

Study explores why human-inspired machines can be perceived as eerie

April 25 2024, by Ingrid Fadelli



The uncanny valley: An artwork representing the recent study. Credit: Karl F. MacDorman.

Artificial intelligence (AI) algorithms and robots are becoming increasingly advanced, exhibiting capabilities that vaguely resemble those of humans. The growing similarities between AIs and humans could ultimately bring users to attribute human feelings, experiences, thoughts, and sensations to these systems, which some people perceive as

erie and uncanny.

Karl F. MacDorman, associate dean of academic affairs and associate professor at the Luddy School of Informatics, Computing, and Engineering in Indiana, has been conducting extensive research aimed at better understanding what could make some robots and AI systems feel unnerving.

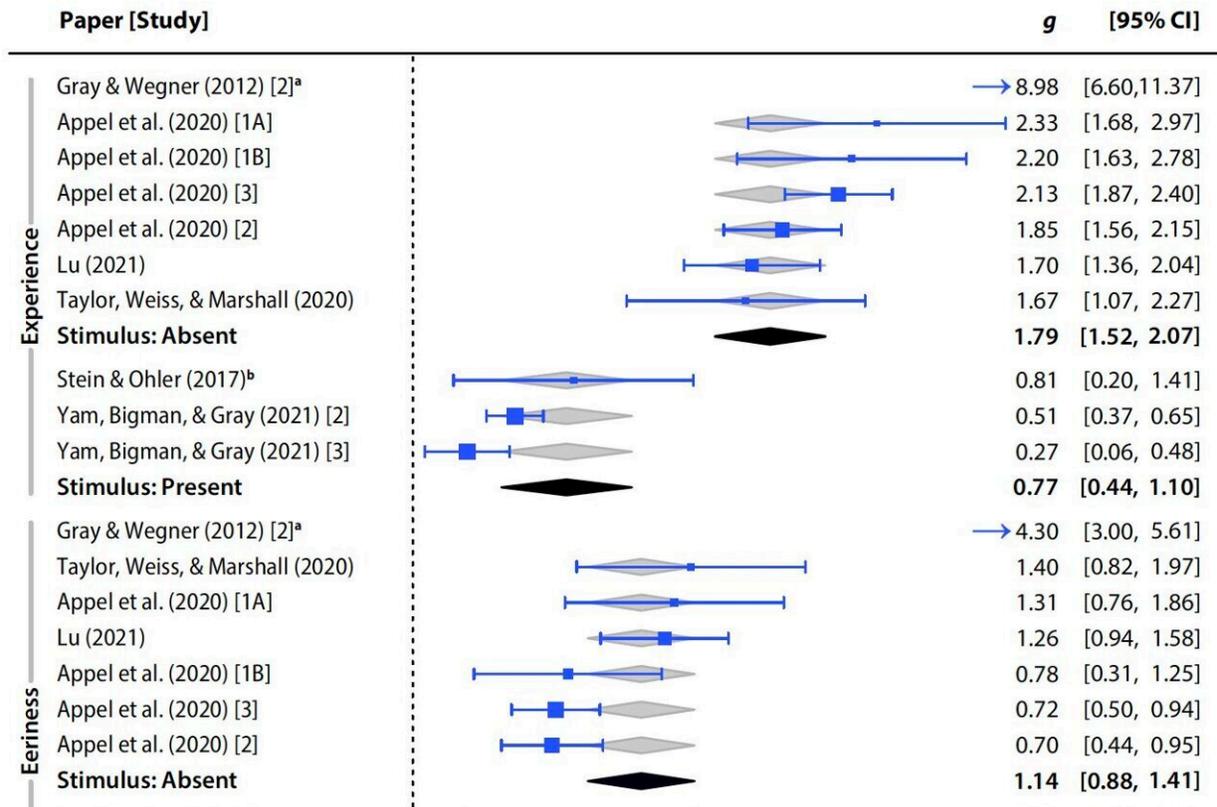
His most recent [paper](#), published in *Computers in Human Behavior: Artificial Humans*, reviews past studies and reports the findings of an experiment testing a recent theory known as "mind perception," which proposes that people feel eeriness when exposed to robots that closely resemble humans because they ascribe minds to these robots.

"For many Westerners, the idea of a machine with conscious experience is unsettling," MacDorman told Tech Xplore. "This discomfort extends to inanimate objects as well. When reviewing Gray and Wegner's [paper](#) published in *Cognition* in 2012, I considered this topic well worth investigating, as its urgency has increased with the rapid rise of AIs, such as ChatGPT and other large language models (LLMs)."

LLMs are sophisticated natural language processing (NLP) models that can learn to answer user queries in ways that strikingly mirror how humans communicate. While their responses are based on vast amounts of data that they were previously trained on, the language they use and their ability to generate query-specific content could make them easier to mistake for sentient beings.

In 2012, two researchers at Harvard University and University of North Carolina carried out experiments that explored the so-called "uncanny valley." This term is used to describe the unnerving nature of robots that have human-like qualities.

When he first read this 2012 paper, MacDorman was skeptical about whether "mind perception" was at the root of the uncanny valley. This inspired him to conduct new studies exploring its existence and the extent to which robots with human-like characteristics are considered eerie.



The mixed-effects meta-regression model with construct (experience or eeriness) and stimulus (present or absent) as moderators. Credit: Karl F. MacDorman.

"Despite other researchers replicating Gray and Wegner's findings, their reliance on the same experimental method and flawed assumptions merely circles back to the initial hypotheses," MacDorman explained.

"Essentially, they assume the conclusions they seek to establish. The uncanny valley is the relation between how humanlike an artificial entity appears and our feelings of affinity and eeriness for it. So, unless the cause of these feelings is our perception of the entity through our five senses, the experiment is not about the uncanny valley."

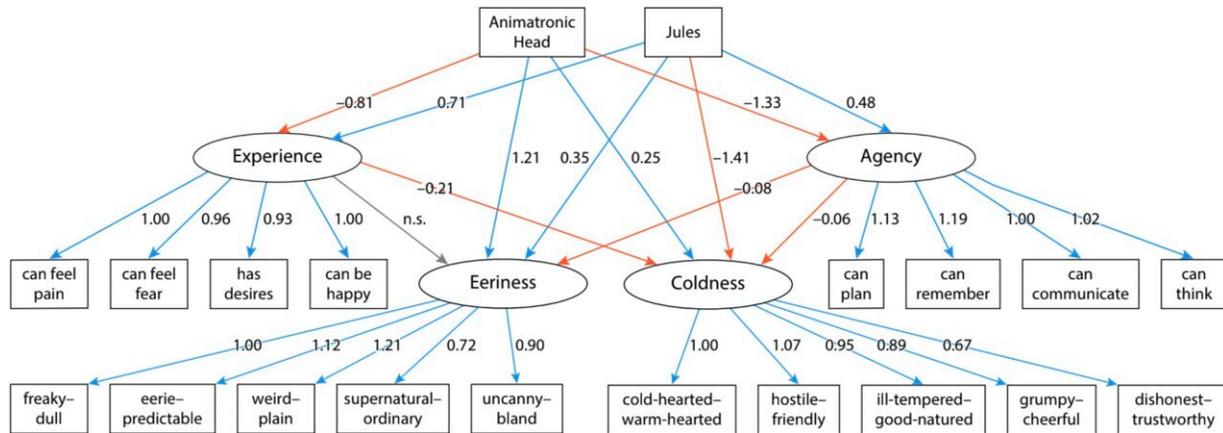
The first objective of MacDorman's recent paper was to pinpoint the possible shortcomings of past "mind perception" experiments. Specifically, his hypothesis was that these experiments disconnect the manipulation of experimental conditions from the appearance of an AI system.

"The manipulation in 'mind perception' experiments is just a description of whether the entity can sense and feel," MacDorman said. "What better way to show this disconnection than by re-analyzing previous experiments? When I did that, I found that machines described as able to sense and feel were much less eerie when they were physically present or represented in videos or virtual reality than when they were absent."

In addition to performing a meta-regression analysis of previous related findings, MacDorman also designed a new experiment that could be used to test mind perception theory. This experiment entails convincing a group of participants who do not think that machines could be sentient that they in fact can be or vice versa, by asking them to read texts and write related essays.

"This experiment allows us to compare how eerie the same robots are when they are viewed by a group of people who ascribe more sentience to them and a group who ascribe less sentience to them," MacDorman said. "If Gray and Wegner (2012) were correct, the group that ascribes more sentience should also find the robots eerier, yet the results show otherwise."

Overall, the results of the meta-analysis and experiment run by MacDorman suggest that past studies backing mind perception theory could be flawed. In fact, the researcher gathered opposite results, suggesting that individuals who attribute sentience to robots do not necessarily find them eerier due to their human resemblance.



A structural equation model showing the stimuli's effect of experience, agency, eeriness, and coldness with Aiko as the reference. Credit: Karl F. MacDorman.

"The uncanny valley is a research area littered with theories purporting to explain the phenomenon," MacDorman said. "There are dozens of such theories, and I must admit, I am one of the main culprits. What I often see is researchers advancing a particular theory, and one of the problems with research in general is that positive results are more likely to be published than negative results.

"What I seldom observed, with [a few exceptions](#), is an attempt to falsify theories or hypotheses or at least to compare their explanatory power."

MacDorman's recent work is among the first to critically examine mind

perception and experiments aimed at testing this theory. His findings suggest that rather than being linked to mind perception, the uncanny valley is rooted in automatic and stimulus-driven perceptual processes.

"The paper shows that the main cause of the uncanny valley is not attributions of conscious experience to machines," MacDorman said. "It also shows that mind perception theory reaches beyond humanoid robots and can be applied to disembodied AI like ChatGPT. This is a positive result from the meta-regression analysis on the 10 experiments found in the literature."

This recent study gathered interesting new insight about the so-called uncanny valley and the link between perceptions of mind and how eerie a [robot](#) is perceived to be. This insight contributes to the understanding of how humans perceive robots and could inform the development of future AI systems.

"Although attributions of mind are not the main cause of the uncanny valley, they are part of the story," MacDorman added. "They can be relevant in some contexts and situations, yet I don't think that attributing mind to a machine that looks human is creepy. Instead, perceiving a mind in a machine that already looks creepy makes it creepier. However, perceiving a mind in a machine that has risen out of the [uncanny valley](#) and looks nearly human makes it less creepy.

"Exploring whether there is strong support for this speculation is an area for future research, which would involve using more varied and numerous stimuli."

More information: Karl F. MacDorman, Does mind perception explain the uncanny valley? A meta-regression analysis and

(de)humanization experiment, *Computers in Human Behavior: Artificial Humans* (2024). [DOI: 10.1016/j.chbah.2024.100065](https://doi.org/10.1016/j.chbah.2024.100065)

© 2024 Science X Network

Citation: Study explores why human-inspired machines can be perceived as eerie (2024, April 25) retrieved 5 May 2024 from <https://techxplore.com/news/2024-04-explores-human-machines-eerie.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.