

A DigitalCAT: A Fusion of Creativity Assessment Theory and HCI

Dr. Genevieve M. Cseh, Karl K. Jeffries
University of Central Lancashire
School of Art, Design & Fashion
Victoria Building
Preston PR1 2HE
GMCseh@uclan.ac.uk,
KJeffries@uclan.ac.uk

Dr. Mark Lochrie, Prof. Paul Egglestone
University of Central Lancashire
Media Innovation Studio
Media Factory
Preston PR1 2HE
MLochrie@uclan.ac.uk
PEgglestone@uclan.ac.uk

Allan A. Beattie
University of Aberdeen
Web Team
Edward Wright
Aberdeen AB24 3QY
allan.beattie@abdn.ac.uk

As a 'gold standard' creativity assessment method, it is important to reflect on the digital future of Amabile's Consensual Assessment Technique (CAT). Over the past three decades, the CAT has given creativity researchers a formal tool on which to build a true science out of the subjective phenomenon of creativity in a vast range of domains. Research practice, however, has relied on primitive paper-based tools or only rudimentary digital technology. As a result, it is high time a more sophisticated, standardized research tool is developed to greatly facilitate future creativity research and assessment – a DigitalCAT – building on expertise from the design research, psychology and human-computer interaction (HCI) disciplines.

Consensual Assessment Technique. Creativity. Digital. Tinder. Interface. Social.

1. INTRODUCTION

In creativity research, one of the most established methods of measuring creativity is the *Consensual Assessment Technique* (CAT: Amabile, 1982; 1996), and after over 30 years of its use, it is high time for creativity researchers to reflect on its future, particularly within a digital era.

2. BACKGROUND

Since the 1950s, a number of methods have been used to assess creativity (Batey & Furnham, 2006), but the consensus opinion of domain expert judges (using their subjective opinions based on tacit knowledge and expertise) is arguably the most common, holistic, and ecologically valid assessment method available (Baer & McKool, 2009). It is also highly reliable when procedures are carefully controlled and standardized - i.e., expert judges tend to agree well on what is creative or not in their field (Amabile, 1982; 1996). Certain guidelines must be followed – e.g., judges must be domain experts, items to be rated must be rated relative to one another on a scale, and no prior training or instruction can be given to judges. However, beyond these basic rough guidelines, there is no official strategy for implementation. The CAT can be applied to any setting and has been applied to settings from the arts to the military. However, the adaptability of the measure is also one of its

weaknesses, as this has led to inconsistency and corner-cutting in its application, compromising the integrity of creativity research. A single digitised CAT accessible to all would help to standardise the CAT and improve research overall.

Research practice to date has tended towards primitive paper-based tools or only very limited digital technology. Thus, it requires physical administration, manual data collection and input, and its application is often tedious and burdensome for both judges and researchers, discouraging busy industry practitioners from participating in academic research or educational assessment.

In practice, the traditional CAT procedure involves judges independently viewing either physical paper copies/cards, or at best a PDF slide presentation of a series of items to be assessed (e.g., artworks), and familiarising themselves with the items prior to making relative comparative judgements about their creative merit on a scale, on paper.

Jeffries (2015), in applying the CAT to the assessment of graphic design creativity, evolved the procedure by giving judges a laminated A3 rating sheet and a set of miniature thumbnail cards of the artworks they had viewed. Cards were physically placed on the rating sheet, and judges proceeded to sort the artwork visually: i.e., first into low, medium and high, followed by more refined rating of works from one to six (high-low within each category). This

clarified and visualised the rating process for judges and allowed for a freeform tactile experience. Prior to this, most instructions to judges were text-based.

Building on expertise in design research, psychology and HCI, the authors suggest it is time to develop a more sophisticated research tool: a DigitalCAT.

3. DIGITALCAT

A DigitalCAT is envisioned as a cross-platform application that would allow judges to directly record ratings for each artwork via tablet, mobile, laptop or desktop. Judges' rating data would be captured in a database where a report is generated on collective ratings. This can be fed into other systems as required (such as SPSS statistical analysis software or Excel) via an Application Program Interface (API).

The first known prototype DigitalCAT was built by web developer Allan Beattie at the University of Aberdeen as part of the first author's PhD (Cseh, 2014; Fig. 1).

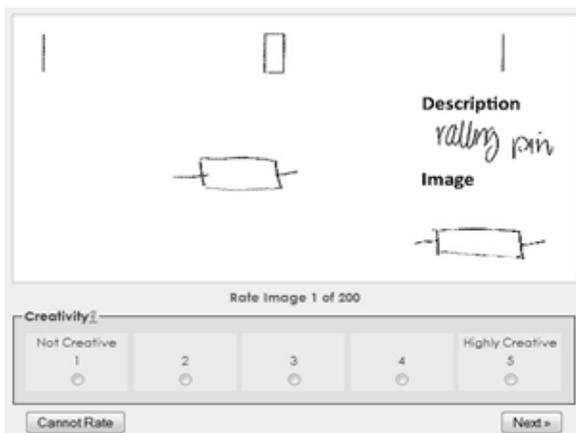


Figure 1. First DigitalCAT prototype web rating platform, built by Allan Beattie at the University of Aberdeen (Cseh, 2014).

This system allowed judges to remotely rate artworks created as part of psychological research, using a web-based form rating system connected to a PHP/MySQL database. The survey-style interface was simple, effective and allowed judges to rate at a time and from a place of their choosing, allowing for much easier recruitment and data collection as no lab time was required, and data input was automatic. However the setup of each CAT study required considerable developer input, and users were limited in how much they could manipulate the images to be rated. The next stage of development aims to truly create a WYSIWYG content management system to enable easier back-end setup and front-end user interaction.

Currently, funding is being sought to develop this DigitalCAT further, with a series of user testing workshops planned to beta test a new prototype.

Regarding software design plans, the new DigitalCAT will more fully explore user interface opportunities offered by digitisation. Moving to digital allows for creative new interfaces: e.g., design interfaces inspired by current social platforms such as Facebook, game interfaces and dating applications like Tinder, coupled with related research from within UX/UI studies. For example, the potential of the swiping motion of Tinder could be adopted to traverse large numbers of items in a relatively short space of time, and the voting/liking/reacting application of Facebook applied to rating procedures. Large click areas and zoom capabilities could be used to improve usability, especially on small mobile devices, allowing judges to sort images as in Jeffries' graphical paper-based rating system. Potentially when a judge swipes up, across or down, this translates into high, medium and low respectively. Upon performing the gesture, a further popup is presented to capture finer details of the sort, such as the top or bottom of that category, or a drag-and-drop option for greater interactivity. There is potential for a new social component as well, allowing judges greater instant feedback on completion, to view how other judges have rated the same works, which may increase engagement and interest.

To ensure all items are fairly rated, items would be randomized each time the application is loaded. This is an important consideration to CAT scientific protocol, and relatively straightforward to achieve digitally, but a significant practical burden with current paper-based tools. The DigitalCAT will need to be responsive to and universally accessible on a variety of platforms/devices. A consideration which must be carefully tested is the ability to adequately view and make comparisons between large datasets on commonly-used devices such as small mobile phone screens. Future research would be required to ensure that the DigitalCAT is as natural to use as the physical paper-based task, conserving or possibly even improving the validity and reliability of the measure with the move to digital.

4. CONCLUSIONS

The potential benefits of a DigitalCAT are considerable compared to current paper-based tools. Digitisation eases the burden of data collection and entry for researchers, and ensures that CAT research is released from the confines of geography, office working hours and physical attendance by judges, thus minimising practical barriers to participation in research and educational/workplace assessment. A DigitalCAT would help standardize and control methodology, increasing the scientific integrity of creativity research based on CAT results. Thus digitisation of the CAT has the potential to greatly improve both academic and practical creativity assessment.

3. REFERENCES

- Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, 43(5), 997–1013. doi:10.1037//0022-3514.43.5.997
- Amabile, T. M. (1996). *Creativity in context*. Boulder, CO: Westview.
- Baer, J., & McKool, S. S. (2009). Assessing creativity using the Consensual Assessment Technique. In *Handbook of research on assessment technologies, methods, and applications in higher education* (pp. 1–13). doi:10.4018/978-1-60566-667-9.ch004
- Batey, M., & Furnham, A. (2006). Creativity, intelligence, and personality: A critical review of the scattered literature. *Genetic, Social, and General Psychology Monographs*, 132, 355-429.
- Cseh, G. M. (2014). *Flow in visual creativity*. Ph.D. Dissertation. University of Aberdeen, Aberdeen, Scotland. Advisor(s): David G. Pearson, Louise H. Phillips. uk.bl.ethos.641464
- Jeffries, K. K. (2015). A CAT with caveats: Is the Consensual Assessment Technique a reliable measure of graphic design creativity? *The International Journal of Design Creativity and Innovation*. doi: 10.1080/21650349.2015.1084893