Formative Feedback For Improved Student Performance Through Adaptive Comparative Judgment

Dr. Scott R. Bartholomew, Purdue University, West Lafayette

Scott R. Bartholomew, PhD. is an assistant professor of Engineering/Technology Teacher Education at Purdue University. Previously he taught Technology and Engineering classes at the middle school and university level. Dr. Bartholomew’s current work revolves around Adaptive Comparative Judgment (ACJ) assessment techniques, student design portfolios, and Technology & Engineering teacher preparation.

Dr. Greg J. Strimel, Purdue University, West Lafayette

Dr. Greg J. Strimel is an assistant professor of engineering/technology teacher education in the Purdue Polytechnic Institute at Purdue University in West Lafayette, Indiana. His prior teaching experience includes serving as a high school engineering/technology teacher and a teaching assistant professor within the College of Engineering & Mineral Resources at West Virginia University.

Dr. Esteban Garcia
Emily Yoshikawa, Purdue University, West Lafayette
Liwei Zhang

Liwei Zhang is a master student at the Department of Technology Leadership and Innovation at Purdue University. In 2016, she received her Bachelor’s degree in Industrial Engineering from Purdue University, West Lafayette, IN. Her research focuses on assessment in STEM education.
Formative Feedback For Improved Student Performance through Adaptive Comparative Judgment

Scott R. Bartholomew, Purdue University  
Greg J. Strimel, Purdue University  
Esteban Garcia Bravo, Purdue University  
Liwei Zhang, Purdue University  
Emily Yoshikawa, Purdue University

Abstract  
Understanding the best practices of providing, receiving, and improving the formative feedback process in design is critical to improving student creative graphics education. Situated in a university-level computer graphics course this research studied the impacts on student performance of students engaged in adaptive comparative judgment (ACJ), as a formative learning and assessment tool, during several open-ended design problems. Students participated in ACJ, acting as judges of peer work and providing and receiving feedback to, and from, their peers. This paper will examine the relationships between using ACJ and student achievement and will specifically visit the implications of situating ACJ in the midst of an open-ended graphic design project. Further, this paper will explore the potential of using ACJ as a formative assessment and feedback tool.

Key Words: Adaptive Comparative Judgment, Design Graphics, Design Assessment

Introduction  
The current methods to assess design problems and projects are often unreliable and difficult to use in practice due to the creative and ill-structured nature of these problems (Bartholomew, 2017; Pollitt, 2004; 2012). Adaptive comparative judgment (ACJ) is a relatively new assessment approach designed for assessing open-ended problems, such as design portfolios and design projects (Pollitt, 2012). Rather than using rubrics and scoring with a set of criteria like other assessment methods, graders in an ACJ-setting make holistic comparative judgments by comparing two artifacts and picking which one is better. The decision between two items is significantly more reliable for humans to make than a subjective decision based on a predetermined criteria (Pollitt, 2004; 2012). ACJ does not involve judges grading each piece of student work but rather tasks judges with repeatedly making a professional decision on which of two pieces displayed is better. Further, in ACJ as judge’s progress through series of judgments an embedded algorithm iteratively selects similarly-judged pairs of items for the judges to compare until a reliable rank-order of all items is produced. Research around using ACJ, as an assessment tool, has found ACJ to be more reliable than traditional assessment methods (Kimbell, 2012; Pollitt, 2004; 2012) and has demonstrated promise in a variety of grade levels and subjects (Bartholomew & Yoshikawa, under review).

In addition to improved reliability in assessment through the ACJ process, judges also have an opportunity to provide feedback, or a justification around their decision, in conjunction with each judgment they complete. This feedback has the potential to become a valuable source of guidance and direction. Although ACJ was originally designed for summative uses only and as a tool for professional assessors and teachers, this research examined the implications of students acting as the judges of their peers work in the ACJ process formatively. Situated in a
university-level computer graphics course this research investigated the impact on student performance as students engaged in ACJ as judges—both providing and receiving feedback to and from their peers—during a graphics design task. Further, this paper emphasizes a formative-approach to ACJ with students using ACJ at the conclusion of several design projects and both providing and receiving feedback to and from their peers in preparation for future projects. Each of these feedback sessions further acted as a formative-learning opportunity for students as they prepared for their culminating project.

Statement of the Problem

Feedback has been part of the world of art and design for centuries at the ateliers of the craftsmen. The learning process in design scenarios traditionally occurred through the dialogue of apprenticeship (Collins, Brown, & Newman, 1989). In contemporary arts and design curriculum this type of feedback is known as “studio critique” or simply as “critique” (Barrett, 2000). In the past many classrooms have used paper-based assessments to facilitate small group critiques within a larger group (Garcia Bravo, Ashby, & Exter, 2016). Regardless of past success, we sought to investigate ACJ as a potential tool and further analyze student’s perceptions and successes with it. The holistic approach to assessment of ACJ stimulated interesting questions about 2D graphic design as a whole; for example, can we use ACJ formatively to help students improve their designs and expand their awareness of visual design principles? Exploring ACJ was based on an implication that good designs may be a function, based on the consensus, of a group as opposed to the opinions of a few extensively-trained individuals (e.g., teachers).

Further, we felt it necessary to explore the use of ACJ in light of existing issues around scalability and efficiency which limit the feasibility of implementing many approaches formative feedback and assessment (Carless, Salter, Yang, & Lam, 2011). Specifically, we wanted to research if ACJ could potentially be used to alleviate some of these burdens and facilitate a more efficient approach to formative assessment. Therefore, we investigated the implications of using ACJ as a tool for assisting in formative assessment and peer feedback in design problems with the following research questions guiding our study.

Research Questions

RQ1: What impact, if any, does the use of ACJ, as a peer formative assessment tool, have on future student performance on subsequent design projects?

RQ2: What is the relationship, if any exists, between the final ACJ rank-order of student projects and the scores received through traditional scoring approaches?

Graphic Design Project Assessment

Many students are new to the idea that design is a continuous process of iteration and that implementing design projects in problem-based learning scenarios requires students to go through a process of reflection, critique and revision (Larmer, Boss, & Mergendoller, 2015). In this research we specifically introduced students to the concept of re-iteration using ACJ as a tool for peer and self-assessment. We wanted to investigate if ACJ would allow students to experience this continuum formatively through ranking their own, and their peers’, designs. After each ACJ-session, and the accompanying receipt of peer-feedback, students had an additional week to revise their designs before their final submission.
A university-level design course provided the context for this study with students working on creative design projects which allowed them to place personal ownership into their work while also helping them to engage more fully (Nichols, 2016) and tasking them with creatively producing an artistic artifact. The course objectives included combining typography, images, colors and other design elements to create organized visual relationships, adaptive digital images for print, web, animation and software apps, and creating new ideas and formulating them into visual concepts. Superficially, a design project was chose as design projects also have been found to improve student experience and learning (Knight, Carlson, & Sullivan, 2007) while presenting the challenge of following the assignment criteria and also striving to express unique creativity (Laamanen & Seitamaa-Hakkasen, 2014).

The majority of assessments for design projects seem to emphasize a criterion-based approach with most of these assessments revolving around a rubric (Markham, 2011; Panadero & Jonsson, 2013). Despite the popular use of criterion based assessments, these often fall short of correctly measuring the student progress as well as the desired learning objectives (Reynolds, Livingston, Willson, & Willson, 2010). Moore and Fitz (1993) recommended evaluating digital graphic design according to accepted Gestalt principles (e.g., figure-ground segregation, symmetry, closure, proximity, good continuation, and similarity). This process involves “Gestalt,” which is the holistic impression that an image first creates. This is followed by the cognitive process of breaking this first impression of the image down into several other areas for assessment (Moore & Fitz, 1993). As ACJ relies on a holistic comparative judgment we posited that ACJ provided an opportunity of assessing the totality (Gestalt) of each product through a relatively new instructional technique.

**Formative Feedback**

Formative assessment is an approach which allows both students and teachers to evaluate the progress and understanding of the students in situ (Black & Wiliam, 2009). As an important aspect of the learning process in working towards mastery learning is feedback received during formative assessments (Cotton, 2017), this resulting feedback is a key element of formative assessment (Shute, 2008). While a variety of formative feedback practices exist (e.g., teacher inquiry, peer feedback, and student self-inquiry, Luckin, Clark, Avramides, Hunter, & Oliver, 2017) research has demonstrated that this process has a variety of benefits (Ardington & Drury, 2017).

However, despite the positive outcomes that are associated with formative feedback, teachers have continued to struggle with best practices of the formative assessment (Black & Wiliam, 2009). Barriers such as insufficient time and resources, incorrect understandings of feedback, and inability to scale-up the process of both giving and receiving feedback, all hinder teachers abilities to full-utilize formative feedback in education settings. Efforts aimed at addressing these concerns have offered a variety of suggestions including outlines for formative feedback (Gibbs & Simpson, 2005) and clarification to the balance between offering and receiving feedback (Bailey & Garner, 2010). One commonly-used approach to formative feedback which capitalizes on the teacher-student ratio for scaling up the formative feedback approach is to use students for peer formative feedback.

**Peer Formative Feedback.** Peer feedback can be an effective tool which expedites the feedback process while also creating a shared understanding in working towards a common goal (Russell, Van Horne, Ward, Bettis, & Gikonyo, 2017). As students participate in peer feedback processes around a similar assignment studies have shown that they perform better than the
students without this formative practice (Li & Gao, 2016; Li, Liu, & Steckelberg, 2010). Further, students that take part in peer formative feedback at various phases of a project experienced both improved feedback quality and better assignment product scores as a result of their participation (Gielen & de Wever, 2015).

Additional benefits associated with participation in peer formative feedback setting include increased critical thinking (Sluijsmans, Dochy, & Moerkerke, 1998), a better understanding of class material (Stefani, 1994), a more analytical approach towards assignment criteria (Nicol, Thomson, & Breslin, 2014), and increased engagement (Jurado, 2011). In addition to the identified benefits further research has shown that students readily recognize these, and other benefits, to participating in these exercises (Jones & Alcock, 2014).

In addition to the wide variety of benefits associated with peer formative feedback settings there have also been several challenges documented. These include arguments which arise as students have continued to struggle with the credibility of peer feedback over a perceived expert, such as a teacher (McCarthy, 2017) and discrepancies between more extensive and specific feedback received from tutors or teachers - as compared with peers (Hamer, Purchase, Luxton-Reilly, & Denny, 2015). While these findings describe perceptions of students around quality of feedback and the provider, additional studies have shown that the effectiveness of formative feedback may not be based on the quality or characteristics of the feedback, but rather may be dependent on the recipient of the feedback and their reaction (Walker, 2015).

**Methodology**

Eighty-five students from two class sections in an introductory level graphic design course participated in this study. The students were assigned into two groups by enrolled section, with one (N=44) assigned as the “control” group which used traditional paper-based peer critique method, and the other assigned as the “experimental group” (N=41) which used ACJ for peer review. All students in each group completed four design projects and used the same syllabus, course calendar, and were instructed by the same teacher. For each project, students in both groups submitted their first draft for peer assessment (either through paper-based critique or ACJ), and were allowed one week before the final submission to improve their design based on the feedback they got from the in-class peer critique process.

The paper group, who used a traditional paper-based peer critique method, printed out their draft and brought one copy to class. These students were then assigned into groups of five to six and each was given a peer critique worksheet (See Figure 1) and copies of design drafts from another group. To review peer work, students discussed each item as a group, marked points for each project criterion, and left comments and suggestions on the associated-worksheet. Each student received their worksheet with the comments at the conclusion of this exercise and each critique session lasted about one hour.
Figure 1. Paper-based critique

Design Review: Project 1
Reflection and Feedback Sheet

Name: __________________________  Date: __________________________

This project... (mark with and X or √)
Combines shapes, colors and other design elements to create organized visual relationships (Composition)

Uses an adequate image format, color space and resolution (Criteria)

Overall this project is (Rank from 1 to 5)
5= An abstract design that is aesthetically pleasing
4= Looks random and disorganized
3= Creative / Engaging
2= Simple / Boring
1= Looks random and disorganized

Project strengths

Color use is good. Simplicity, Patterns appealing.
- The yellow/orange background is appealing.
- organized

Project weaknesses

Shape placement (green diamond). Colors could be more consistent. Unnecessary shapes (small blue triangles). Tension in center
- Moving the green diamond could make a huge difference.

Which design principles can you identify on this project?

Emphasis, balance, Economy, Harmony, Variety
The experimental (ACJ) group, who participated in ACJ for peer critique, submitted digital copies of their design draft for each assignment. Their designs were uploaded to an ACJ session and all students completed 10 comparative judgments of their peers’ work and left comments and feedback on each item they reviewed (see Figure 2). The ACJ system (called CompareAssess; DigitalAssess, 2017) generated a rank order, Rasch-value misfit statistics for each judge and item, and aggregated all comments associated with each item. Each student received the feedback from their peers, collected through the ACJ session, and was provided time to improve their product.

Figure 2. ACJ critique

At the end of study, a panel of all the course instructors’ completed two ACJ sessions: 1) an ACJ session with student work from both control and experimental groups from the first assignment, and 2) an ACJ session with student submissions from both the control and experimental groups for the final assignment. The first session was done to check for comparability between groups and the second was done to investigate potential impacts on final student performance with relation to ACJ or traditional feedback participation.

Findings

Following the research all data was collected, conditioned, and analyzed in line with the guiding research questions. The findings were separated by research question and will be presented here.

RQ1: What impact, if any, does the use of ACJ, as a peer formative assessment tool, have on future student performance on subsequent design projects?
The first research question investigated how ACJ, used as a learning tool for self and peer design project evaluation, might influence students’ performance. Due to the nature of the course we could not employ a truly experimental design and thus utilized a quasi-experimental research approach. Further, we recognize that a variety of differences could potentially exist between sections include in our research; however, it was not feasible to investigate every potential difference, therefore we emphasized student performance on the first design assignment as an indicator of their graphic design project skill to check for comparability between the control and experimental groups. Using the data, we investigated comparability between the control and experimental sections with the first assignment scores from students in each group and, second, we investigated potential impacts of participation in the intervention on student achievement using the last assignment scores from each student from both groups.

The output from the first ACJ session was a rank-order of all student work from both the control and the experimental sections for the first assignment. To investigate comparability we conducted an independent samples t-test and the results revealed that there was a significant difference in the ranks for the control group (M = 40.98, SD = 27.44) and the experimental group (M = 52.70, SD = 26.53); t(83) = 1.99, p = .05, d = 0.43, with the control group averaging a lower (better) rank than their peers in the experimental group.

Recognizing these differences we conducted another independent samples t-test with the rank order from all the final projects from each section at the conclusion of the course. The results revealed that while students were not comparable to begin the semester (the control group performed significantly better) they were comparable at the conclusion of the semester (d < .005). The control group (M = 44.66, SD = 27.08) and the experimental group (M = 44.78, SD = 24.00) were not significantly different in how they performed on the final assignment; t(83) = .022, p = .98.

These results suggest that the ACJ intervention may have been impactful as the students in the experimental group were initially significantly worse-ranking than their peers but following the intervention these students improved and were not significantly different that their peers at the conclusion of the course. While not conclusive these results suggest positive benefits from using ACJ as a formative tool for learning and assessment in a graphics course.

RQ2: What is the relationship, if any exists, between the final ACJ rank-order of student projects and the scores received through traditional scoring approaches?

Recognizing that ACJ is not an approach which is currently utilized in these courses we sought to identify the relationship, if any, between the rank-order produced through ACJ and the traditional scores assigned by the instructors using rubric-based approaches currently in place. A Spearman correlation between the final rank-order of student work from the final project and the traditional scores assigned by instructors revealed a significant correlation (r = -.65, p < .001). It should be noted that a lower rank is associated with a better product (e.g., rank 1 is better than rank 5) and the negative correlation suggests this inverse relationship between the traditional scores received and the rank-order produced through ACJ.

These findings suggest that the results produced through ACJ are comparable with the results produced through traditional approach to assessing open-ended graphic design projects for students. Thus, ACJ may be a potential alternative for assessment which produces similar results to currently practiced assessment techniques.
Discussion and Conclusion

Fully recognizing the exploratory nature of this project (e.g., two classes at one university) we understand the findings of this research may not be applicable to all settings. However, our results indicate that ACJ may be a feasible tool for improving student performance on design projects. Further, the results from ACJ appear to align well with traditionally-utilized and accepted approaches to assessment currently in place. In addition to the reported findings anecdotally, we noted several benefits to ACJ and its’ use in class. In the process of conducting this study, we observed that students were overall engaged with the ACJ critique. For example, Figure 3 shows students discussing their judgments with their peers while keeping track of the live rankings on the projector screen. Future research will address a qualitative assessment of student’s perceptions on critiques with or without ACJ.

Figure 3. Students watching the rankings emerge as they completed the ACJ session

We also noted the scale-up-ability of ACJ as a means of greatly expanding the feedback for students. We posit that one of the reasons that students were able to improve in the experimental section may be related to the fact that they received significantly more feedback from their peers than with the paper rubrics as a result of the ACJ engine which facilitated an expedited process for providing this feedback. More feedback and an associated-ranking from the session appeared to provide a potentially-solid foundation for students to make decisions (e.g., we noted in observations that students that ranked poorly in the ACJ sessions were more prone to redesign their projects as a result of that ranking). We noticed that ACJ seamlessly expanded students’ abilities to both provide and receive feedback as they were exposed to a variety of student work and their feedback was collected and dispersed through the software tool.

As identifying good design is an important skill for emerging practitioners and evaluating design is an important learning objective which demonstrates students’ awareness of the Gestalt principles, ACJ appeared to enhance this objective as students were exposed to more projects and provided more feedback as a result of their participation than their peers in the control group. Further, we noted that students appeared to enjoy the process of comparing pieces of work and
were influenced by the exposure they received to different ideas through the ACJ process. While students in the control group were also exposed to different ideas their exposure was much less as it was confined to the group they were working in. The benefits of working through the comparisons, exposure to a variety of ideas, and providing/receiving feedback all appeared to be beneficial to the experimental group students and future research into each of these aspects of the ACJ process may shed additional light on each of these opportunities for assisting students.

In light of our experience and findings we recommend ACJ for consideration as a potentially-powerful tool for improving the peer formative assessment process in graphic design settings. Exposure to peer work, increased feasibility in scale-up for both providing and receiving feedback, similar results to traditionally-utilized methods, and significant increases in student achievement following participation all suggest serious consideration for the place of ACJ in design projects and classrooms.
References


Barrett, T. (2000). Studio critiques of student art: As they are, as they could be with mentoring. Theory into Practice, 39(1), 29-35.


