0.023$). This result indicates that students in the experimental group who had a high germane cognitive load score also performed well on the relevant question of ternary phase diagrams (question 2B).

Conclusions

The results of cognitive load surveys in these two chemical engineering classes suggest that humorous/relatable contexts or analogies can increase germane cognitive load. The consistency of the cognitive load results also indicate that the modified cognitive load survey adapted for this work is a useful tool for probing cognitive load distributions in technical classes. The increase in germane cognitive load, which is associated with improved educational outcomes, occurs if the humorous/relatable example is isomorphic with the engineering lesson being taught. If the intrinsic cognitive load is sufficiently low (below approximately 5 using the instrument adapted here), then the germane cognitive load will be sufficiently large even in traditional educational approaches. If the intrinsic cognitive load is higher (above approximately 5), integration of a humorous/relatable example into the educational lesson can improve the germane cognitive load, and this improvement did correlate with improved learning outcomes. However, this humorous/relatable example must be structurally isomorphic with the technical problem being studies. Such a humorous/relatable intervention may not be necessary to achieve a high germane cognitive load if the intrinsic load is sufficiently low. The approach to formulating generalized humor for the engineering or technical classroom described herein is general and easily applied. This approach requires the simple juxtaposition of a technical and non-technical concept that can produce an integrated intervention which will improve learning if the technical and non-technical concepts are sufficiently isomorphic in structure. The isomorphic structure coupled with the incongruity of a technical and non-technical juxtaposition fulfill the requirements of a generalized humor theory. The resulting humor can add the additional benefits of increased immediacy and engagement, which may be even more important for public STEM outreach as opposed to engineering education.

Acknowledgments

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. Helpful discussions with Lew Lefton and the students of the Georgia Tech Humor Genome Project were instrumental in developing the juxtaposition theory of humor used in this work.

References

Aragon, S. R. (2003). Creating social presence in online environments. *New directions for adult and continuing education*, *2003*(100), 57-68.

Arbaugh, J. B. (2001). How instructor immediacy behaviors affect student satisfaction and learning in web-based courses. *Business Communication Quarterly*, *64*(4), 42-54.

Attardo, S. (1997). The semantic foundations of cognitive theories of humor. *Humor: International Journal of Humor Research*.

Attardo, S. (2017). The general theory of verbal humor. *The Routledge handbook of language and humor*, 126-142.

Bassok, M., Holyoak, K.J. (1989). Interdomain Transfer between Isomorphic Topics in Algebra and Physics. *Journal of Experimental Psychology: Learning Memory, and Cognition, 15(1)*, 153-166.

Berk, R. A. (2002). *Humor as an instructional defibrillator: Evidence-based techniques in teaching and assessment*. Stylus Publishing, LLC.

Crisp, G., Nora, A., Taggart, A. (2009). Study Characteristics, Pre-College, College, and Environmental Factors as Predictors of Majoring and Earning a STEM Degree: An Analysis of Students Attending a Hispanic Serving Institution. *American Educational Research Journal, 46(4)*, 924-942.

D'Amico, M.M., Dika, S.L., Elling, T.W., Algozzine, B., Ginn, D.J. (2014). Early Integration and Other Outcomes for Community College Transfer Students. *Research Higher Education, 55*, 370-399.

Duncan, W. J. (1985). The superiority theory of humor at work: Joking relationships as indicators of formal and informal status patterns in small, task-oriented groups. *Small Group Behavior*, *16*(4), 556-564.

Fisher, M. S. (1997). The effect of humor on learning in a planetarium. *Science Education*, *81*(6), 703-713.

Freitas, F. A., Myers, S. A., & Avtgis, T. A. (1998). Student perceptions of instructor immediacy in conventional and distributed learning classrooms. *Communication Education*, *47*(4), 366-372.

Gass, M., Priest, S. (1993) Using Metaphors and Isomorphs to Transfer Learning in Adventure Education. *Journal of Adventure Education and Outdoor Leadership, 10(4)*, 18-23.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of educational psychology*, *90*(3), 414.

Hu, D. L., Lefton, L., & Ludovice, P. J. (2017). Humor Applied to STEM Education. *Systems Research and Behavioral Science*, *34*(3), 216-226.

Hurley, M. M., Dennett, D. C., & Adams, R. B. (2011). *Inside jokes: Using humor to reverse-engineer the mind*. MIT press.

Kalyuga, S. (2009). Instructional designs for the development of transferable knowledge and skills: A cognitive load perspective. *Computers in Human Behavior*, *25*(2), 332-338.

Kalyuga, S. (2011). Cognitive load theory: How many types of load does it really need?. *Educational Psychology Review*, *23*(1), 1-19.

Leppink, J., Paas, F., Van der Vleuten, C. P., Van Gog, T., & Van Merriënboer, J. J. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior research methods*, *45*(4), 1058-1072.

Ludovice, P. J., Lefton, L. E., & Catrambone, R. (2010). Improvisation for engineering innovation. In *ASEE Annual Conference and Exposition, Conference June* (pp. 20-23).

Ludovice, P., Lefton, L., & Catrambone, R. (2013). Humorous Improvisation Tailored for Technical Innovation. In *Proceedings of the American Institute of Engineering. Education: SE Regional Meeting*.

McMorris, R.F., Boothroyd, R.A., Pietrangelo, D.J. (1997). Humor in Educational Testing: A Review and Discussion. *Applied Measurement in Education, 10(3)*, 269-297.

Morrison, B. B., Dorn, B., & Guzdial, M. (2014, July). Measuring cognitive load in introductory CS: adaptation of an instrument. In *Proceedings of the tenth annual conference on International computing education research* (pp. 131-138). ACM.

Raskin, V. (2012). *Semantic mechanisms of humor* (Vol. 24). Springer Science & Business Media.

Shurcliff, A. (1968). Judged humor, arousal, and the relief theory. *Journal of personality and social psychology*, *8*(4p1), 360.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science*, *12*(2), 257-285.

Vance, C. M. (1987). A comparative study on the use of humor in the design of instruction. *Instructional Science*, *16*(1), 79-100.

Wang, X. (2015). Pathway to a baccalaureate in STEM fields: Are community colleges a viable route and does early STEM momentum matter?. *Educational Evaluation and Policy Analysis*, *37*(3), 376-393.

Wanzer, M. B., & Frymier, A. B. (1999). The relationship between student perceptions of instructor humor and students' reports of learning.

Woo, J. C. (2014). Digital Game-Based Learning Supports Student Motivation, Cognitive Success, and Performance Outcomes. *Journal of Educational Technology & Society*, *17*(3).

Ziv, A. (1983). The influence of humorous atmosphere on divergent thinking. *Contemporary Educational Psychology*, *8*(1), 68-75.