12-2-2023

Is a Fool With a(n AI) Tool Still a Fool? An Empirical Study of the Creative Quality of Human–AI Collaboration

Sebastian Weber
University of Bremen, Germany, sebweber@uni-bremen.de

Bastian Kordyaka
University of Bremen, Germany, kordyaka@uni-bremen.de

Raphael Palombo
University of Bremen, Germany, palombo@uni-bremen.de

Dominik Siemon
LUT University, Finland, dominik.siemon@lut.fi

Bjoern Niehaves
University of Bremen, niehaves@uni-bremen.de

Is a Fool With a(n AI) Tool Still a Fool? An Empirical Study of the Creative Quality of Human–AI Collaboration

Research in progress

Sebastian Weber
Faculty 3 - Mathematics and Computer Science
University of Bremen, Germany
Bremen, Germany
Email: sebweber@uni-bremen.de

Bastian Kordyaka
Faculty 3 - Mathematics and Computer Science
University of Bremen, Germany
Bremen, Germany
Email: kordyaka@uni-bremen.de

Raphael Palombo
Faculty 3 - Mathematics and Computer Science
University of Bremen, Germany
Bremen, Germany
Email: palombo@uni-bremen.de

Dominik Siemon
School of Engineering Science
LUT University, Finland
Lahti, Finland
Email: dominik.siemon@lut.fi

Bjoern Niehaves
Faculty 3 - Mathematics and Computer Science
University of Bremen, Germany
Bremen, Germany
Email: niehaves@uni-bremen.de

Abstract

Recent computational advancements in artificial-intelligence (AI)-based systems have pushed forward the frontiers of creativity. The collaboration between humans and AI promises to unleash new creative potential. Accordingly, the proposed study aims to provide a better understanding of what is needed to accept generative AI as a co-creative partner by researching the necessary skills via the collaboration between humans and AI. The cross-sectional study will examine the interaction effect of domain knowledge and prompt engineering skills with regard to the creative quality of human–AI-generated content. By researching these relationships, the study will contribute to the literature on human–AI collaboration and creativity by broadening our understanding of how prompt engineering skills and domain knowledge affect the creative quality of human–AI-generated content. The findings will have practical implications for adopting AI as a co-creative partner and should provide a more nuanced view of the concept of human-AI-based creativity.

Keywords: creativity, artificial intelligence (AI), generative AI, human–AI collaboration, prompt engineering

1 Introduction

The human-like perception of machine output has long been relevant in many fields, ranging from computer science, philosophy, and psychology to information systems (IS) research (Cole 2004; Schuetz and Venkatesh 2020; Schuetzler et al. 2021). In this line of research, tests have also been proposed for deciding whether machines (i.e., artificial-intelligence-based systems) exhibit the capabilities to mimic humans. Such tests include the Turing Test, proposed by Turing (1950), and the Lovelace Test, proposed by Bringsjord et al. (2003). By mimicking how humans process and interpret information, artificial-intelligence (AI)-based systems are designed to better understand and respond to the needs and wants of humans. One crucial example stems from the domain of creativity. Indeed, creativity has been one of the most important skills in the development of humankind. Creativity represents the pillar of progress, and it is hence crucial to discover how far AI-based systems can go with regard to creative output generation.

In this line, AI-based systems are mostly perceived as being limited in their ability to fully process and interpret the world around them, and they lack the necessary skills to make decisions and judgments based on their experiences alone when tested on their ability to successfully mimic human capabilities (Hulman et al. 2023; Powell 2019). However, AI-based systems – which are described as systems that implement recent computational advancements that reflect human intelligence (Berente et al. 2021) – are now expanding the frontiers with regard to capabilities such as creativity. Current advancements in the field of generative AI can transform text-to-text, text-to-audio, text-to-image, text-to-video, text-to-code, or text-to-3D (GitHub Inc. 2022; OpenAI 2022; Poole et al. 2022). After input has been provided, these AI-based technologies are capable of generating convincing output. For instance, an artist recently won an art contest using the generative AI program Midjourney (Roose 2022). This event highlights the fact that a symbiotic collaboration between humans and AI can unleash new creative potential. However, while the artist in this case did have knowledge in his domain, in the future, it will become increasingly common for several jobs to be taken over by one person, even when this person does not have expertise everywhere.

It is hence essential to grasp the requirements for successful collaboration with generative AI in the realm of creative endeavors. On the one hand, domain knowledge is paramount to producing meaningful creative output in human–AI collaborations (Baer 2015). A deep understanding of the subject matter ensures that the AI's responses are both contextually relevant and accurate. On the other hand, prompt engineering skills are equally vital and could even mitigate lacking domain knowledge. Prompt engineering refers to the formal search for prompts that elicit creative output from language models for a specific task and represents both a burgeoning field and an invaluable skill (Liu and Chilton 2022). With today's rapid advancements and the vast array of applications for generative AI, users are simultaneously creators and consumers. The fast and broad-based adoption of this technology – as can be seen, for example, in the 100 million active users of ChatGPT within just two months of its debut (Hu 2023) – underscores the significance of the technology. In terms of crucial dependent variables, such as long-term usage and successful human–AI collaboration, both domain expertise and proficient prompt engineering skills play a central role in perceiving AI as a co-creative partner.

The proposed study therefore aims to shed light on the acceptance of generative AI as a co-creative partner by contributing to two research gaps: a) the impact of domain knowledge on creative content quality produced with generative AI and b) the ability of prompt engineering skills to mitigate lacking domain knowledge. For this purpose, we will conduct a cross-sectional study that researches the impact of prompt engineering skills and domain knowledge on the creative quality of textual content. Our results will shed light on the following two research questions (RQs):

> RQ1: How crucial is domain knowledge when it comes to producing creative output through human–AI collaboration?
>
> RQ2: How does the level of prompt technical skills affect creative output for people with lower levels of domain knowledge versus people with medium or higher levels of domain knowledge?

By answering these RQs, we will contribute to research on human–AI collaboration in creativity-based tasks (Feuerriegel et al. 2023) both in the IS domain and beyond. This paper is structured as follows: First, in the section on related work, we introduce AI-based creativity and human–AI collaboration to derive hypotheses. Then, we describe the methodology, including the procedure, measurements, data collection, and data analysis that we will use. Finally, the paper closes with a short outlook section.

2 Related Work

2.1 Creativity & AI

Creativity plays an important role in innovation and competitiveness. It is a critical skill for individuals and organizations to develop in order to adapt to changing environments and to stay ahead of the competition. Creativity can lead to new discoveries, which can improve the quality of life for individuals and society as a whole. Hence, it is no wonder that technology-supported creativity has gained momentum, which is also evident in the literature on creativity support systems (Wang and Nickerson 2017). However, recent AI-based systems capabilities represent a paradigm shift in technology-enabled creativity (Runco 2023a, 2023b). For instance, the Lovelace Test (a modified version of the Turing Test) was designed to evaluate the creativity and originality of computer programs in terms of the ability of AI to mimic human creativity. The test was proposed by Bringsjord et al. (2003) as a way of measuring the level of creative intelligence in a computer program. Against the background of the recent advancements in generative AI, a new version was also proposed (Riedl 2014). In general, the Lovelace Test aims to determine whether a computer program can create something new beyond what it has been explicitly programmed to do. The test does so by requiring the program to generate a creative artifact, such as a piece of art or text, which is then evaluated by human judges. The judges rate the artifact in order to determine whether it satisfactorily meets all the predetermined criteria. In this regard, no specific details on the criteria for a standardized evaluation yet exist. However, the creativity literature has already elaborated on this issue by defining creativity and proposing evaluation criteria (Dean et al. 2006; MacCrimmon and Wagner 1994; Plucker et al. 2004; Runco 2023a, 2023b). With regard to the definition, creativity has been defined as requiring both originality and effectiveness (i.e., ideas should be novel and useful) (Runco and Jaeger 2012). Recent research suggests that with the advent of AI, a further distinction should be made between artificial creativity and human creativity (Runco 2023a, 2023b) because the creativity of AI is viewed as inauthentic and perhaps also unintentional.

Yet, with regard to evaluating creative output, the dimensions of creativity remain the same. Dean et al. (2006) proposed measuring the creativity of ideas in four dimensions that reflect the term's definition: novelty, feasibility, relevance, and specificity. Novelty refers to the originality or uniqueness of an idea or solution, feasibility refers to easy implementation without violating any constraints, relevance refers to usefulness in solving a problem, and specificity refers to the degree to which an idea is worked out in detail. In this context, research has tested the performance of generative AI, with mixed findings (Chakrabarty et al. 2023; Guzik et al. 2023; Koivisto and Grassini 2023). On the one hand, results indicate that generative AI possesses the ability to generate new, unique, and unexpected ideas that match or exceed levels of human creativity (Guzik et al. 2023; Koivisto and Grassini 2023), but on the other hand, humans remain superior in terms of creative output (Chakrabarty et al. 2023; Koivisto and Grassini 2023). In conclusion, studies have thus far only researched the creative performance of generative AI in comparison with humans with mixed results and have neglected human–AI collaboration, which is the current modus operandi. Nevertheless, all existing studies that have examined AI vs. humans have required humans to provide the initial prompt for the AI. Therefore, it is important to also investigate the decisive factors that enable AI to produce creative output (in cooperation with humans).

2.2 Human–AI Collaboration in Creativity-Based Tasks

Collaboration is a dynamic and interactive process in which multiple parties actively participate in shared activities in order to achieve one or more common objectives (Bedwell et al. 2012). Human–AI collaboration specifically refers to the partnership between humans and AI-based systems, with the two entities working together as "teammates" to solve problems as opposed to solely automating routine tasks, as was done in the past (Lai et al. 2021). With the rise of AI (Schuetz and Venkatesh 2020), human–AI collaboration is an emerging area of research that has gained significant attention in recent years, particularly in the field of IS and human–computer interaction (Dellermann et al. 2019; Wang et al. 2020). The idea behind human–AI collaboration is to leverage the complementary strengths of humans and AI in order to achieve better results than either entity could achieve alone (Daugherty and Wilson 2018). In our context of producing creative output with generative AI, this concept is particularly relevant because recent technological advances offer strong potential to augment humans in this context. With generative AI, humans have the role of providing input, modifying this input, and evaluating it. In this line, prompt engineering skills and domain knowledge with regard to the task are crucial variables that should be scrutinized with regard to creative output.

Domain Knowledge. In the realm of IS research, the rise of generative AI for specialized tasks is evident. The effectiveness of AI – particularly in tasks that demand creativity – is intrinsically linked to domain knowledge. Not only is this knowledge foundational for guiding data input, but it is also critical in refining, modifying, and evaluating the output that the AI systems produce. The essence of domain-specific knowledge extends far beyond the confines of AI applications and is a cornerstone in the broader field of creativity research. Exemplary studies by Baer (2012), Feldhusen (2010), Huang et al. (2017), Mayer (2010), and Silvia et al. (2009) have all emphasized the profound impact that domain knowledge has on enhancing creativity. Moreover, research has unveiled discernibly low inter-correlations in creativity ratings across varied domains, which suggests that domain knowledge primarily magnifies creativity within its inherent domain (Baer 2015). As the frontiers of generative AI expand and we continue to harness their capabilities, the use of domain expertise in user interfaces with these systems is becoming increasingly crucial in shaping the trajectory of outcomes and ensuring alignment with domain-specific nuances and requirements. However, AI could serve as a buffer against the limitations posed by lacking domain knowledge. Advanced AI-based systems that are equipped with vast datasets and sophisticated algorithms can access and process information far beyond human capabilities (Schuetz and Venkatesh 2020). In certain situations, this ability could compensate for the user's lack of domain expertise. By suggesting a context, offering insights, or even simulating potential outcomes based on patterns from vast datasets, AI can provide users with a scaffolded experience, thereby bridging gaps in these users' domain knowledge. Moreover, the iterative feedback loop between users and AI can serve as a learning mechanism (Grassini 2023) whereby users gradually augment their domain understanding through constant interaction. Hence, we hypothesize:

> H1: Domain knowledge should be found to be positively associated with the creative quality of the output.

Prompt Engineering Skills. Providing input and modifications is referred to as prompt engineering. More specifically, this practice can be defined as the formal search for prompts that retrieve desired output from language models for a given task (Liu and Chilton 2022). Researchers and practitioners alike now tackle the open problem of prompt engineering for large pre-trained models. Most work in prompt engineering has concentrated on the text generation problem from natural language processing. The term *prompt engineering* originally came from a popular online post about GPT3 (a large language model) and its capabilities in writing creative fiction (Gwern 2020). Gwern (2020) suggested that prompt engineering models could become a new paradigm of interaction in which users would only need to discover how to prompt a model in order to elicit the specific knowledge and abstractions that are necessary for completing tasks. While momentum has begun to build in the field of prompt engineering for text generation purposes, it is important to research the impact of prompt engineering skills on creative output with regard to creative quality. Prompt engineering is a valuable form of interaction to study in detail because it can help users develop mental models of generative models. In a paper investigating human–AI interactions in a gameplay setting, Gero et al. (2020) demonstrated that people can construct mental models of AI through repeated interactions with a model that help them understand the AI's knowledge distribution and behavior. Using prompts to generate image evidence of AI knowledge is a further way of reducing uncertainty with AI, which is one of the fundamental challenges in human–AI interaction (Yang et al. 2020). On the Internet, a variety of users have already scrutinized prompt engineering and discussed specific keywords and phrases that can help them understand the knowledge distribution of the generative AI models in order to discover what input can help tune models to their goals. For example, adding text before a task – such as "write this as an expert in domain X" – yields a different output than would otherwise be achieved. These tricks and many others along the same vein have established a growing trend within online communities in getting the best out of prompts. Research has also elicited similar ideas by proposing guidelines on prompt engineering in the domains of text-to-text- (White et al. 2023) and text-to-image-generative AI (Liu and Chilton 2022). Recent results highlight that optimized prompts lead to better solutions (Yang et al. 2023). In this vein, this skill could be crucial when it comes to fully unleashing the potential of generative AI. Hence, we hypothesize:

> H2: Prompt engineering skills should be found to moderate the relationship between domain knowledge and the creative quality of the output. Specifically, among participants with low levels of domain knowledge, those with higher levels of prompt engineering skills should be found to produce creative output that is comparable to that of participants with higher levels of domain knowledge.

In summary, the literature has shown that domain-specific knowledge is an important predictor of creativity. However, in times of generative AI, new opportunities continue to arise. To the best of our knowledge, the literature on creativity has so far only compared humans and AI - but not important factors in the context of collaboration on creative tasks. AI could possess complementary capabilities in mitigating the lack of domain knowledge. Prompt engineering, however, is an important capability for unlocking the full potential of AI. Figure 1 summarizes the proposed research model to test our hypotheses.

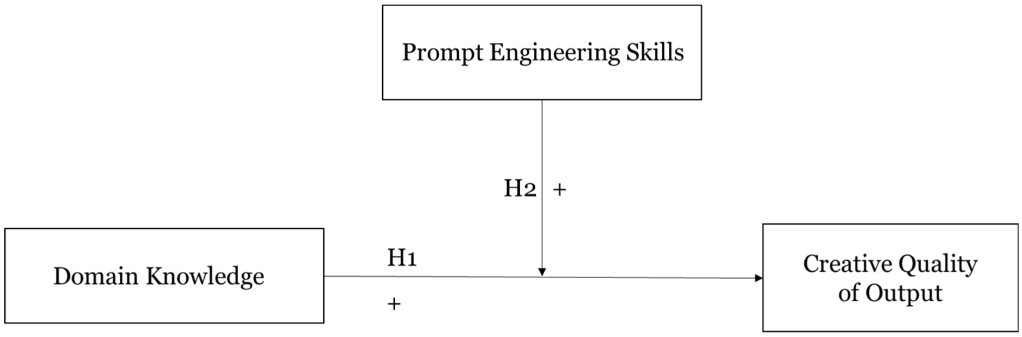

Figure 1. Research Model

3 Methodology

3.1 Participants

We plan to recruit a diverse sample of participants in order to ensure a range of domain knowledge and prompt engineering skills. The sample will be selected from different sources (e.g., university students, online platform participants, etc.). In line with the rule of thumb presented by Green (1991), we will acquire at least 107 participants.

3.2 Measurements

In order to measure independent variable domain knowledge, we will use a validated scale from Wang and Nickerson (2019) that contains seven items (e.g., "Compared with the average person, I do not know much about…"), which we will adapt to suit our context of the tasks. The instrument will use a seven-point Likert scale (ranging from *1 = strongly disagree* to *5 = strongly agree*). We will also adapt this scale to ChatGPT use as alternative self-report scale of the moderator prompt engineering skills.

Furthermore, to measure the moderator prompt engineering skills based on abilities, we will generate a test (i.e., measurement) based on the catalogue proposed by White et al. (White et al. 2023). To do so, we will adapt the approach by MacKenzie and Podsakoff (2011), which will begin with a conceptualization of the construct and end with the validation of the measurement instrument.

In terms of evaluating the dependent variable creativity, we will follow Dean et al. (2006), who advise four dimensions for rating creativity: novelty, feasibility, relevance, and specificity. The instrument will use a seven-point Likert scale (ranging from *1 = not novel / feasible / relevant / specific at all* to *7 = highly novel / feasible / relevant / specific*) to evaluate the creativity of the generated output.

3.3 Procedure

The procedure will be structured in two parts: First, the creative output will be produced, and afterward, the output will be evaluated.

For the first step, each participant will be given the task of reading an online newspaper article from the sports domain (without a title) for 5 minutes. Afterward, the participants will have the task of interacting with ChatGPT (GPT-4) for 10 minutes in order to generate one creative title for the article that yields the maximum number of clicks. The procedure will be repeated with an online newspaper article from the politics domain. Subsequently, the participants will fill out a survey regarding their prompt engineering skills, domain knowledge, and demographics.

In the second step, the headlines will be collated in a random order and presented (anonymously) to a panel of expert judges for creativity assessment.

3.4 Data Collection & Analysis

The goal of this research is to understand what is needed to view perceive generative AI as a co-creative partner. To do so, we conduct a cross-sectional survey study researching the effect of the independent variable domain knowledge and the moderator prompt engineering skills on the dependent variable creative quality. The primary objective of the data analysis will be to discern the influence of domain knowledge and prompt engineering skills on creative output and to ascertain whether prompt engineering skills moderate the effect of domain knowledge on creativity in ChatGPT-generated headlines. Initial analyses will entail computing descriptive statistics – including means and standard deviations –, for domain knowledge, prompt engineering skills, and creative output scores as well as identification of possible confounders to take measures accordingly. Then, an interaction term will be generated for the moderator analysis by multiplying the scores on domain knowledge and prompt engineering skills. Subsequently, a hierarchical multiple regression will be conducted, with creative output as the dependent variable. In the initial step, domain knowledge and prompt engineering skills will serve as predictors. In the following step, the interaction term will be added to the model. This hierarchical approach will enable main effects to be examined prior to considering the interaction effect. We will begin with a pre-test in order ensure that our chosen domain is suitable. Based on this determination, possible further adaptions will be made in order to conduct the study adequately.

4 Outlook

Overall, we aim to address the current stream of technological advancements and the impact of these advancements on creativity. By answering the RQs, we will theoretically contribute to the literature on human–AI collaboration and creativity. We aim to broaden our understanding of how essential skills – such as prompt engineering – can generate qualitative creative output using generative AI. Furthermore, we aim to better understand how prompt engineering skills might mitigate negative creativity-related conditions (e.g., low levels of domain knowledge). For practitioners, the findings will also provide substantial added value regarding the adoption of generative AI for use both as a co-creative partner and in job design, and they should additionally yield a more nuanced view of creativity. Future research could then build on our findings in various ways. First, external validity could be increased by researching further tasks. Second, the pathway to creativity could be scrutinized more closely. In this regard, better understanding whether and how inspiration (Böttger et al. 2017; Weber et al. 2023) is evoked would represent an interesting path forward. Finally, longitudinal studies could be conducted that evaluate human–AI collaboration in the creativity domain over time.

5 References

Baer, J. 2012. "Domain Specificity and the Limits of Creativity Theory," *The Journal of Creative Behavior* (46:1), pp. 16-29 (doi: 10.1002/jocb.002).

Baer, J. 2015. *Domain Specificity of Creativity*, Academic Press.

Bedwell, W. L., Wildman, J. L., DiazGranados, D., Salazar, M., Kramer, W. S., and Salas, E. 2012. "Collaboration at work: An integrative multilevel conceptualization," *Human Resource Management Review* (22:2), pp. 128-145 (doi: 10.1016/j.hrmr.2011.11.007).

Berente, N., Gu, B., Recker, J., and Santhanam, R. 2021. "Managing Artificial Intelligence," *MIS Quarterly* (45:3), pp. 1433-1450.

Böttger, T., Rudolph, T., Evanschitzky, H., and Pfrang, T. 2017. "Customer Inspiration: Conceptualization, Scale Development, and Validation," *Journal of Marketing* (81:6), pp. 116-131 (doi: 10.1509/jm.15.0007).

Bringsjord, S., Bello, P., and Ferrucci, D. 2003. "Creativity, the Turing Test, and the (Better) Lovelace Test," in *The Turing Test*, J. H. Fetzer and J. H. Moor (eds.), Dordrecht: Springer Netherlands, pp. 215-239 (doi: 10.1007/978-94-010-0105-2_12).

Chakrabarty, T., Laban, P., Agarwal, D., Muresan, S., and Wu, C.-S. 2023. "Art or Artifice? Large Language Models and the False Promise of Creativity,"

Cole, D. 2004. "The Chinese Room Argument," available at https://plato.stanford.edu/entries/chinese-room/, accessed on Feb 27 2023.

Daugherty, P. R., and Wilson, H. J. 2018. *Human + Machine*: *Reimagining work in the age of AI*, Boston, Mass.: Harvard Business Review Press.

Dean, D., Hender, J., Rodgers, T., and Santanen, E. 2006. "Identifying Quality, Novel, and Creative Ideas: Constructs and Scales for Idea Evaluation," *Journal of the Association for Information Systems* (7:10), pp. 646-699 (doi: 10.17705/1jais.00106).

Dellermann, D., Ebel, P., Söllner, M., and Leimeister, J. M. 2019. "Hybrid Intelligence," *Business & Information Systems Engineering* (61:5), pp. 637-643 (doi: 10.1007/s12599-019-00595-2).

Feldhusen, J. F. 2010. "The Role of the Knowledge Base in Creative Thinking," in *Creativity and Reason in Cognitive Development*, J. C. Kaufman and J. Baer (eds.), Cambridge University Press, pp. 137-144 (doi: 10.1017/CBO9780511606915.009).

Feuerriegel, S., Hartmann, J., Janiesch, C., and Zschech, P. 2023. "Generative AI," *SSRN Electronic Journal* (doi: 10.2139/ssrn.4443189).

Gero, K. I., Ashktorab, Z., Dugan, C., Pan, Q., Johnson, J., Geyer, W., Ruiz, M., Miller, S., Millen, D. R., Campbell, M., Kumaravel, S., and Zhang, W. 2020. "Mental Models of AI Agents in a Cooperative Game Setting," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, R. Bernhaupt, F. '. Mueller, D. Verweij, J. Andres, J. McGrenere, A. Cockburn, I. Avellino, A. Goguey, P. Bjørn, S. Zhao, B. P. Samson and R. Kocielnik (eds.), Honolulu HI USA. 25 04 2020 30 04 2020, New York, NY, USA: ACM, pp. 1-12 (doi: 10.1145/3313831.3376316).

GitHub Inc. 2022. "Your AI pair programmer," available at https://github.com/features/copilot, accessed on Nov 14 2022.

Grassini, S. 2023. "Shaping the Future of Education: Exploring the Potential and Consequences of AI and ChatGPT in Educational Settings," *Education Sciences* (13:7), p. 692 (doi: 10.3390/educsci13070692).

Green, S. B. 1991. "How Many Subjects Does It Take To Do A Regression Analysis," *Multivariate behavioral research* (26:3), pp. 499-510 (doi: 10.1207/s15327906mbr2603_7).

Guzik, E. E., Byrge, C., and Gilde, C. 2023. "The originality of machines: AI takes the Torrance Test," *Journal of Creativity* (33:3), p. 100065 (doi: 10.1016/j.yjoc.2023.100065).

Gwern. 2020. "GPT-3: Creative Fiction," available at https://www.gwern.net/GPT-3, accessed on Feb 27 2023.

Hu, K. 2023. "ChatGPT sets record for fastest-growing user base - analyst note," available at https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/, accessed on Feb 28 2023.

Huang, P.-S., Peng, S.-L., Chen, H.-C., Tseng, L.-C., and Hsu, L.-C. 2017. "The relative influences of domain knowledge and domain-general divergent thinking on scientific creativity and mathematical creativity," *Thinking Skills and Creativity* (25), pp. 1-9 (doi: 10.1016/j.tsc.2017.06.001).

Hulman, A., Dollerup, O. L., Mortensen, J. F., Fenech, M., Norman, K., Støvring, H., and Hansen, T. K. 2023. *ChatGPT- versus human-generated answers to frequently asked questions about diabetes: a Turing test-inspired survey among employees of a Danish diabetes center*.

Koivisto, M., and Grassini, S. 2023. "Best humans still outperform artificial intelligence in a creative divergent thinking task," *Scientific reports* (13:1), p. 13601 (doi: 10.1038/s41598-023-40858-3).

Lai, Y., Kankanhalli, A., and Ong, D. 2021. "Human-AI Collaboration in Healthcare: A Review and Research Agenda," in *Proceedings of the 54th Hawaii International Conference on System Sciences*, T. Bui (ed.), Hawaii International Conference on System Sciences (doi: 10.24251/HICSS.2021.046).

Liu, V., and Chilton, L. B. 2022. "Design Guidelines for Prompt Engineering Text-to-Image Generative Models," in *CHI Conference on Human Factors in Computing Systems*, S. Barbosa, C. Lampe, C. Appert, D. A. Shamma, S. Drucker, J. Williamson and K. Yatani (eds.), New Orleans LA USA. 29 04 2022 05 05 2022, New York, NY, USA: ACM, pp. 1-23 (doi: 10.1145/3491102.3501825).

MacCrimmon, K. R., and Wagner, C. 1994. "Stimulating Ideas Through Creative Software," *Management Science* (40:11), pp. 1514-1532 (doi: 10.1287/mnsc.40.11.1514).

MacKenzie, and Podsakoff. 2011. "Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques," *MIS Quarterly* (35:2), p. 293 (doi: 10.2307/23044045).

Mayer, R. E. 2010. "The Role of Domain Knowledge in Creative Problem Solving," in *Creativity and Reason in Cognitive Development*, J. C. Kaufman and J. Baer (eds.), Cambridge University Press, pp. 145-158 (doi: 10.1017/CBO9780511606915.010).

OpenAI. 2022. "Examples," available at https://beta.openai.com/examples/, accessed on Nov 17 2022.

Plucker, J. A., Beghetto, R. A., and Dow, G. T. 2004. “Why Isn't Creativity More Important to Educational Psychologists? Potentials, Pitfalls, and Future Directions in Creativity Research,” *Educational Psychologist* (39:2), pp. 83-96 (doi: 10.1207/s15326985ep3902_1).

Poole, B., Jain, A., Barron, J. T., and Mildenhall, B. 2022. “DreamFusion: Text-to-3D using 2D Diffusion,”

Powell, J. 2019. “Trust Me, I'm a Chatbot: How Artificial Intelligence in Health Care Fails the Turing Test,” *Journal of Medical Internet Research* (21:10), e16222 (doi: 10.2196/16222).

Riedl, M. O. 2014. “The Lovelace 2.0 Test of Artificial Creativity and Intelligence,”

Roose, K. 2022. “The Shift: An A.I.-Generated Picture Won an Art Prize. Artists Aren’t Happy.,” available at https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html, accessed on Feb 27 2023.

Runco, M. A. 2023a. “AI can only produce artificial creativity,” *Journal of Creativity* (33:3), p. 100063 (doi: 10.1016/j.yjoc.2023.100063).

Runco, M. A. 2023b. “Updating the Standard Definition of Creativity to Account for the Artificial Creativity of AI,” *Creativity Research Journal*, pp. 1-5 (doi: 10.1080/10400419.2023.2257977).

Runco, M. A., and Jaeger, G. J. 2012. “The Standard Definition of Creativity,” *Creativity Research Journal* (24:1), pp. 92-96 (doi: 10.1080/10400419.2012.650092).

Schuetz, S., and Venkatesh, V. 2020. “"Research Perspectives: The Rise of Human Machines: How Cognitive Computing Systems Challenge Assumptions of User-System Interaction ",” *Journal of the Association for Information Systems*, pp. 460-482 (doi: 10.17705/1jais.00608).

Schuetzler, R. M., Grimes, G. M., Giboney, J. S., and Rosser, H. K. 2021. “Deciding Whether and How to Deploy Chatbots,” *MIS Quarterly Executive*, pp. 1-15 (doi: 10.17705/2msqe.00039).

Silvia, P. J., Kaufman, J. C., and Pretz, J. E. 2009. “Is creativity domain-specific? Latent class models of creative accomplishments and creative self-descriptions,” *Psychology of Aesthetics, Creativity, and the Arts* (3:3), pp. 139-148 (doi: 10.1037/a0014940).

Turing, A. M. 1950. “I.—Computing Machinery and Intelligence,” *Mind* (LIX:236), pp. 433-460 (doi: 10.1093/mind/LIX.236.433).

Wang, D., Churchill, E., Maes, P., Fan, X., Shneiderman, B., Shi, Y., and Wang, Q. 2020. “From Human-Human Collaboration to Human-AI Collaboration,” in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, R. Bernhaupt, F. '. Mueller, D. Verweij, J. Andres, J. McGrenere, A. Cockburn, I. Avellino, A. Goguey, P. Bjørn, S. Zhao, B. P. Samson and R. Kocielnik (eds.), Honolulu HI USA. 25 04 2020 30 04 2020, New York, NY, USA: ACM, pp. 1-6 (doi: 10.1145/3334480.3381069).

Wang, K., and Nickerson, J. V. 2017. “A literature review on individual creativity support systems,” *Computers in Human Behavior* (74), pp. 139-151 (doi: 10.1016/j.chb.2017.04.035).

Wang, K., and Nickerson, J. V. 2019. “A Wikipedia-based Method to Support Creative Idea Generation: The Role of Stimulus Relatedness,” *Journal of Management Information Systems* (36:4), pp. 1284-1312 (doi: 10.1080/07421222.2019.1661095).

Weber, S., Klassen, G., Wyszynski, M., and Kordyaka, B. 2023. “Illuminating the Predictive Power of Gamification to Inspire Technology Users,” in *Mensch und Computer 2023,* Rapperswil, Switzerland, ACM (doi: 10.18420/muc2023-mci-ws08-378).

White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., and Schmidt, D. C. 2023. *A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT.*

Yang, C., Wang, X., Lu, Y., Liu, H., Le V, Q., Zhou, D., and Chen, X. 2023. “Large Language Models as Optimizers,”

Yang, Q., Steinfeld, A., Rosé, C., and Zimmerman, J. 2020. “Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design,” in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, R. Bernhaupt, F. '. Mueller, D. Verweij, J. Andres, J. McGrenere, A. Cockburn, I. Avellino, A. Goguey, P. Bjørn, S. Zhao, B. P. Samson and R. Kocielnik (eds.), Honolulu HI USA. 25 04 2020 30 04 2020, New York, NY, USA: ACM, pp. 1-13 (doi: 10.1145/3313831.3376301).

Copyright