



Evaluating Creativity in Undergraduate Fashion Illustrations: An Instrument Reliability Analysis

Charles Freeman, M.S., Mississippi State University, Starkville, MS and Lisa McRoberts, PhD,
Louisiana State University, Baton Rouge, LA

apparel, assessment, creativity, design

In the introduction to *The International Handbook of Creativity*, Kaufman and Sternberg (2006) credit a lack of methodologically sufficient creativity research to the difficulty of studying creativity and/or acceptance of creativity studies as a fad, lacking scientific rigor. With various complexities arising from creativity research, grounded theoretical application coupled with empirical results contributes to scholarly advances in creative design research. Thus, a component of understanding and developing methodologies of design practice is establishing standards of rigor for evaluation. In a follow-up to a pilot study presented at the 2011 ITAA Conference, the purpose of this investigation is two-fold: (a) to validate the use of a consensual assessment technique to evaluate the creative output of undergraduate design students and (b) to examine the relationships, if any, between thinking abilities and consensual assessment.

Methodology

A convenience sample of 32 undergraduate students enrolled in comparable pattern making courses from two separate southeastern universities was chosen. The Abbreviated Torrance© Test for Adults (ATTA) by Scholastic Testing Services was used to assess overall creativity levels (Creativity Index). Following administration of the ATTA, students completed a design brief within a 45-minute time frame. Modeled after heuristic task criteria outlined by Amabile (1996), instructions stated to create a three piece ensemble from the same landscape photograph provided by the researcher. Providing students with an 11 X 14” sheet of acrylic paper and various art supplies, a facilitator instructed students to “be creative” and feel free to discuss any ideas with your peers. After collecting completed illustrations, three independent judges used the Consensual Assessment Technique (CAT) (Amabile, 1996) to evaluate the creativity levels of the designs illustrated.

The ATTA is an abbreviated version of the original Torrance Test for Creative Thinking© to be used with adults. After completing the pen and paper version, the tests were scored by the researchers based on scoring guidelines provided by Scholastic Testing Service. Each of the participants received a composite score for four creative abilities, as well as an overall creativity index. Based on the creativity index score and creative abilities, the participants’ creativity levels are categorized on a scale ranging from 1 (minimal) to 7 (substantial).

Structured on a Likert scale (1 very low to 7 very high), CAT uses three or more evaluators, with a demonstrated proficiency in a specific domain, to independently assess the creativity and technical qualities of creative projects. Out of 15 total scale items, the seven items measuring the creativity construct are: creativity, novel use of materials, novel idea, effort evident, variation in shapes, detail and complexity. Technical quality measures included: technical goodness, organization, neatness, evidence of planning, balance, representationalism, symmetry and expression. In an effort to eliminate testing effects, each judge received a sealed

packet with instructions for evaluation. Individual measures and the order of evaluations were randomized between judges. Additional criterion for assessment included: independent evaluation, proficiency in the domain, subjective interpretation of the scale items and comparison of the illustrations against others in the sample versus an industry standard or personal ideal.

Results and Discussion

Creativity levels of the 32 undergraduate students were normally distributed ranging from minimal (creativity level = 1) (n = 3) to substantial (creativity level = 7) (n = 3), with a mean of 4.03. Over half of participants ranked from below average to above average. Spearman correlations of creativity levels and creative product assessments indicate no significant relationships between creativity index rankings based on the ATTA and the overall summation scores on the consensual assessment by judges. Similar to the pilot study reported on in 2011, these results support earlier research conducted by Torrance (1962) and Amabile (1996), suggesting the inadequacy of a divergent thinking skills test to predict future creative production or performance.

Treating evaluators as items (often done in creativity research), coefficient alphas were used to determine levels of interrater reliability. Alpha results of independent evaluators using the consensual assessment technique were acceptable ($\alpha = .79$) for items measuring creativity. Conversely, on measures of technical quality interrater reliabilities decrease ($\alpha = .55$). Of the items evaluated, variation in shapes ($\alpha = .72$), details ($\alpha = .72$) and creativity ($\alpha = .70$) showed the highest reliabilities. Technical quality measures including organization ($\alpha = .27$), evidence of planning ($\alpha = .17$) and representationalism ($\alpha = .12$) represented the lowest interrater reliabilities.

Based on these findings, the consensual assessment technique provides apparel educators and scholars an effective tool for evaluating creativity of undergraduate student projects using a panel of judges, despite low reliabilities on measures of technical quality. While judges may come to a consensus on creativity of a project, technical merit and execution remain divisive factors in evaluation. Amabile (1996) argues the validity of the consensual assessment technique can only be maintained by the inclusion of technical measures, despite a lack of agreement, in order to prevent technical factors affecting creativity judgments.

Creativity often occurs within a contextual situation with many outside factors including motivations, environmental conditions, and domain relevant skills influencing creative production for evaluation. The researchers suggest further investigations into additional factors influencing creativity evaluation using a consensual assessment technique, including professional projects, increased number of evaluators and a comparison between expert and novice judges.

References

- Amabile T. (1996). *Creativity in Context*. Boulder, CO: Westview Press.
- Kaufman, J. C., & Sternberg, R. J. (Eds.). (2006). *The International Handbook of Creativity*. Cambridge, NY: Cambridge University Press.
- Goff, K. & Torrance, E.P. (2002). *The Abbreviated Torrance Test for Adults (ATTA)*. Bensenville, IL: Scholastic testing Services.
- Torrance, E.P. (1962). *Guiding Creative Talent*. Englewood Cliffs, NJ: Prentice-Hall.