



Robot discovers that lying about a betrayal helps to rebuild trust

33

TECHNOLOGY 29 March 2019



Should you trust a robot?
Sebo/Krishnamurthi/Scassellati

By Donna Lu

Betraying someone and lying about it is pretty low – especially if you're a robot. An experiment on interactions between humans and robots reveals just how easy it is for robots to regain trust after a misdeed by fibbing. The *devious robot* in question is NAO, a 58-centimetre-high humanoid that moves and interacts with people. Sarah Sebo at Yale University and her colleagues arranged for 82 people to compete for points against NAO in an asteroid-shooting computer game.

In some rounds, a special asteroid blaster was awarded to either the human or NAO. It could be used to get bonus points when shooting an asteroid or for temporarily immobilising the opponent, yielding no points.

Ahead of 10 games played together, the robot promised it wouldn't immobilise the player, but betrayal was programmed into its nature – it always broke its promise in round three.

To test the best route back from betrayal, the robot framed the immobilisation as a mistake with half of the human players. "Oh no! I hit the wrong button," it would say. With the other players, it yelled: "Yes! You're immobilised!".

Read more: [Replacing your boss with a cruel robot could make you concentrate more](#)

It then either apologised or denied it had broken its promise to the players in each group.

People were twice as likely to take revenge by immobilising the robot in the following round if they thought the robot's betrayal was on purpose. Similarly, if NAO simply denied its actions, people were also twice as likely to seek revenge immediately afterwards.

"Personally, I think robots should be 100 per cent honest and not lie ever," says Sebo, but the results suggest an advantage in lying to a person to optimise their trust. Framing a betrayal as incompetence could lead to a robot being treated better by people.

It is an ethical dilemma that we will increasingly face as robots become more sophisticated and humans start to view them as capable of having independent motives, says Sebo.

The team presented the work at the [International Conference on Human-Robot Interaction](#) in Daegu, South Korea, earlier this month.

More on these topics: [robot](#)

Magazine issue 3224, published 6 April 2019

33



DOWNLOAD

BUY IN PRINT

< Previous article

A crucial population of lions has lost much of its genetic diversity

Next article >

Mathematicians have found a new way to multiply two numbers together

MORE FROM NEW SCIENTIST



Offspring from older sperm are fitter and age more slowly



Russia's plan to unplug from the internet shows cyberwar is escalating

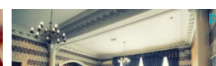
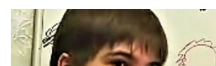


The last black leopard photographed in Kenya was born in New York



The truth about intelligence: What makes someone smarter than others?

PROMOTED STORIES



SCOTTISH MORTGAGE INVESTMENT TRUST
Wanted. Dreamers, visionaries and revolutionaries.
FIND OUT MORE

- TRENDING LATEST VIDEO FREE
- We contain microbes so deeply weird they alter the very tree of life 1
- David Attenborough finally talks climate change in prime time BBC slot 2
- We may finally know what causes Alzheimer's - and how to stop it 3
- Dead whale found with 40 kilograms of plastic in its stomach 4
- New species of human discovered in a cave in the 5



Information Document is available by contacting us.

> FIND OUT MORE



[Gallery] Genius Russian Boy Shares A Prediction About Space That's Making Experts
(Money Versed)



How To Improve Diabetes Naturally?
(ouremedy.com)



Glasgow - Alexander Thomson
(Booking.com)



Don't Trust Equity Release? Over 60s Should Know The Truth
(Telegraph Equity Release)

Recommended by @utbrain | >

Information Document is available by contacting us.

> FIND OUT MORE



[Contact us](#) | [Help](#) | [About us](#) | [Privacy & cookies](#) | [Terms & conditions](#)

[Advertise](#) | [Write for us](#) | [Events](#) | [Science jobs](#) | [Syndication](#) | [RSS feeds](#)

[Gift subscriptions](#) | [Student subscriptions](#) | [Educational subscriptions](#) | [Corporate subscriptions](#)

GET THE APP



FOLLOW US



© Copyright New Scientist Ltd.

[Back to top](#)

