
Developing Instrumentation for Assessing Creativity in Engineering Design

Abstract

A perceived inability to assess creative attributes of students' work has often precluded creativity instruction in the classroom. The Consensual Assessment Technique (CAT) has shown promise in a variety of domains for its potential as a valid and reliable means of creativity assessment. Relying upon an operational definition of creativity and a group of raters experienced in a given domain, the CAT offers the field of engineering education an assessment method that has demonstrated discriminant validity for dimensions of creativity as well as for technical strength and aesthetic appeal. This paper reports on a web-based adaptation of the CAT for rating student projects developed during a weeklong engineering camp. Images of resulting scale models, technical drawings, and poster presentation materials were displayed on a website which was accessed by a team of seven independent raters. Online survey software featuring a series of Likert-type scales was used for ratings. The raters viewed project images on larger computer screens and used iPads to input their assessments. This effort extended the accessibility of the CAT to raters beyond limitations of geographic location.

Keywords: Engineering Design, Creativity, Consensual Assessment Technique

The need for promoting creative thinking and innovative problem solving in classrooms has been established in the literature (National Research Council, 2002; Todd & Shinzato, 1999). Not only is creativity seen as an essential component of human cognition, but its promotion is essential to a global economy and creating globally competitive citizens (Kaufman, Baer, Cole, & Sexton, 2008). It is vital that teachers are able to effectively impart 21st century skills to our students, including creative and innovative skills (Fatt, 2000; P21, 2010). The cultivation of our high school students as innovative and creative problem solvers for today's technological problems has become a focus for STEM education in the 21st century (Dede, 2010; Fatt, 2000; P21, 2010).

Cameron D. Denson (cddenson@ncsu.edu) is Assistant Professor of Technology, Engineering and Design Education; Jennifer K. Buelin (jkbuelin@iteea.org) is Director of Digital Initiatives Division at ITEEA; Matthew D. Lammi (matthew_lammi@ncsu.edu) is Assistant Professor of Technology, Engineering and Design Education in the Department of STEM Education; and Susan D'Amico (susan_damico@ncsu.edu) is Coordinator of Engineering K-12 Outreach Extension, in the College of Engineering at North Carolina State University.

Engineering and technology education classrooms are uniquely positioned to offer a potentially fertile environment for developing students' problem-solving abilities and creative behavior (Lewis, 2005). With an emphasis on problem-based learning and open-ended questions, instructors of technology, engineering, and science education can provide students with a milieu conducive to the promotion of creativity. This is especially true for informal environments in which teachers are not bound by the standards-based restrictions of formal classroom settings.

Though the need for promoting creativity has been established in the literature, the task of fostering creativity and creative problem-solving skills can prove challenging amidst the classroom expectations of explicit objectives and measurable outcomes (Buelin-Biesecker & Weibe, 2013). This is especially difficult within the current goal framework of the average K–12 public school classroom, a context in which engineering education is gaining traction with the release of the *Next Generation Science Standards* (NGSS Lead States, 2013). Part of the challenge is that teachers may view creative students as “inattentive and disruptive,” tending to “wander away from the regular paths of thought” (Lau & Li, 1996, p. 348). Without effective measures of creativity and validated instruments for the assessment of creativity, the teaching of creativity will continue to face scrutiny amongst teachers. Much of this scrutiny can be attributed to a lack of research dedicated to developing strategies that help teachers identify creativity and assess creative attributes of student work (Lewis, 2005). It is the researchers' contention that the lack of validated assessment measures for creativity and a perceived inability to assess creative attributes of students' work has precluded the teaching and learning of creativity in STEM classrooms (Buelin-Biesecker & Weibe, 2013; Lewis, 2009).

Studies have shown, however, that the reliable assessment of creativity in students' design work is possible (Amabile, 1996; Hennessey, Amabile, & Mueller, 2011; Hickey, 2001). This paper highlights a novel approach to creative assessment of engineering design products in secondary classrooms. This paper reports on the results of using the Consensual Assessment Technique (CAT) for creativity assessment in an engineering design setting. CAT offers promise for the assessment of creativity in a myriad of different domains. Relying upon an operational definition of creativity, the CAT has proven to be a valid and reliable means of assessment. However, CAT's accessibility has been limited by the need for a group of expert raters experienced in a given domain to rate design products on site. To address this issue, this paper will report on a web-based adaptation of the CAT for rating student projects. If functional, the web-based version of the CAT offers the field of engineering education an assessment method that has demonstrated discriminant validity for dimensions of creativity as well as for technical and aesthetic appeal.

Background Literature

Creativity

When sorting through the profuse definitions and conceptual frameworks available for discussing the concept of creativity, it is useful to identify those most applicable to the task at hand; in this case, the topic of interest is the potential for fostering students' creativity in hands-on problem-solving activities in engineering design settings. Two types of definitions are useful to this discussion. Hennessey, Amabile, and Mueller (2011), whose work in creativity assessment has had tremendous influence upon the design of this study, offered the following:

Conceptual definition of creativity A product is considered creative to the extent that it is both a novel and appropriate, useful, correct, or valuable response to an open-ended task. (p. 253)

Operational definition of creativity A product or response is considered creative to the extent that appropriate observers independently agree that it is creative. Appropriate observers are those familiar with the domain in which the product was created or the response articulated. (p. 253)

Hennessey et al.'s (2011) conceptual definition is a useful guide for evaluating student products in technology and engineering education because student products and design processes will vary widely due to many factors and problems are often open ended. The definition assimilates many prior conceptual definitions (Cropley, 1999) and can be helpful in clarifying to students what is being asked of them when they are told that creativity is a part of their grades. The operational definition establishes the framework and justification for the use of Amabile's (1983) Consensual Assessment Technique (CAT) for evaluating creativity and other dimensions of student responses to open-ended design and problem-solving activities: If knowledgeable raters independently, and with an acceptable level of interrater reliability, determine that a student product is creative in its context, then by definition, it is. The creative outcomes sought in the engineering design curriculum will be assessed using this method for three major dimensions (creativity, technical strength, and aesthetic appeal) and for nine additional subdimensions (novel idea, novel use of materials, complexity, organization, neatness, effort evident, liking, pleasing use of shape or form, and pleasing use of color or value). Factor analysis reveals the CAT's discriminant validity, in effect revealing whether creativity was measured apart from other characteristics of students' work.

Consensual Assessment Technique (CAT)

The CAT is an evaluation tool used by creativity researchers for assessment of creative products by panels of raters. The method “is based on the assumption that a panel of independent raters familiar with the product domain, persons who have not had the opportunity to confer with one another and who have not been trained by the researcher, are best able to make such judgments” regarding “the nature of creative products and the conditions that facilitate the creation of those products” (Hennessey et al., 2011, p. 253).

Amabile (1996) describes consensual assessment as a technique of judging creativity based on an operational, rather than conceptual, definition of creativity. Amabile states that “a product or response is creative to the extent that appropriate observers independently agree it is creative. Appropriate observers are those familiar with the domain in which the product was created or response articulated” (Amabile, 1982; as cited in Amabile, 1983, p. 31). Recent studies have advanced Amabile’s work by applying the CAT in different contexts, including assessing the creativity of children’s musical compositions and nonparallel creative products (Baer, Smith, & Allen, 2004; Hickey, 2001).

The application of the CAT for making inferences about students’ work, and subsequent inferences about pedagogical strategies used in producing that work, depends upon acceptance of an operational definition of creativity, which is described above. Interrater reliability “quantifies the closeness of scores assigned by a pool of raters to the same study participants. The closer the scores, the higher the reliability of the data collection method” (Gwet, 2008, p. 29). As Hennessey et al. (2011) explained,

In the case of the consensual assessment technique, reliability is measured in terms of the degree of agreement among raters as to which products are more creative, or more technically well done, or more aesthetically pleasing than others. (p. 253)

By definition, interjudge reliability in this method is equivalent to construct validity: if appropriate judges independently agree that a given product is highly creative, then it can and must be accepted as such. (p. 256)

In order to claim that creativity is being isolated and measured apart from other characteristics of students’ work, it is essential to demonstrate an instrument’s discriminant validity. Items related to creativity will ideally receive consistently different ratings from items related to categorically different types of items. Many studies using the CAT have followed Amabile’s (1983) three clusters of dimension types (creativity, technical strength, and aesthetic appeal) and have included ratings of multiple related subdimensions (Buelin-Biesecker & Weibe, 2013). Figure 1 provides a list of subdimensions associated with each of the three major dimensions. Factor analysis determines the CAT’s

discriminant validity; optimally, items within each of those three clusters will consistently load together.

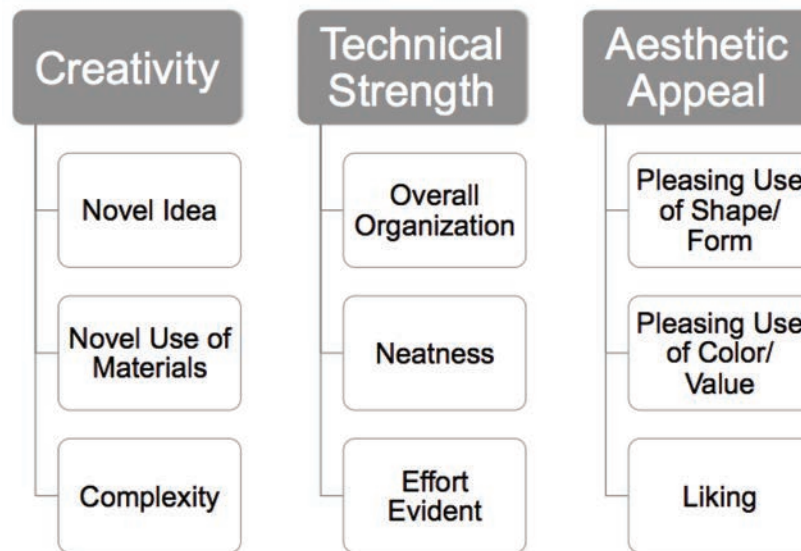


Figure 1. Subdimensions associated with each major dimension measured.

The rating instrument provided raters with a brief description of each subdimension. The creativity prompt was described this way: “Using your own subjective definition of creativity, the degree to which the design is creative.” Those subdimensions associated with creativity throughout Amabile’s body of work on the CAT include novel idea (the degree to which the design explores a unique and interesting idea), novel use of materials (the degree to which the use of materials is unique and interesting), and complexity (the level of complexity in the design).

The technical strength prompt was described this way: “The degree to which the work is good technically.” Those subdimensions associated with technical strength throughout Amabile’s body of work include overall organization (the degree to which the work shows good organization), neatness (the amount of neatness shown in the work), and effort evident (the amount of effort that is evident in the product).

The aesthetic appeal prompt was described this way: “In general, the degree to which the design is aesthetically appealing.” Those subdimensions associated with aesthetic appeal throughout Amabile’s body of work on the CAT include pleasing use of shape or form (the degree to which there is a pleasing use of shape or form in the design), pleasing use of color or value (the degree to which

the design shows a pleasing use of color or value), and liking (your own subjective reaction to the design; the degree to which you like it).

Informal Learning Environments

The informal learning environment framing the following study is classified as a programmed setting. Informal learning environments can be categorized into three major settings: (a) “everyday experiences,” (b) “designed settings,” and (c) “programmed settings” (Kotys-Schwartz, Besterfield-Sacre, & Shuman, 2011, p. 1). Programmed settings are characterized by “structures that emulate [or complement] formal school settings—planned curriculum, facilitators . . . , and a group of students who continuously participate in the program” (Kotys-Schwartz et al., 2011, p. 2). It is estimated that during the schooling years of students, 85% of their time will be spent outside of a classroom (Gerber, Cavallo, & Marek, 2001). This illustrates the importance of providing opportunities for learning that are outside of the traditional learning environment. Informal learning environments provide these opportunities and have been an integral part of education for years (Martin, 2004). The continued study of informal learning environments may provide insight into ways that the nation can address the issue of STEM education reform (Kuenzi, 2008). The merits of informal learning environments are known (Gerber et al., 2001), however little research is available that addresses their role in the cultivation of creativity. Informal environments were deemed appropriate for the exploration of creativity in this study because they are not bound by the standard-based restrictions of formal learning environments. However, it is argued that results from this study have implications for both informal and formal learning environments.

Description of the Innovation

Digital CAT interface

Creativity assessment conducted using the CAT has traditionally followed similar implementation processes: students create products that are collected by researchers, spread around a single physical space, and viewed and assessed in that space by one rater at a time until the ratings were completed. It may prove valuable to expand the accessibility of consensual assessment beyond the traditional method characterized by displaying student projects throughout a physical space and having raters complete the assessments in person. For this study, the researchers developed a web-based assessment interface consisting of (a) an overview video displaying all project images for raters to view prior to the rating session; (b) a website built for the display of project images and documentation; and (c) a web-based version of the consensual assessment instrument, accessed by raters via iPad while viewing the project website on desktop computers. The web-based version of the CAT consisted of images of

modeled artifacts resulting from the engineering design challenge (see Figures 2 and 3).

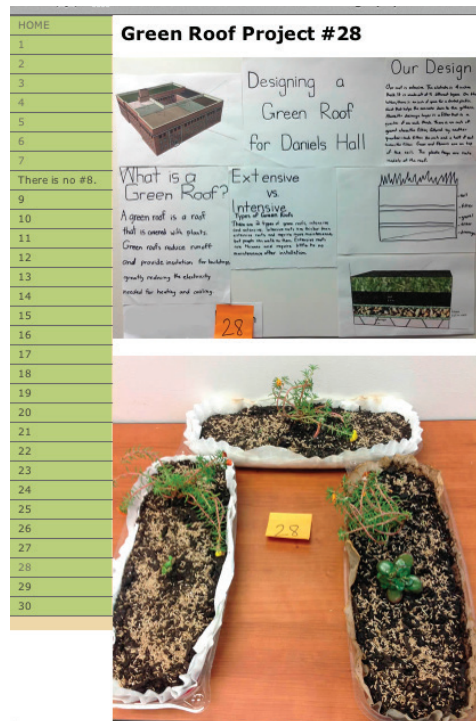


Figure 2. Green roof project website. This figure illustrates the website used by project raters for viewing each of the green roof projects. Photographs of presentation posters and physical models were included for each project.

Consensual Assessment Form (Form A)

**Please view all products before making any ratings.*

**Please rate products relative to each other, rather than to some absolute standard.*

Project Number (1-30):

Overall organization
[The degree to which the work shows good organization.]

9 Best
8
7
6
5
4
3
2
1 Worst

Overall aesthetic appeal
[In general, the degree to which the design is aesthetically appealing.]

9 Best
8
7
6
5
4
3
2
1 Worst

Effort evident
[The amount of effort that is evident in the product.]

9 Best
8
7
6
5
4
3
2
1 Worst

Figure 3. Consensual assessment instrument for iPad. This figure illustrates the interface used by project raters for making online consensual assessment ratings on 12 dimensions of students' projects.

Example Applications

For an example of the interface that the raters were using for assessment, please refer to the following URL: <http://www4.ncsu.edu/~jkbuelin/index.html>.

Please follow the link below for an example of the web-based version of the Consensual Assessment Technique (CAT) for the iPad: <http://tinyurl.com/GreenRoofCAT>.

Procedures

Engineering Summer Camp

Founded in 1999 as an extension of the Women in Engineering Program, the Engineering Summer Camps at North Carolina State University offers week-long day and residential engineering camps each summer for rising 3rd through 12th grade students' interested in experiencing engineering, science and technology. Participants for this study attended a multidisciplinary coed day

camp session for rising 9th and 10th grade students. Student campers paid a fee to participate in the engineering summer camps; however, financial aid was available to those demonstrating need. Approximately 90 students were placed in design teams of three students, providing the study with 30 student groups. The demographic data for the participants were as follows: 63% male, 37% female, 53% Caucasian, 18% African American, 11% Asian, 4% Hispanic, 4% Native American, 6% other, and 3% didn't respond. Participants were not provided remuneration for their participation in this study.

Three secondary school educators, one middle school and two high school teachers with backgrounds in science or math were selected as instructors for the engineering summer camp. Instructors were responsible for 30 students each, equaling 10 student groups. The instructors provided guidance and instruction for the student teams while facilitating the engineering design experience. Six staff camp counselors, undergraduate engineering students, assisted the teacher team leads as mentors and role models to the participants. Six staff high school assistants also supported the engineering summer camp by providing materials and logistical support.

Throughout the week, a variety of hands-on activities were presented, providing a glimpse into the broad scope of opportunities available in engineering. The main weeklong project was the Green Roof Design Challenge, designing an intensive green roof for a campus building that would absorb rainwater, provide insulation for a building, and serve as a beautiful, natural green place that students, faculty, and visitors can enjoy. The project included three steps: (1) Create a very detailed design, complete with technical drawings; (2) create a working scale model of the final design; and 3) prepare a brief 3–5 minute presentation about the design.

In order to complete the project, the campers were provided with the following instructional guidance:

- Learn About Green Roofs
- Substrate Proof of Concept Design
- Test the Substrate Design
- Conceptual Model
- Mathematical Model
- Graphical Model
- Working Model
- Presentation

Fieldtrips to a local arboretum to view plant options and to a nearby building with a working green roof were included in the week of camp.

After receiving their team assignments and a brief introduction to the engineering summer camp, student teams received their green roof engineering design challenge on Day 1 of the 5-day camp. Each day throughout the week, teams participated in ancillary activities designed to promote critical-thinking

and problem-solving skills. These activities included experimentation, analysis, mathematical modeling, and other engineering ways of thinking and doing.

In groups of three, each team was “responsible for defining, developing, and testing a design which takes into account all relevant specifications and constraints” for a proposed green roof on campus. Besides a rooftop schematic, the students were not given any more guidance on the design brief. The design challenge was left ambiguous for the student designers so that they could further formulate the problem, take deeper ownership of the design, engage in questioning, and express creativity.

Additionally, the teams were asked to produce a series of modeling artifacts as part of the design requirements. The models that the teams produced included a conceptual model, a mathematical model, a graphical model, and a working model illustrating their design solution (Lammi & Denson, 2013). The modeling artifacts gave the students something tangible to which they could work while giving the instructors and teaching assistants opportunities to offer concrete feedback and assessment. This design process culminated in team presentations to all camp participants, staff, and students’ families on Day 5.

Following the presentations, photographs of students’ working models and presentation materials were taken. Images were catalogued by project number on a website built for rater access. Once raters were contracted as participants they were given instructions via email as well as the project website URL, and each rater was given a unique CAT survey URL.

Methods

The primary research question for this study was whether the digital interface developed for this implementation of the Consensual Assessment Technique would yield strong ($\alpha > 0.75$) interrater reliability among the seven raters for the 12 dimensions measured. A secondary question concerning the digital instrument’s discriminant validity was also investigated because it is essential to determine whether raters are evaluating creativity apart from other dimensions of projects, such as technical strength and aesthetics.

To secure raters for this study, researchers developed an online solicitation, which explicitly detailed in the criteria that raters needed to be familiar with the engineering design process and experienced in teaching high school aged students. It was important that raters understood the nuances of assessing engineering design products while still understanding the quality of work to be expected from high school age students. Below is the solicitation that prospective raters received:

STEM Education faculty at NCSU request the participation of project raters for an investigation into the assessment of creativity in high school students' engineering design projects.

Raters should be familiar with engineering design processes and should have some knowledge of learners aged 14–17. It is not necessary for raters to have taught high school engineering design in a formal classroom setting.

Ratings will be performed digitally, simultaneously using an iPad and a desktop or laptop computer connected to the Internet. No travel is required for participation; however, an equipped workspace will be provided on the NCSU campus if requested. Compensation of \$50 will be provided for time spent conducting ratings. The estimated time for completion of ratings is approximately 1–2 hours.

The raters included a high school teacher currently teaching Project Lead the Way (PLTW) with over 9 years of teaching experience, a professor with joint appointments in engineering and technology education, a National Board certified science teacher with over 19 years teaching experience, a former engineer and current middle school assistant principal, a high school teacher who has taught at the summer engineering camp for five previous years, an engineering camp director with National Board certification as a science teacher, and a 6th grade science teacher with 13 years teaching experience.

Raters were asked to commit approximately 2 to 3 hours to a rating session during which they would evaluate student projects on dimensions such as creativity, aesthetic value, and technical strength. Raters were compensated with a \$50 honorarium for their participation.

After the camp ended and documentation of student products was organized on the rater website, raters were provided with the URL for the website and a link to the rating form. They were given the following instructions:

Please begin the rating process by reading the problem definition contained in the student's artifacts and viewing the short video on the project landing page. This video is an overview of the images you will find on the website. It serves as an introduction to the products created by the students, and it will give you a sense of the range of abilities represented in the sample. It is essential to our methodology that you look over all the products prior to rating any projects, and that you rate projects relative to each other rather than making ratings based on some absolute standard. In other words, consider what the camp students were able to do given time, instruction, supplies, etc., rather than what you think they should be able to do.

To ensure a consistent rating experience, raters were offered loaner iPads, laptops, and office space in which to conduct ratings if needed.

Assessment of Innovation

To test interrater reliability, Cronbach's alpha was calculated using adult raters' scores for the 12 separate dimensions rated. It can be seen in Table 1 that all 12 items have reliabilities greater than .70 and that ten of the 12 have reliabilities greater than .80. This includes creativity, with an interrater reliability of 0.86. According to the Landis and Koch (1977) scale, a reliability coefficient between 0.61 and 0.80 is "substantial," and agreement above 0.80 is "almost perfect" (p. 165).

Table 1
Cronbach's Alpha for 12 Dimensions Measured

Dimensions of Judgment	Cronbach's α
Creativity	0.8642
Aesthetic Appeal	0.8786
Technical Strength	0.7126
Complexity	0.7818
Liking	0.8557
Novel Idea	0.8453
Novel Use of Materials	0.8808
Shape/Form	0.8422
Color/Value	0.8914
Organization	0.8269
Neatness	0.8149
Effort Evident	0.8387

In order to evaluate the discriminant validity for this implementation of the CAT, factor analysis was conducted on the mean ratings of the 12 dimensions of judgment (promax rotation). Factor analysis suggested the emergence of three factors, as shown in Table 2, corresponding, albeit not perfectly, with Amabile's (1983) paradigm (Figure 1). Although only one factor emerged with an eigenvalue higher than 1.0, consideration of the scree plot (Figure 4) similarly suggests the emergence of three factors, indicated by the rate of change in magnitude of the eigenvalues for Factors 1–3. Factor 1 includes creativity and its three subjacent items: novel idea, novel use of materials, and complexity (as well as liking, effort evident, and technical strength). Factor 2 comprises overall aesthetic appeal and its three subjacent dimensions: pleasing use of color or value, pleasing use of shape or form, and liking (as well as novel use of materials and creativity). Factor 3 includes technical strength and two out of three of its subjacent dimensions: overall organization and neatness. This suggests that the raters were able to distinguish between the features of creativity, technical strength, and aesthetic appeal. The clusters as provided by the factor analysis align very closely with Amabile's (1983) three clusters of

dimension types. This provides strong evidence that the raters were able to distinguish between creative characteristics of design and other characteristics (i.e., aesthetic appeal) of the students' green roof designs. It should be noted, however, that factor analysis is far more stable with larger sample sizes than that of this study; therefore, further testing would be necessary in order to make claims about this instrument's discriminant validity.

Table 2
Factor Loading of 12 Dimensions, Promax Rotation

Dimensions of Judgment	Factor 1: Creativity	Factor 2: Aesthetic Appeal	Factor 3: Technical Strength
Creativity	0.61	0.43	0.04
Aesthetic Appeal	0.12	0.79	0.18
Technical Strength	0.79	-0.23	0.41
Color/Value	-0.13	1.02	0.06
Complexity	0.90	0.01	0.08
Effort	0.58	0.21	0.33
Liking	0.55	0.48	0.07
Neatness	0.04	0.18	0.83
Novel Idea	0.88	0.22	-0.09
Novel Materials	0.52	0.57	-0.05
Organization	0.14	0.20	0.72
Shape/Form	0.19	0.75	0.14

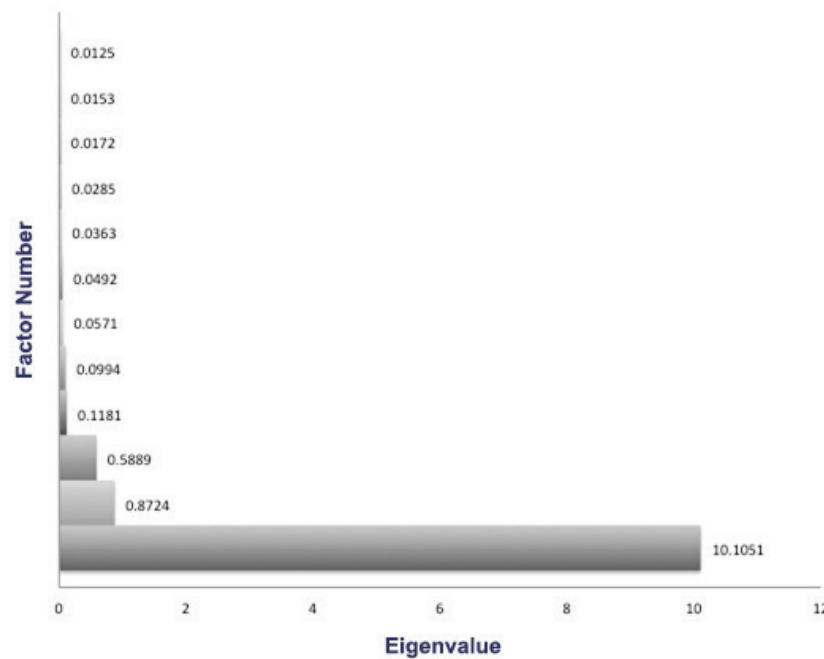


Figure 4. Scree plot for the 12 dimensions measured. This figure illustrates the rate of change in magnitude of the eigenvalues for all 12 components. The slope flattens considerably beyond the third component, suggesting the retention of three factors.

Conclusion and Summary

Despite the skepticism that various stakeholders (e.g., teachers, students, parents, administrators) have been known to display, a growing body of research supports the assertion that creativity can be reliably recognized and assessed in a formal classroom setting. The Consensual Assessment Technique shows promise for the assessment of creativity in the domain of engineering design education. The web-based CAT tools used in this study allow instructors to bypass the limitations posed by implementing consensual assessment in a single physical location. The likelihood of obtaining well-qualified raters is improved, and logistical challenges such as displaying a large number of student projects simultaneously are ameliorated. Using the web-based version of the CAT still produced interrater reliability among the seven raters that was consistently high for all 12 dimensions of judgment measured in this study, and, despite its relative instability with a small sample size, factor analysis suggests that raters were able to recognize and assess creativity apart from other characteristics of

student projects. These findings are important to discussions of how curricula and assessment methods might evolve in engineering design education.

A need for the promotion of creative thinking and innovative problem solving has been identified in the research literature (National Research Council, 2002; Todd & Shinzato, 1999), and the importance of creativity in engineering education has become well documented in recent years (Amato-Henderson, Kemppainen, & Hein, 2011). This study builds upon the work of Amabile (1996), Hennessey et al. (2011), Hickey (2001), and others in confirming that creativity can be recognized by raters who are knowledgeable in a domain and that it can be reliably assessed in the classroom. The promotion of engineering students' abilities to think creatively and to effectively communicate their innovative design ideas is fundamentally important. As these findings add to a research base that continues to show creativity can reliably be assessed, engineering instructors are encouraged to include creativity as an explicit objective in their design challenges.

Overview of Future Work

Further study is needed to develop practical classroom projects and assessment instruments for pre-engineering and engineering students and instructors that will spur students toward meeting their creative potential. One challenge for formal learning environments is that the current system can provide raw scores per dimension and project from the slider scale input. The user is required to download and manipulate raw data, and the mean score (between 1 and 9) does not directly translate to a reportable grade. The development of a streamlined software or website template would be beneficial because this method requires the time, resources, and ability to compile images into an accessible format that is not too cumbersome for raters and it requires familiarity and access to an online survey instrument. The promotion of creativity in engineering design settings still faces many logistical questions as well that has to be addressed. The time and planning needed to secure seven "expert" raters has to be considered. Unlike the researchers in this study, teachers may not have the latitude or budget to pay raters of student projects. In light of these challenges, researchers are encouraged by the preliminary results of assessing creativity in engineering design products.

Larger scale investigation could be useful in exploring potential benefits of self and peer evaluation to student achievement as well as to classroom creativity assessment. Additional investigation is needed into effective methods for training students to act as peer raters. Consistently high levels of interrater reliability found in preliminary cross-domain studies have laid a groundwork for pedagogical investigations comparing, for example, the effects of variables such as design processes, pedagogical strategies, and design prompts on engineering students' creative outcomes. Gender tendencies might also be of interest in similar future studies of larger samples because prior studies have intermittently

shown girls receiving significantly higher creativity scores than boys (Amabile, 1983; Hennessey et al., 2011). Results of this study add to the body of literature on creative assessment through continued research with the engineering summer camp. A future study will investigate the reliability and validity of the digital interface CAT using 144 student participants, which formed 48 student groups. In addition, researchers will investigate students' creative self-efficacy and explore its relationship with creative outcomes as determined by the CAT.

Acknowledgements

We would like to thank Dr. Laura Bottomley for her tireless work, and we would also like to thank the Engineering Place for allowing us the opportunity to work with them on this project.

References

- Amabile, T. M. (1983). *The social psychology of creativity*. New York, NY: Springer-Verlag. doi:10.1007/978-1-4612-5533-8
- Amabile, T. M. (1996). *Creativity in context*. Boulder, CO: Westview Press.
- Amato-Henderson, S., Kempainen, A., & Hein, G. (2011). *Assessing creativity in engineering students*. Paper presented at the 41st ASEE/IEEE Frontiers in Education Conference, Rapid City, SD. Retrieved from <http://fie-conference.org/fie2011/papers/1440.pdf>
- Baer, R. A., Smith, G. T., & Allen, K. B. (2004). Assessment of mindfulness by self-report: The Kentucky Inventory of Minfullness Skills. *Assessment, 11*, 191–206. doi:10.1177/1073191104268029
- Buelin-Biesecker, J. K., & Weibe, E. N. (2013). *Can pedagogical strategies affect students' creativity? Testing a choice-based approach to design and problem-solving in technology, design, and engineering education*. Paper presented at the American Society for Engineering Education Annual Conference & Exposition, Atlanta, GA. Retrieved from http://www.asee.org/file_server/papers/attachment/file/0003/3381/PedagogicalStrategiesCreativity.pdf
- Cropley, A. J. (1999). Definitions of creativity. In M. A. Runco & S. R. Pritzker (Eds.), *Encyclopedia of creativity* (Vol. 1, pp. 511–524). San Diego, CA: Academic Press.
- Dede, C. (2010). Comparing frameworks for 21st century skills. In J. Bellanca & R. Brandt (Eds.), *21st century skills: Rethinking how students learn* (pp. 51–76). Bloomington, IN: Solution Tree Press.
- Fatt, J. P. T. (2000). Fostering creativity in education. *Education, 120*(4), 744–757.
- Gerber, B. L., Cavallo, A. M. L., & Marek, E. A. (2001). Relationships among informal learning environments, teaching procedures and scientific reasoning ability. *International Journal of Science Education, 23*(5), 535–549. doi:10.1080/09500690116971

- Gwet, K. L. (2008). Intrarater reliability. In R. B. D'Agostino, L. Sullivan, & J. Massaro (Eds.), *Wiley encyclopedia of clinical trials* (pp. 1–13). Hoboken, NJ: Wiley. doi:10.1002/9780471462422.eoct631
- Hennessey, B. A., Amabile, T. M., & Mueller, J. S. (2011). Consensual assessment. In M. A. Runco & S. R. Pritzker (Eds.), *Encyclopedia of creativity* (2nd ed., Vol. 1, pp. 253–260). San Diego, CA: Academic Press.
- Hickey, M. (2001). An application of Amabile's consensual assessment technique for rating the creativity of children's musical compositions. *Journal of Research in Music Education*, 49(3), 234–249. doi:10.2307/3345709
- Kaufman, J. C., Baer, J., Cole, J. C., & Sexton, J. D. (2008). A comparison of expert and nonexpert raters using the consensual assessment technique. *Creativity Research Journal*, 20(2), 171–178. doi:10.1080/10400410802059929
- Kotys-Schwartz, D., Besterfield-Sacre, M., & Shuman, L. (2011). *Informal learning in engineering education: Where we are - where we need to go*. Paper presented at the 41st ASEE/IEEE Frontiers in Education Conference, Rapid City, SD. Retrieved from <http://fie-conference.org/fie2011/papers/1235.pdf>
- Kuenzi, J. J. (2008). *Science, technology, engineering, and mathematics (STEM) education: Background, federal policy, and legislative action* (Congressional Research Service Report No. RL33434). Retrieved from <http://www.fas.org/sgp/crs/misc/RL33434.pdf>
- Lammi, M., & Denson, C. D. (2013). *Pre-service teacher's modeling as a way of thinking in engineering design*. Paper presented at the 120th American Society for Engineering Education Annual Conference & Exhibition, Atlanta, GA. Retrieved from <http://www.asee.org/public/conferences/20/papers/5867/download>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1) 159–174. doi:10.2307/2529310
- Lau, S., & Li, W.-L. (1996). Peer status and the perceived creativity: Are popular children viewed by peers and teachers as creative? *Creativity Research Journal*, 9(4), 347–352. doi:10.1207/s15326934crj0904_6
- Lewis, T. (2005). Creativity—A framework for the design/problem solving discourse in technology education. *Journal of Technology Education*, 17(1), 35–52. Retrieved from <http://scholar.lib.vt.edu/ejournals/JTE/v17n1/pdf/lewis.pdf>
- Martin, L. M. W. (2004). An emerging research framework for studying informal learning and schools. *Science Education*, 88(S1), S71–S82. doi:10.1002/sce.20020
- National Research Council. (2002). Equipping the federal government to counter terrorism. In National Research Council, *Making the nation safer:*

- The role of science and technology in countering terrorism* (pp. 335–356).
Washington, DC: National Academies Press.
- NGSS Lead States. (2013). *Next generation science standards: For states, by states*. Washington, DC: National Academies Press.
- Partnerships for 21st Century Skills. (2010). *Partnerships for 21st century skills*.
- Todd, S. M., & Shinzato, S. (1999). Thinking for the future: Developing higher-level thinking and creativity for students in Japan—and elsewhere. *Childhood Education*, 75(6), 342–345.
doi:10.1080/00094056.1999.10522054