## Hey! What's New? 2025-54

## Are We Any Good at Spotting AI Fakes?

With AI-generated images on the rise, it's becoming increasingly important for people to be able to spot them. But just how good are we at doing so? According to an article in *KelloggInsight*, research led by Kellogg's Matthew Groh, a deepfake expert, found that people were able to distinguish between real and fake images in about five out of every six images they saw.

But not all AI-generated photos are obviously fake. Sometimes careful observation can reveal a clue that something looks off, but other times, AI-generated photos are indistinguishable from real photographs. So, how often and in what contexts can people tell the difference between real and fake photos?

Gro and his team, which conducted a large experiment involving more than 50,000 participants, found that people's accuracy varied widely depending on the complexity of a photo, the kinds of distortions it contained and the amount of time participants spent looking at it.

For the study, the team created a dataset of 149 real photographs curated from the internet and 450 images generated using AI text-to-image tools including Midjourney, Firefly, and Stable Diffusion. Both the real and AI-generated images depicted similar scenarios, and the AI-generated images were selected from a larger dataset of 3,000 AI-generated images based on images that looked most photorealistic to the team.

The team then set up an online experiment, where participants viewed a random arrangement of these images and indicated whether they thought each was real or fake. After they viewed five images, participants were then randomized into one of five groups that viewed each of the remaining images for a certain amount of time: one second, five seconds, ten seconds, twenty seconds, and unlimited time.

Overall, participants correctly identified AI-generated images 76% of the time and real photographs 74% of the time. "That's very much in line with what other experiments have found," says Groh. "It's halfway between random guessing and perfect identification."

Accuracy increased with more-complex images, like photos of groups – because there was a bigger chance of the AI platform getting something wrong, like weird-looking hair or inconsistent lighting. "Specifically, AI-generated simple portraits were harder to detect than group photos," the research found.

Perhaps unsurprisingly, longer viewing times also increased accuracy. But the magnitude of the increase was impressive: with one second of display time, participants were accurate 72% of the time for AI images; at five seconds, accuracy increased to 77%, and at ten seconds, to 80%.

"That's a big jump in accuracy," Groh says. "If you just take a few more seconds to look at an image, you will be much better at determining if it is AI-generated. It's a simple intervention that anyone can use."

Groh and his colleagues then created a taxonomy of common issues associated with Algenerated images, from functional implausibilities, such as a woman holding a sandwich sideways, to stylistic artifacts like waxy skin.

In general, participants were the least accurate at categorizing AI images that had functional implausibilities. In contrast, anatomical errors like unrealistic body proportions were easiest to spot

The team ran a second experiment with a second batch of images generated from the same prompts as the first dataset but without any human curation. And, this time, participants were much more accurate at categorizing the images.

The team plans to use this data and their new taxonomy to create interventions that can help people get better at distinguishing real from fake photos. "We want to guide people where to look, especially if they only have a few seconds," Groh says. "That way, you can engage their attention to help them make more-informed decisions so they don't fall for deepfake images."

For a whole lot more, check out When Put to the Test, Are We Any Good at Spotting AI Fakes?