Deepfake videos prompt false memories of films in half of participants

July 13 2023

The authors suggest we don't need technical advances to distort memory, we can do it very easily and effectively using non-technical means. Credit: Anemone123, Pixabay, CC0 (creativecommons.org/publicdomain/zero/1.0/)

In a new study, deepfake video clips of movie remakes that don't actually exist prompted participants to falsely remember the films—but simple text descriptions of the fake movies prompted similar false

memory rates. Gillian Murphy of University College Cork, Ireland, and Lero, the Science Foundation Ireland Research Centre for Software, and colleagues presented these findings in the open-access journal *PLOS ONE.*

Deepfake videos are clips that have been altered using artificial intelligence (AI) technology to switch out one person's voice or face with that of a different person. Tools for making deepfakes have recently become much cheaper and more accessible, amplifying conversations about potential creative applications as well as potential risks—such as spreading misinformation and manipulating viewers' memories.

To explore potential risks and benefits, Murphy and colleagues invited 436 people to complete an online survey that included watching deepfake videos of fake movie remakes starring different actors, such as Will Smith as the character Neo—originally played by Keanu Reeves—in The Matrix and Brad Pitt and Angelina Jolie in The Shining. Other fake remakes in the study were Indiana Jones and Captain Marvel.

For comparison, participants also saw clips of real remakes, including Charlie & The Chocolate Factory, Total Recall, Carrie, and Tomb Raider. As well, in some cases, participants read text descriptions of the remakes instead of watching the deepfakes. Participants were not told that the deepfakes were false until later in the survey.

In line with prior studies, the deepfake videos and descriptions prompted false memories of the fake remakes, with an average of 49 percent of participants believing each fake remake was real. Many participants reported they remembered the fake remakes being better than the originals. However, false memory rates from the text descriptions were similarly high, suggesting that deepfake technology may not be more powerful than other tools at distorting memory.

Most participants reported being uncomfortable with deepfake technology being used to recast films, citing concerns including disrespecting artistic integrity and disrupting shared social experiences of movies.

These findings could help inform future design and regulation of deepfake technology in films.

The authors add, "While deepfakes are of great concern for many reasons, such as non-consensual pornography and bullying, the current study suggests they are not uniquely powerful at distorting our memories of the past. Though deepfakes caused people to form false memories at quite high rates in this study, we achieved the same effects using simple text. In essence, this study shows we don't need technical advances to distort memory, we can do it very easily and effectively using non-technical means."

  More information: Gillian Murphy et al, Face/Off: Changing the face of movies with deepfakes, *PLOS ONE* (2023). DOI: 10.1371/journal.pone.0287503

Provided by Public Library of Science