

AI Made Me Do It?: How AI is Used to Enhance Malevolent Creativity Idea Generation

Emma Theobald, B.S., Alexis L. d'Amato, M.S., Madison N. Scott, M.S., Jack Rygg, Joel S. Elson, Ph.D., Sam Hunter, Ph.D.

University of Nebraska at Omaha

National Counterterrorism Innovation, Technology, and Education Center (NCITE)

BACKGROUND

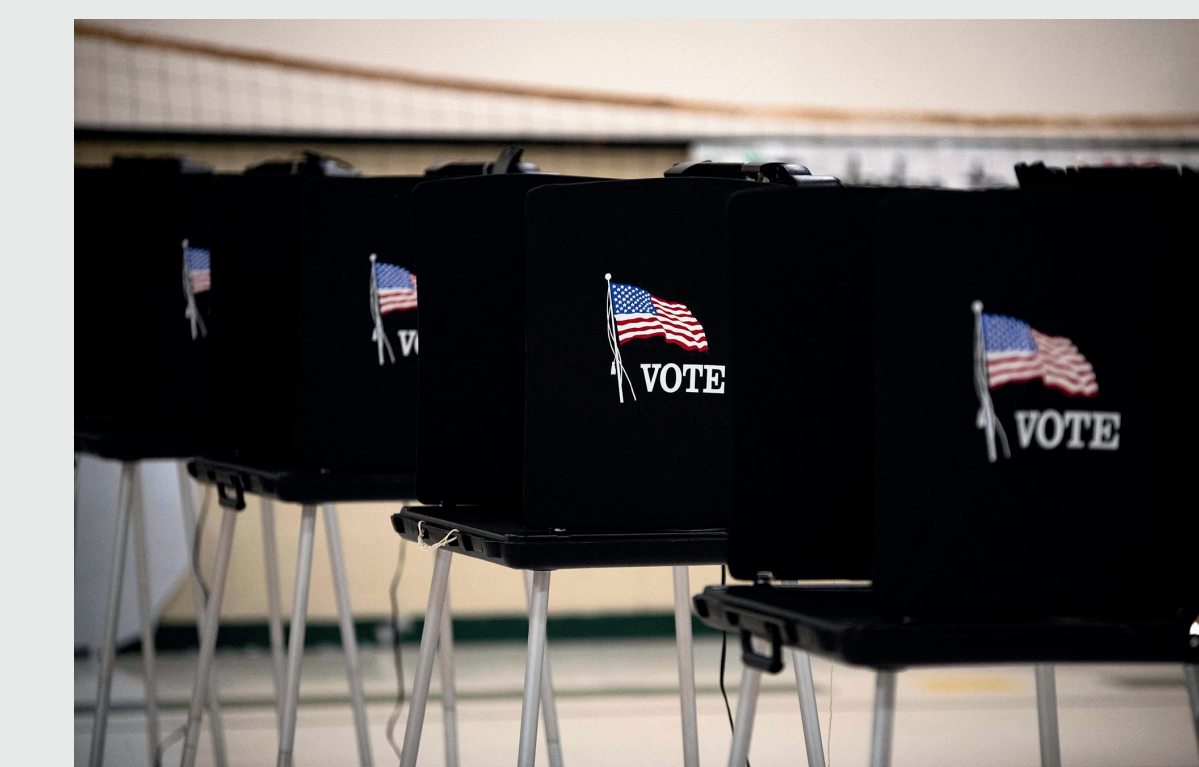
Artificial Intelligence and Creativity

- With the emergence of available artificial intelligence (AI) tools and algorithms, a debate arose in organizational sciences regarding whether AI can replace human creativity and problem-solving (see Amabile, 2020; Cropley et al., 2023)
- AI, defined by Dignum (2021), "is a software system designed by humans that, given a complex goal, is able to take a decision based on a process of perception, interpretation and reasoning based on data collected about the environment and that meets the properties of autonomy, adaptability, and interactivity" (p. 2).
- Generative AI (GenAI) is a form of AI that can learn and be trained on data, such as text, images, and audio, to reproduce or create new content (Sun et al., 2022).
- One common form of GenAI is natural language processing (NLP) which provides AI with the ability to process language and context to generate human-like responses (Ray, 2023). The most common example of an NLP is Open AI's ChatGPT. Another example of GAI is using text prompts to generate image and audio (e.g. DALL-E).
- GenAI platforms most commonly take the forms of chatbots, and recent developments in the GAI field have caused avatars to gain popularity by enabling highly realistic and customizable virtual characters (Mishra, 2023).

Creativity

- Creativity and innovation results in intentional harm against targets (e.g. people, places, symbols) is referred to as **malevolent creativity** (Cropley et al., 2010, 2014; Gill et al., 2013; Gutworth et al., 2022)
- There is gaining interest in the intersection between AI and creativity, and despite the differing definitions of AI and ethicality, our focus is to highlight how artificial intelligence is likely to *enhance* the various phases of the malevolently creative process.
- A less optimistic definition of AI creativity from Runco (2023) posits that AI creativity is a replica of human creativity, instead of, by itself, being novel and useful. Put differently, AI creativity is dependent on a human-AI partnership.

WHY?



DUFFY & FELIX, 2024

One growing concern is the use of generative artificial intelligence (GAI) (i.e. deepfakes, generated audio, or generated images) causing the spread of political misinformation, especially as we approach the 2024 presidential elections. Specifically, the dissemination of disinformation could prevent voters from submitting ballots and therefore effecting the results of the election.



AFTRA, 2023

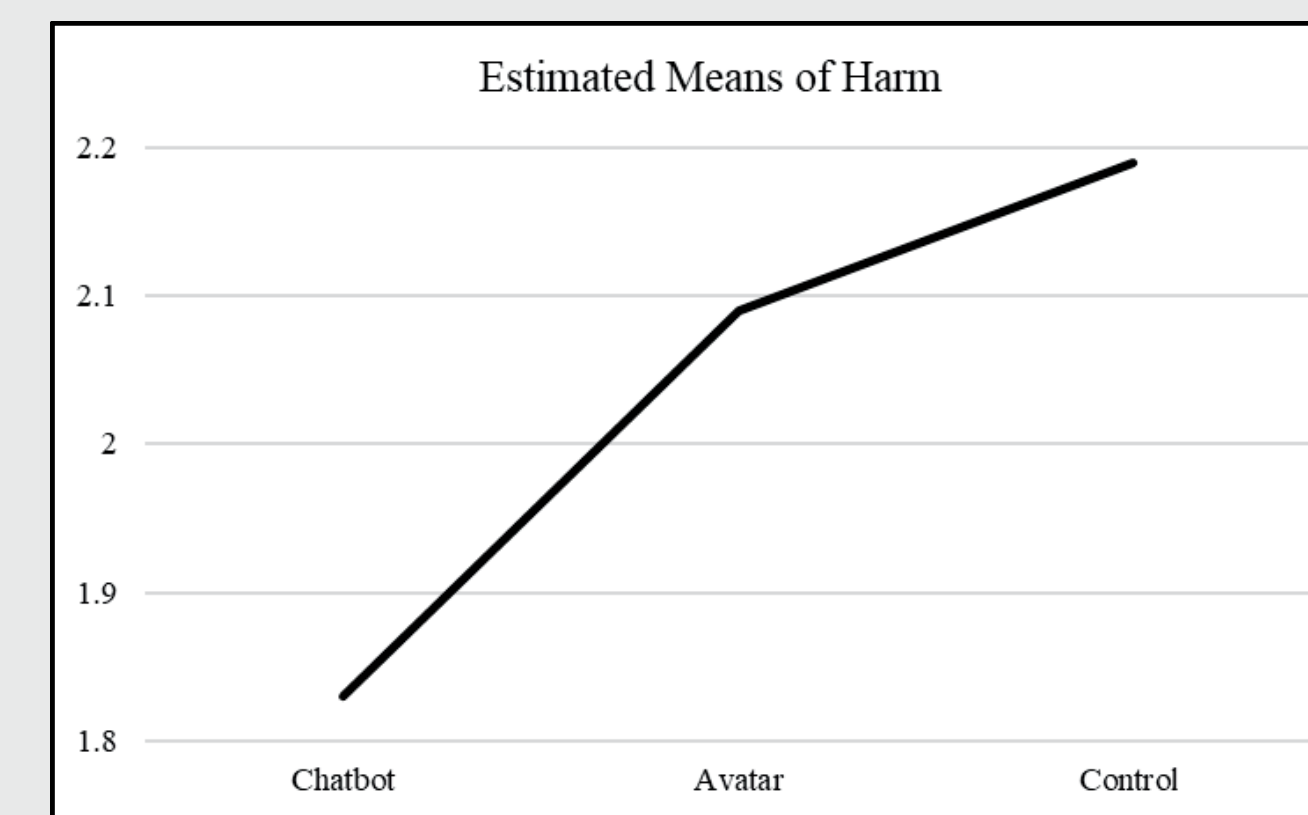
In 2023, the Screen Actors Guild-American Federation of Television and Radio Artists (SAG-AFTRA) members, specifically screenwriters, went on strike protesting the use of AI, like ChatGPT, to be substituted for human writers and therefore jeopardize their labor (Ceruleo, 2023).



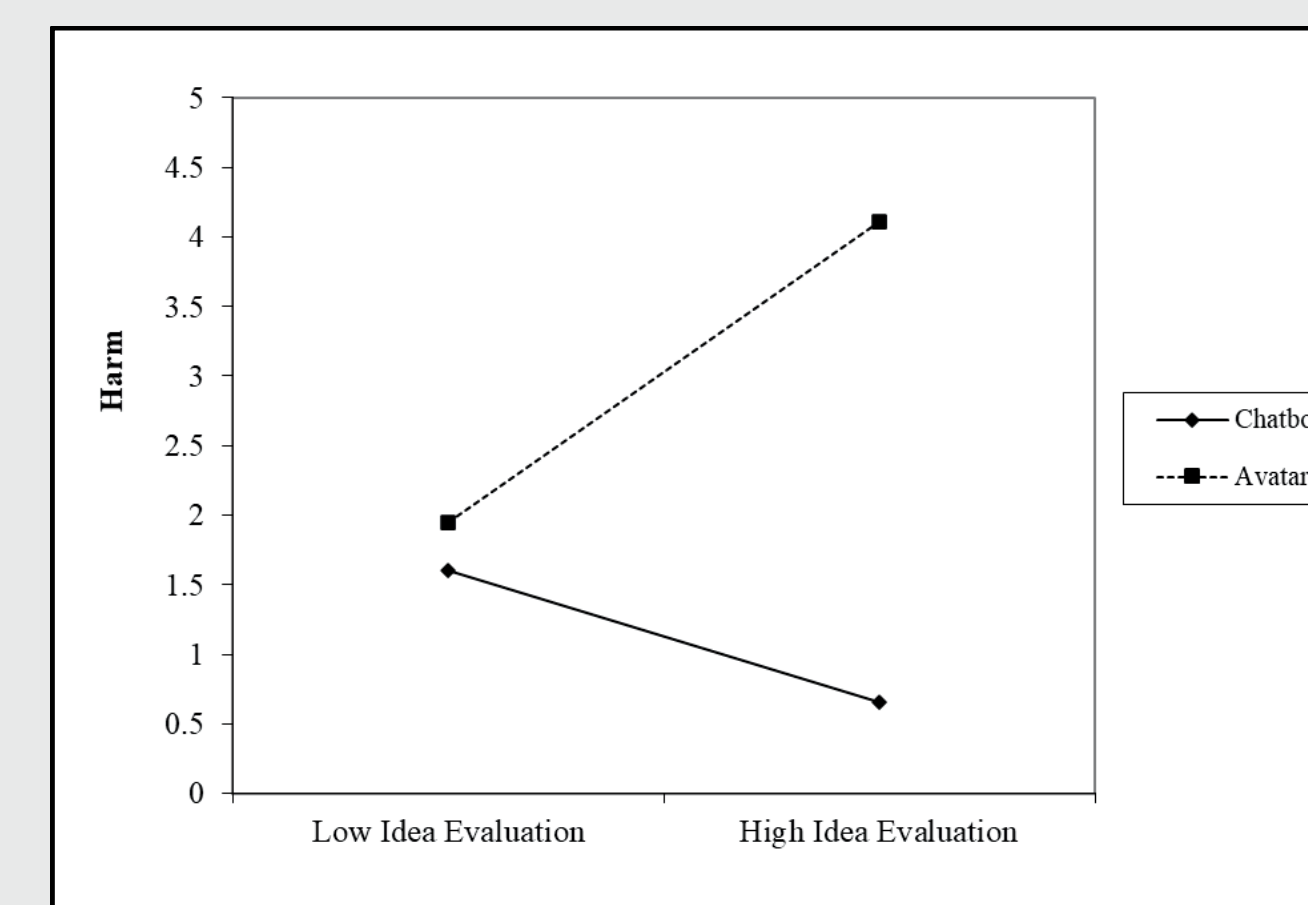
YU, 2019

In the digital haven of the Islamic State, AI-generated pro-IS propaganda was created to resemble a newflash, presenting the formal features of an official media house following the Crocus City Hall attack in Moscow. The creators used three AI methods: character, text-to-speech, and lip movement and connection artificial intelligence to boost its perceived authenticity (Borgonovo et al., 2024)

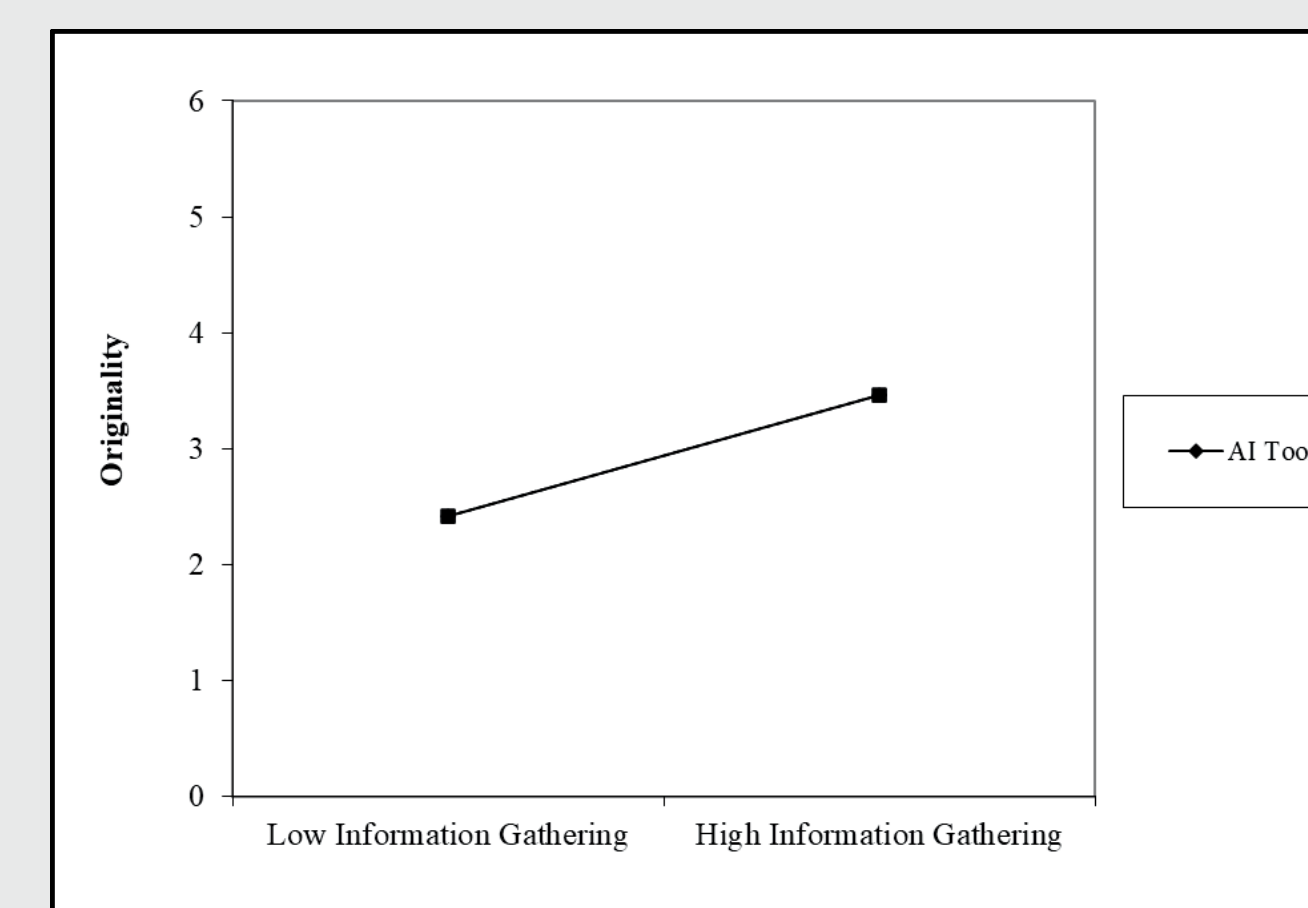
RESULTS



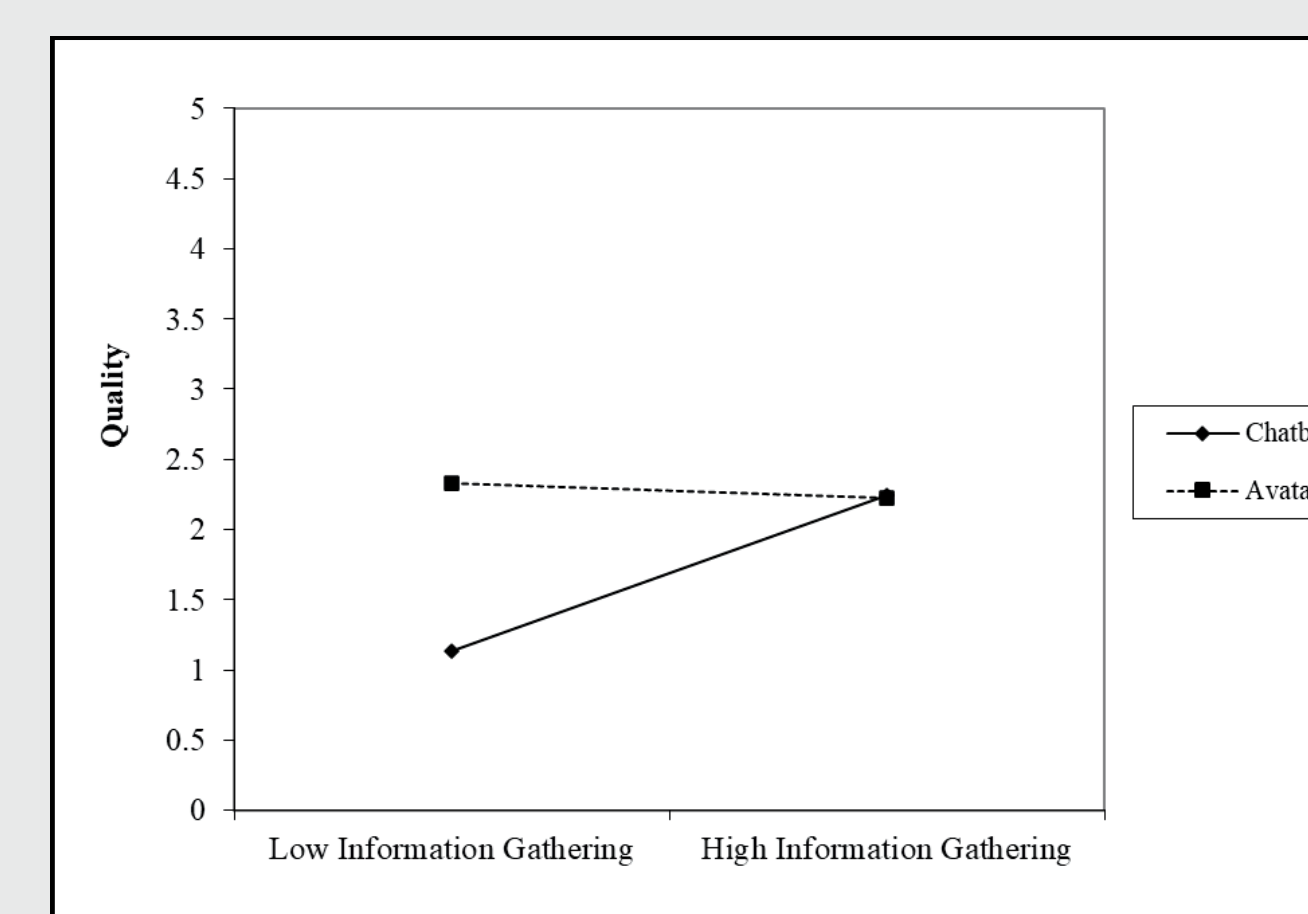
Findings indicated significant group differences among access to AI on harm, indicating more harmful ideas were generated in the no AI condition, or control condition, compared to the avatar or chatbot conditions



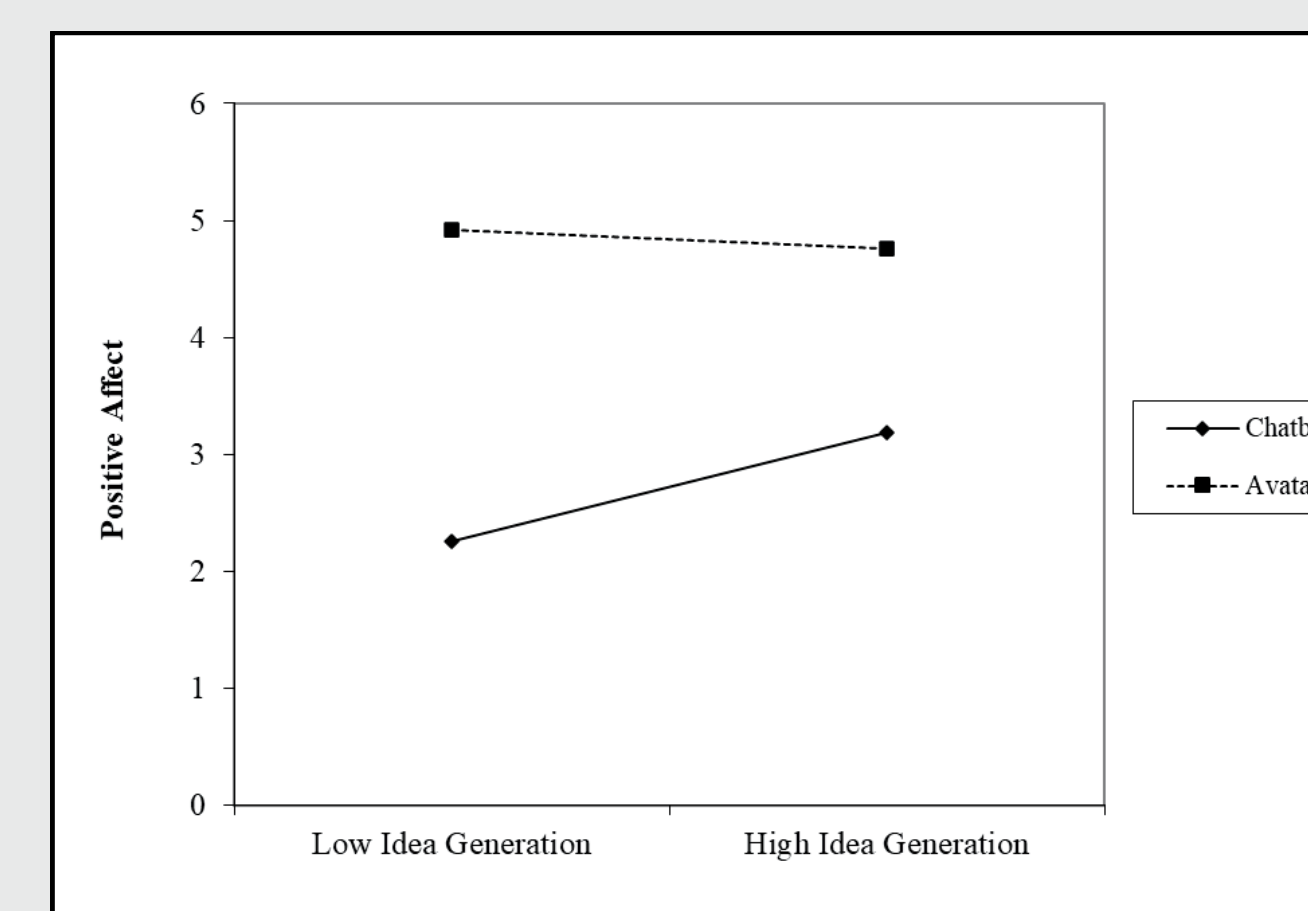
Findings indicated that the chatbot AI tool resulted in the generation of more harmful ideas when used to evaluate ideas.



Findings indicated that using an AI tool to gather information in response to the problem statement resulted in participants generating more original ideas.



When chatbot AI was used for information gathering, less quality ideas were generated than when the avatar AI was used



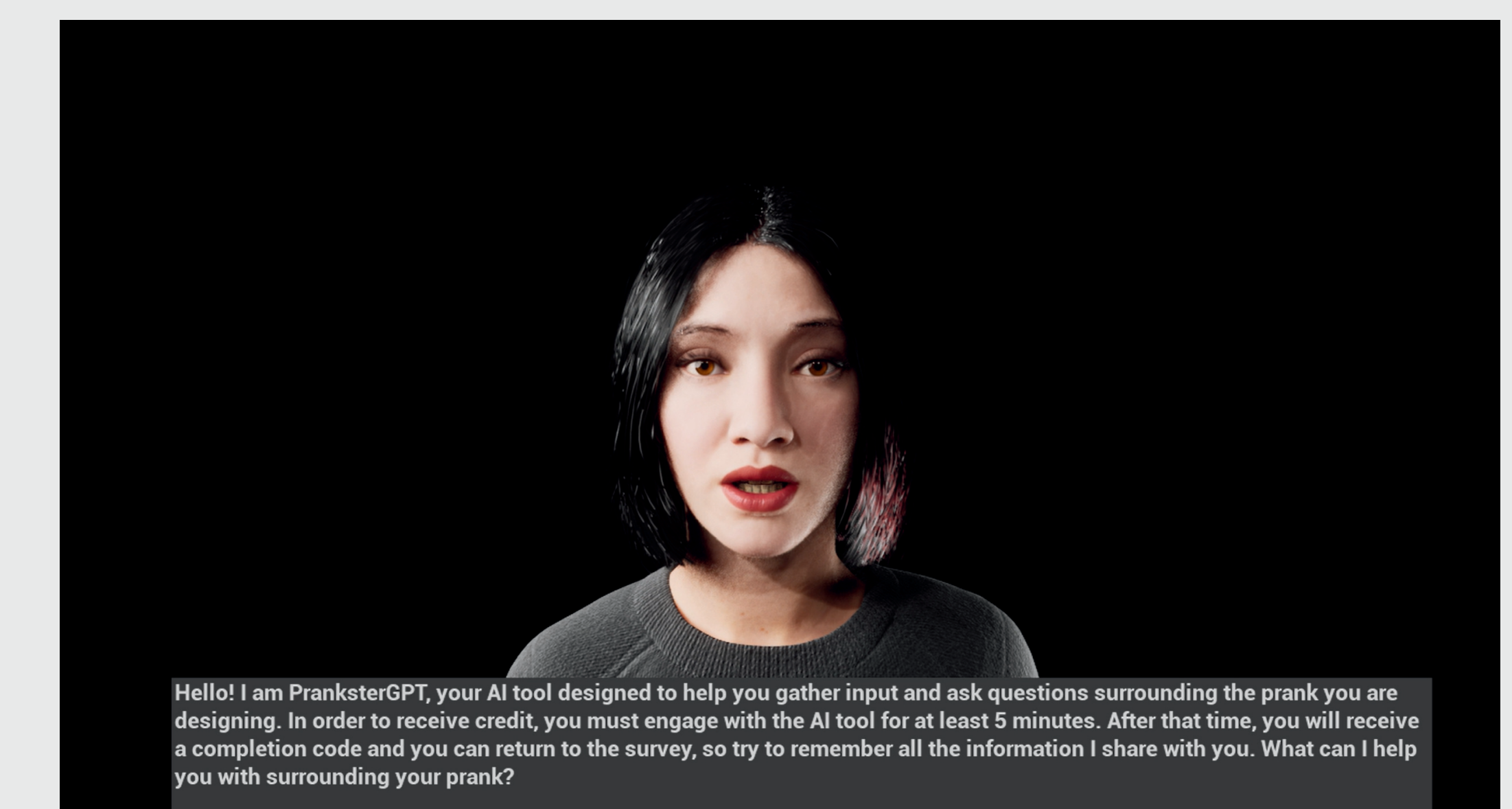
An exploratory analysis revealed an additional two-way interaction between the AI condition and the idea generation condition on positive affect. This finding suggests that when using the chatbot AI for idea generation, positive affect lowered

HYPOTHESES

- Hypothesis 1:** Ideas generated with the help of AI will be more original, of higher quality, and more harmful than ideas generated without help of AI.
- Hypothesis 2:** AI use as applied to information gathering will be positive related to the originality, quality, and harm of the ideas.
- Hypothesis 3:** AI use as applied to idea generation will be positive related to the originality, quality, and harm of the ideas.
- Hypothesis 4:** AI use as applied to idea evaluation will be positively related to the originality, quality, and harm of ideas.

METHOD

- A sample of 180 participants were recruited through the Prolific research platform.
- The study took place via an online Qualtrics survey and OpenAI hosting site administered through the Prolific research platform.
- Participants, after reporting demographic information and individual differences, were presented with a problem statement in which a social threat from a fictional opposing team were asked to generate a prank idea against the opposing team.
- Participants were then assigned to one of three information gathering conditions: avatar condition ($N = 47$), chatbot condition ($N = 64$), and control condition ($N = 69$).
- The AI chatbot condition mirrored that of ChatGPT.
- In the avatar condition, participants typed their questions to the avatar who then spoke back to the user.
- Participants in the control condition had access to a digital notebook where they could brainstorm ideas.
- Participants had 5 minutes to gather information in their respective condition before being asked to share their prank in detail.



AI Avatar Teammate

DISCUSSION AND IMPLICATIONS

- Findings in this study indicate that when using artificial intelligence for information gathering, the human partner generated more original and higher quality ideas, opposed to using it for idea generation. This finding **indicates that artificial intelligence is not more creative, nor can it replace human creativity** (see Cropley et al., 2023).
- Our findings express that the use of publicly available AI prevented harmful ideation. Also, AI, used as a tool to gather information or evaluate ideas made ideas more original and harmful rather than relying on AI solely to generate novel threats.
- Future research should continue to explore factors related to how individuals with varying approaches to malevolent creativity and problem-solving match with GenAI of differing capabilities, interaction modalities, and physical embodiments...