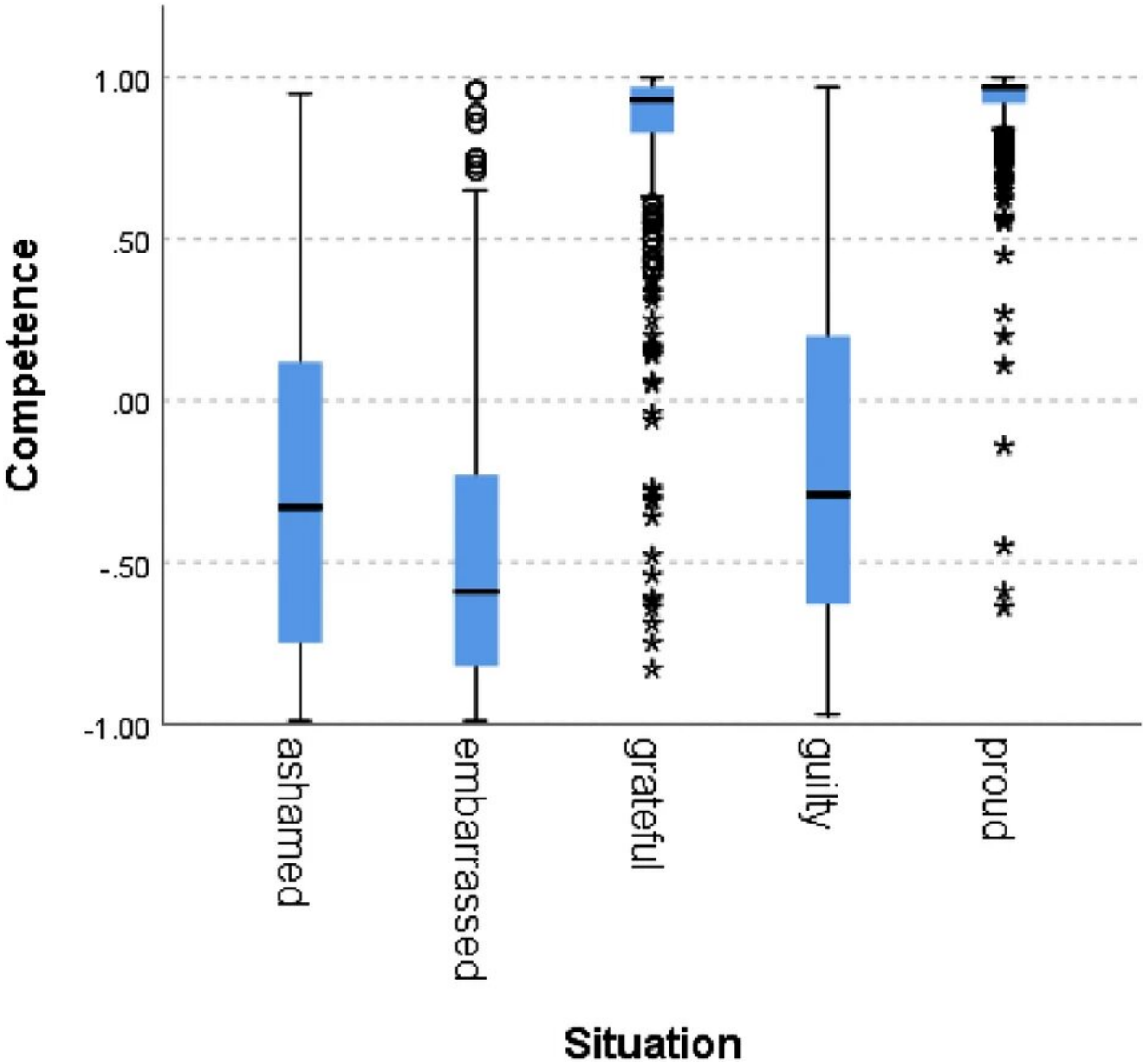


Researchers build an AI system to identify social norm violations

July 20 2023



Competence scores for the different situation types. Credit: *Scientific Reports*

(2023). DOI: 10.1038/s41598-023-35350-x

A researcher at Ben-Gurion University of the Negev has designed an AI system that identifies social norm violations. The project is one of the first to tackle the automatic identification of social norm violations. While many social norms exist worldwide, social norm violation boils down to only a few general categories.

Prof. Yair Neuman and his engineer Yochai Cohen built the system using GPT-3, zero-shot text classification, and automatic rule discovery. The system used a binary of ten social emotions as categories.

DARPA commissioned The Computational Cultural Understanding (CCU) program to create cross-cultural language understanding technologies to improve a Department of Defense operator's [situational awareness](#) and interactional effectiveness. Cross-cultural miscommunication not only derails negotiations, but also can be a contributing factor leading to war, according to DARPA's explanation of the rationale for the program.

Their findings were published recently in *Scientific Reports*.

Prof. Neuman and his engineer trained the system to identify ten social emotions: competence, politeness, trust, discipline, caring, agreeableness, success, conformity, decency, and loyalty. The system successfully characterized a written situation under one of these ten classifiers and could perceive if it was positive or negative.

The system was tested on two massive datasets of short texts and empirically proved the validity of the models.

"This is a preliminary work, but it provides strong evidence that our approach is correct and can be scaled up to include more [social norms](#)," says Prof. Yair Neuman.

Prof. Neuman is head of The Functor Lab in the Department of Cognitive and Brain Sciences at BGU.

More information: Yair Neuman et al, AI for identifying social norm violation, *Scientific Reports* (2023). [DOI: 10.1038/s41598-023-35350-x](https://doi.org/10.1038/s41598-023-35350-x)

Provided by Ben-Gurion University of the Negev

Citation: Researchers build an AI system to identify social norm violations (2023, July 20) retrieved 28 April 2024 from <https://techxplore.com/news/2023-07-ai-social-norm-violations.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.