
Evaluating Collaborative, Creative Contributions

Jonas Frich

frich@cc.au.dk

Centre for Digital Creativity

Aarhus University

Aarhus, Denmark

ABSTRACT

Throughout the last 70 years, it has been the main goal for creativity research to establish a reliable measure of creativity, which would lend itself to formal evaluation. From Guilford's psychometric methods in the 1950s over Amabile's domain dependent consensual methods in the 1980s to today's automated semantic-based measures fueled by the improvement of neural networks and computational power, this problem continues to be at the centre of creativity-related research. This paper presents research on how to evaluate creativity and presents suggestions for future computer-based evaluation forms.

KEYWORDS

Creativity, evaluation, collaboration, CST

ACM Reference Format:

Jonas Frich. 2019. Evaluating Collaborative, Creative Contributions. In *CHI'19, Extended Abstracts, May 04–09, Glasgow, Scotland UK*. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI'19, Extended Abstracts, May 04–09, 2019, Glasgow, Scotland UK

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...\$15.00

<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

INTRODUCTION

"If you can't measure it, you can't improve it." - Perhaps this famous quote by Peter Drucker explains why evaluating creativity has been so central to the study of creativity throughout the last 70 years. While improvements in creative performances may yield significant impacts in terms of solving everything from the UN's 17 sustainable development goals, to developing the next new iPhone or writing the next best-seller, these rarely serve well as measures in a scientific context. Firstly, success is determined historically by peers, the marketplace, and society as a broad, which makes it rather impractical considering both academic life-cycles and tool development pace, and secondly, the sheer number of influencing variables makes it almost impossible to determine cause and effect. As a consequence of this, we need proxies to tell us something about the creative performance.

Rhodes influential four Ps of creativity model [11] highlights the *product*, *process*, *press* and the *person* as the cornerstones of creativity research, and these may also be perceived as four relevant proxies for evaluating collaborative, creative contributions from a crowd. However, many of the original evaluations were conceived in a time where creativity research either focused on gifted individuals or the underlying cognitive processes of a single person [12], and may not be fitting for the use in the context of crowd-sourced creativity. Perhaps we should reconsider how to further develop these methods to: **evaluate the quality of contributions from the crowd?**

Using the four Ps, this paper will present relevant ways of evaluating creativity before moving forward with a discussion of how some of these methods may be appropriated or improved to cater to crowd-based creativity. The work in this paper builds on both practical experience with conducting empirical creativity-related HCI research and previous reviews of creativity-related HCI literature [5] and Creativity Support Tools [6] completed by my colleagues and me at the Centre for Digital Creativity, Aarhus University. Finally, the paper will finish with three points for future crowd-powered creativity support systems, which could be relevant for discussion in the workshop.

TYPES OF EVALUATIONS

Product

"When an idea becomes embodied into tangible form it is called a product" [x] and since creativity is often defined within a specific domain, one way of evaluating the final product is to have it assessed by expert judges within the given field. This idea was formalized by Amabile in 1982 when she first introduced the Consensual Assessment Technique (CAT) which is now well-established as a powerful tool in the field. In its basic form, it relies on a panel of judges with domain expertise rating the creativity of a given outcome or product, with an emphasis on the expert part [9]. The CAT has been referred to as the *gold standard* for evaluating creativity by e.g. [2, 9], because it measures the actual creative performance or outcome rather than possible traits, skills or processes correlated with

creativity, and because it is so well validated. It does, however, come with certain drawbacks as clearly stated by Bear & McKool: "It is very resource intensive: assembling groups of expert judges are not simple and it may be expensive." [2].

Process

In this case, the notion of process refers to perception, learning, creative thinking abilities, and communication. However, understanding the unfolding creative process through protocol analyses [7] is not relevant in the context of crowd-sourced creativity. Alternatively, the use of divergent thinking tests is mentioned by Amabile [1] as a possible method for assessing *creativity-relevant processes*, denoting the mental processes or styles of thinking. Torrance Tests of Creative Thinking [13] or Remote Associates Tests [10] are two examples of such tests, which basically measure the ability to come up with many different responses. A recent contribution to this discussion is the introduction of Multi-Trial Creative Ideation, which reinterprets the concept of fluency using the time while capturing individual dynamics in the ideation process [3]. The premise of this new approach is to invoke digital tools in a series of Alternate Use Tasks with predefined stimuli, allowing for automatic collecting of process-indicators and subsequent log-analysis of *exploration*, *production* and *verification* [3].

Press

Press plainly refers to the environment in which any given creative production occurs (incl. the working context and the available tools). While e.g. Amabile has been conducting research on creativity in organizational contexts for years [1], the more relevant contribution for this discussion is that of the creativity support tools. Building on i.a. Csikszentmihalyi's elements of flow and Shneiderman's design principles for CSTs Cherry & Latulipe [4] developed a solid measure for how well a digital tool supported creative in general and in terms of *Enjoyment*, *Expressiveness*, *Immersion*, *Collaboration*, *Exploration* and *Results Worth Effort*, so to enable potential attention to be devoted to further developing support in either area.

Person

Personal traits like temper, intelligence, and openness to experience are obvious parameters to consider alongside domain-relevant knowledge. However, one important feature of crowd-based ideation is to bring a group of people with diverse backgrounds together, and sorting for e.g. expertise, personality or domain-relevant knowledge might thus be counterproductive. A recent measure of called *Forward Flow*, which basically quantifies how much current thoughts depart from previous thoughts within free association using latent semantic analysis, has been shown to be positively associated with creativity [8]. The positive association was found across multiple studies and persisted even when controlled for cognitive capacity [x].

DISCUSSION

Evaluating the quality of contributions from the crowd is no simple matter, and while the *Consensual Assessment Technique* could be suited to develop a baseline for evaluating contributions up against, the scale at which crowd-based ideations occur might render it impractical. *Multi-Trial Creative Ideation* may, however, present a possible way forward by introducing quantifiable process-indicators permitting automation, such as the amount of exploration or production. Likewise, the *Forward Flow* might hold some potential in either term of e.g. screening the individual person in the crowd-based ideation using a rather low-intrusive method. The *Creativity Support Index*, on the other hand, might be used to examine the tools used by both the crowd and the individual at the 'end' of the pipeline, although it is perhaps tailored to a different 'type' of tool (as it is used to evaluate i.e. Google Docs.).

ACKNOWLEDGMENTS

This research has been funded by The Velux Foundations grant: Digital Tools in Collaborative Creativity

REFERENCES

- [1] Teresa M Amabile. 1996. *Creativity in context: Update to the social psychology of creativity*. Hachette UK.
- [2] John Baer and Sharon S McKool. 2014. The gold standard for assessing creativity. *International Journal of Quality Assurance in Engineering and Technology Education (IJQAETE)* 3, 1 (2014), 81–93.
- [3] Baptiste Barbot. 2018. The Dynamics of Creative Ideation: Introducing a New Assessment Paradigm. *Frontiers in psychology* 9 (2018), 2529.
- [4] Erin Cherry and Celine Latulipe. 2014. Quantifying the creativity support of digital tools through the creativity support index. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21, 4 (2014), 21.
- [5] Jonas Frich, Michael Mose Biskjaer, and Peter Dalsgaard. 2018. Twenty Years of Creativity Research in Human-Computer Interaction: Current State and Future Directions. In *Proceedings of the 2018 Designing Interactive Systems Conference*. ACM, 1235–1257.
- [6] Jonas Frich, Lindsay MacDonald Vermeulen, Christian Remy, Michael Mose Biskjaer, and Peter Dalsgaard. 2019. Mapping the Landscape of Creativity Support Tools in HCI. In *Accepted in Proceedings of the 2019 CHI*.
- [7] John S Gero and Udo Kannengiesser. 2004. The situated function–behaviour–structure framework. *Design studies* 25, 4 (2004), 373–391.
- [8] Kurt Gray, Stephen Anderson, Eric Evan Chen, John Michael Kelly, M Christian, John Patrick, and K Lewis. 2019. Forward flow: A new measure to quantify free thought and predict creativity. *American Psychologist* (2019).
- [9] James C Kaufman, John Baer, Jason C Cole, and Janel D Sexton. 2008. A comparison of expert and nonexpert raters using the consensual assessment technique. *Creativity Research Journal* 20, 2 (2008), 171–178.
- [10] Sarnoff A Mednick. 1968. The remote associates test. *The Journal of Creative Behavior* 2, 3 (1968), 213–214.
- [11] Mel Rhodes. 1961. An analysis of creativity. *The Phi Delta Kappan* 42, 7 (1961), 305–310.
- [12] R Keith Sawyer. 2011. *Explaining creativity: The science of human innovation*. Oxford University Press.
- [13] E Paul Torrance. 1972. Predictive validity of the Torrance tests of creative thinking. *The Journal of creative behavior* 6, 4 (1972), 236–262.