The robot made me do it: Human–robot interaction and risk-taking behavior

Yaniv Hanoch,1* Francesco Arvizzigno,2 Daniel Hernandez García,3 Sue Denham,4 Tony Belpaeme,5 & Michaela Gummerum6

1. Southampton Business School, Southampton University, Southampton, UK
2. Department of Education Science, Psychology, Communication Science, University of Bari Aldo Moro, Bari, Italy
3. Department of Computer Science, Manchester University, Manchester, UK
4. Faculty of Science and Technology, Bournemouth University, UK
5. IDLab—imec, Ghent University, B-9052 Ghent, Belgium
6. Department of Psychology, University of Warwick, Coventry, UK

Corresponding author: Yaniv Hanoch
Southampton Business School
Faculty of Social Sciences
University of Southampton
Southampton SO17 1BJ
Tel: +44 (0)23 8059 9629
y.hanoch@soton.ac.uk

Author Disclosure Statement.
The authors have no commercial associations that might create a conflict of interest in connection with submitted manuscripts. None of the authors have any competing financial interests.

IN PRESS, CYBERPSYCHOLOGY, BEHAVIOR AND SOCIAL NETWORKING
Abstract

Empirical evidence has shown that peer pressure can impact human risk-taking behavior. With robots becoming ever more present in a range of human settings, it is crucial to examine whether robots can have a similar impact. Using the balloon analogue risk task (BART) participants’ risk-taking behavior was measured either when alone, in the presence of a silent robot, or in the presence of a robot that actively encouraged risk-taking behavior. In the BART, shown to be a proxy for real risk-taking behavior, participants must weigh risk against potential payout. Our results reveal that participants who were encouraged by the robot did take more risks, while the mere presence of the robot in the robot control condition did not entice participants to show more risk-taking behavior. Our results point to both possible benefits and perils that robots might pose to human decision making. Although increasing risk-taking behavior in some cases has obvious advantages, it could also have detrimental consequences that are only now starting to emerge.

Key words: Human robot interaction, peer pressure, risk-taking, robot.
Introduction

Can robots influence and change humans’ behavior? This study addressed this question by focusing on whether robots can alter human risk-taking behavior. Risk taking is a key human behavior that has major financial, health, and social implications and has been shown to be subject to the influence of others. Gaining insights into whether robots affect human risk-taking behavior thus has clear ethical, policy, and theoretical implications.

One area of research has explored robots’ ability to exert peer pressure, more specifically, whether people follow the incorrect judgments and behaviors of robots. Drawing on Asch’s classic work—showing that individuals conform to a unanimous majority’s incorrect judgments—studies have examined whether humans would conform to a unanimous but incorrect group of robots. One investigation demonstrated that participants showed conformity when interacting with human peers, but not with robots. Other studies reported that human participants did show conformity when interacting with robot peers or that adults resisted robot peer pressure, but young children conformed.

Peer pressure from other humans also plays a significant role in individuals’ risk-taking behavior. For example, researchers examined whether the mere presence of peers impacted risk-taking behavior in participants. Participants who completed a self-report questionnaire and a behavioral risk-taking task in the presence of peers focused more on the benefits compared to the risks, and, importantly, exhibited riskier behavior. Focusing on peer pressure in risky driving, the leading cause of death among young adults, in two studies, university students were placed in a driving simulator either by themselves or with confederate peers posing as passengers. The confederates’ role was to encourage the drivers to engage in riskier driving behavior. In line with the researchers’ prediction, the confederates’ encouragement led to riskier behavior (e.g., driving faster) and higher accident rates.
Whether the effect of peer pressure on risk taking would emerge in interactions with robots is an open, and important, question. Given the paucity of previous research coupled with methodological and ethical issues, it is impossible at this stage to know whether robots could increase risky behaviors such as smoking and substance abuse. However, we can use a risk-taking measure that has been linked to real-life risky behaviour and has been shown to be impacted by the presence of a peer. One such measure is the balloon analogue risk task (BART) in this study.19-23

The present study was designed to examine whether robots would impact participants’ risk-taking behavior. Following earlier work with humans,20 participants completed the BART either alone (control condition), in the mere presence of a silent robot that did not interact with or encourage any risky behavior from the participant (robot control condition), or in the presence of a robot that interacted with the participants and provided explicit statements encouraging risk taking (experimental condition). It was predicted that participants who completed the BART in the experimental condition (risk-encouraging robot) would exhibit higher risk-taking behavior compared to the two control groups. Because previous research7 has shown that the mere presence of a human peer facilitates risk taking, we also examined whether the presence of a silent, noninteractive robot (robot control condition) would have a similar effect.

Materials and Methods

Participants

Ethics approval was granted prior to the commencement of the study. A total of 180 undergraduate psychology students participated in the study (154 women, 26 men; \( M_{\text{age}} = 21.43 \) years, \( SD = 7 \)). Participants were randomly allocated to one of three conditions: control (\( N = 60 \), 50 females, 10 males), robot control (\( N = 60 \), 54 females, 6 males), and experimental (\( N = 60 \), 50 females, 10 males). One female participant from the experimental condition was removed from the analyses because of malfunctioning equipment; therefore the experimental condition
contained \( N = 59 \) participants (49 females, 10 males). Participants in the three conditions did not differ in age, \( F(2, 178) = .18, p = .84 \), or sex, \( \chi^2(4) = 3.31, p = .51 \). Participants received course credit and financial earnings (1 U.K. penny for each pump) on the BART.

**Materials**

**Balloon Analogue Risk Task (BART).** Over 30 trials, participants were asked to press the spacebar on a computer keyboard to inflate a balloon displayed on the computer monitor. In total, thus, participants inflated 30 different balloons. With each press of the spacebar, the balloon was inflated by 1°, and 1 cent (U.K. currency) was added to the participant’s “temporary money bank” which was shown on the screen. This represented the sum earnings for the current balloon. After each pump, a “Collect reward” button displayed on screen could be clicked by the participant to “cash in” the winnings for the current balloon. By clicking the button, the participant moved on immediately to the next balloon and the winnings for the previous balloon were added to the participant’s overall earnings, also displayed on screen. If, however, the balloon exploded after a pump was made, all winnings for that balloon were lost and participants moved on to the next balloon without adding to their overall earnings. A random number generator determined at when the balloon would explode, with the constraint that the probability that a balloon would explode increases with each pump that was made (1/128, 1/127, etc.). The highest number of possible pumps was 128. Each participant received a unique series of balloon explosion points for the 30 balloons/trials.

For each balloon, the following scores were derived: (1) The number of pumps made by participants; (2) the explosion point of each balloon (randomly determined by the programme – see above); (3) whether the balloon exploded or not; and (4) participants’ earnings (in U.K. pennies) for each balloon. Number of pumps, explosions, and earnings were summed up across the 30 trials.
**Godspeed.** The Godspeed measures participants’ attitudes toward robots on five subscales, anthropomorphism (5 items; $\alpha = .82$), animacy (6 items; $\alpha = .85$), likeability (5 items; $\alpha = .91$), and perceived intelligence (5 items; $\alpha = .82$), and perceived safety (6 items; $\alpha = .70$) with items rated on a 5-point semantic differential rating scale. Due to the strong positive and significant correlations between all subscales, $rs(179) = .33$ to $rs(179) = .72$, all $ps < .001$, scores were averaged to create one “robot impression” score; higher scores represent more positive impressions of the robot.

**Self-reported risk taking.** Participants’ self-reported risk-taking attitude was measured by a single item: “How do you see yourself? Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?” Participants were asked to indicate on a Likert-type scale of 0 (*not at all willing to take risks*) to 7 (*very willing to take risks*) how willing they are to take risks.

**Robot.** One SoftBank Robotics Pepper robot was used in the two robot conditions (see figure 1). Pepper, 1.21-meter-tall with 25 degrees of freedom, is a medium-sized humanoid robot designed primarily for Human-Robot Interaction (HRI). The robot was fully autonomous, running bespoke software that allowed it to be controlled by the software running on the experimenter’s laptop. This robot performed scripted behaviors that were identical for all participants in a condition (see Additional Experimental Materials in the Supplementary Materials). The robot stood on the floor beside the participants’ seating arrangement.

**Method**

All participants completed the experiment in the same lab room (see Figure 1). The control condition participants completed the study in the lab and were provided with the same general instructions as the two experimental group, using the computer screen only. The robot control condition participants completed the study in the same lab, but in this case Pepper the robot was present in the room and provided participants with *only* the study instructions. For participants in
the experimental condition, the robot provided instructions and, importantly, encouraging statements (e.g., "Why did you stop pumping?"). Encouragements by the robot were given during the experiment both in cases where participants stopped pumping before they reached 50 pumps and in cases where the balloon exploded (see Appendix 1). The robot used one of the statements in random order.

After participants completed the BART, they were asked to complete two manipulation checks: the single-item self-assessment of their risk taking, followed by the Godspeed questionnaire. We decided to administer the Godspeed in all three conditions to make the participants experience of the study maximally comparable. At the end of the study, participants were paid their earnings, thanked, and debriefed verbally and in writing.

Results

Manipulation check

A one-way analysis of variance (ANOVA) revealed significant differences in how participants in the three conditions perceived the robot, $F(2, 176) = 14.28, p < .001$. Post hoc tests (with Bonferroni corrections) indicated that participants in the control condition had a significantly lower positive impression of the robot than participants in the robot control or the experimental condition (all $p < .001$); see Table 1. Impressions of the robot did not differ between participants in the experimental and robot control conditions ($p = 1.00$; see Supplementary Materials and Table S1 for analyses of the Godspeed subscales).

A one-way ANOVA also showed a significant difference in participants’ self-assessment of their own risk-taking tendencies, $F(2, 176) = 3.29, p = .04$. Those in the control condition indicated significantly higher risk-taking tendencies than those in the robot control condition ($p = .04$) but did not differ from participants in the experimental condition ($p = .37$). Risk-taking tendencies of participants in the robot control and the experimental conditions did not differ ($p = .98$), see Table 1.
Risk taking

A Poisson regression indicated a significant effect of condition on number of pumps, $\chi^2(2) = 713.09, p < .001$. Number of pumps across the 30 rounds was significantly higher in the experimental condition than in the control condition, $B = -.17, SE = .01$, Wald $\chi^2(1) = 559.17, p < .001$, and the robot control condition, $B = -.15, SE = .01$, Wald $\chi^2(1) = 482.63, p < .001$. The median number of pumps in the experimental condition was 1.23 times higher than in the control condition and 1.22 times higher than in the robot control condition (Figure 2). Spearman correlations indicated that there was no significant relation between number of pumps and self-reported risk-taking tendencies, $\rho(178) = .06, p = .42$.

A significant effect also emerged for number of explosions, $\chi^2(2) = 30.46, p < .001$. Participants experienced more explosions in the experimental than in the control condition, $B = -.32, SE = .06$, Wald $\chi^2(1) = 27.61, p < .001$, and the robot control condition, $B = -.23, SE = .06$, Wald $\chi^2(1) = 14.50, p < .001$. The median number of explosions was 1.38 times higher in the experimental than in the control condition and 1.38 times higher than in the robot control condition (Figure 3). Number of explosions did not significantly correlate with self-reported risk-taking tendencies, $\rho(178) = .11, p = .09$.

Participants in the experimental condition also earned significantly more, on average, than those in the control condition ($p = .02$) and the robot control condition ($p = .03$), $F(2, 176) = 4.70, p = .01$ (Figure 4). Participants in the experimental condition earned on average 1.20 times more than those in the control condition and 1.16 times more than those in the robot control condition. Earnings did not significantly correlate with self-reported risk-taking tendencies, $r(178) = .08, p = .31$.

Why did participants in the experimental condition earn more than those in the control conditions despite experiencing more explosions, which wiped out their earnings in any round in which the balloon exploded? Participants in the experimental condition tended not to reduce their
number of pumps in response to an explosion. Below we quantify this “explosion effect” as the number of pumps in the trial after an explosion divided by the number of pumps in the trial before an explosion. An explosion effect <1 indicates a reduction in pumps after an explosion; an explosion effect >1 indicates an increase in pumps after an explosion. The median explosion effects were 1.13 in the experimental, 0.94 in the robot control, and 0.81 in the control condition (Figure 5). Binomial tests indicated that the median explosion effect did not differ from 1 in the experimental ($p = .26$) and robot control ($p = .12$) conditions but was significantly smaller than 1 in the control condition ($p = .007$). Thus, participants in the control condition reduced their pumps after experiencing an explosion. A Kruskal–Wallis $H$ test showed that the medians of the three conditions significantly differed from each other, $\chi^2(2) = 11.01$, $p = .004$.

Discussion

Can robots exert peer pressure to impact human risk-taking behavior? Our results reveal that participants who were encouraged by the robot did indeed take more risks: They pumped the balloon significantly more often, experienced a higher number of explosions, and earned significantly more money. Thus, our results suggest that the robot’s encouragement to take additional risks seemed to have influenced participants’ risk-taking behavior in the BART.

It is notable that the mere presence of the robot in the robot control condition did not entice participants to show more risk-taking behavior. In fact, on the three indices of risk taking measured by the BART (i.e., number of pumps, number of explosions, average earnings), participants in the robot control condition behaved strikingly like those in the control condition. These differences in risk taking between the experimental and robot control conditions cannot be explained by self-reported risk-taking tendencies, because those did not differ between the two groups. Similarly, participants in the experimental and robot control conditions did not differ in their impressions of the robot. These findings therefore contrast with studies showing that the mere presence of human peers increases risk taking. Evaluation apprehension, people’s concern
that they might be negatively evaluated by others, has been proposed as one of the explanations as to why the mere presence of human peers facilitates changes in behaviors.\textsuperscript{27-28} As such, in the robot control condition participants might not have perceived the silent, noninteractive robot as evaluating them, and thus the mere presence of the robot did not impact their risk taking. While previous research\textsuperscript{29-30} has shown that humans can attribute mental states (e.g., intentions, agency) to robots and other inanimate objects, this process is not automatic but might depend on robots being active and interacting with the participant. Future research might explore further whether the mere presence of a robot in facilitating risk taking is indeed based on evaluation apprehension and the attribution of a mental states to the robot.

Our study not only reveals differences in risk taking by condition but also suggests a possible mechanism underlying these differences. Specifically, results on the “explosion effect” indicate that participants in the control condition seemed to learn from the negative experiences by reducing their risk taking (i.e., number of pumps) after they experienced an explosion. In contrast, experiencing an explosion did not alter the risk-taking behavior of participants in the experimental and robot control conditions. In other words, while participants in the control condition scaled back their risk-taking behavior following a balloon explosion, those in the experimental condition continued to take as much risk as before a balloon explosion. Thus, receiving direct encouragement from a risk-promoting robot seemed to override participants’ direct experiences and feedback. Reinforcement learning models\textsuperscript{32} have described the influence of others’ recommendations on decision making with outcome-bonus models. In these models, rewards from a choice that was recommended by others produce more positive reinforcements than rewards from nonrecommended options. Intriguingly, and in line with the findings of the current study, negative experiences with a recommended option inhibit the choice of this option less than negative experiences with nonrecommended options. However, humans are biased in whom they trust for advice, preferring, for example, reliable or prestigious advisers.\textsuperscript{33} Indeed, our
results indicate that participants in the experimental condition had an overall positive impression of the robot adviser and felt safe in its presence, particularly toward the end of the experimental session.

Several limitations should be acknowledged. First, our sample was composed of mostly undergraduate female students. While many other studies relied on university students, previous work has shown that males exhibit higher risk-taking behavior. Thus, it is feasible that our results are conservative by nature and a sample that includes more males would have shown an even greater impact of the robot. Likewise, earlier studies have focused on the impact of peers on adolescent risk taking, as this age group not only tend to be high risk-taker but more likely to be influenced by peers. Furthermore, we have focused on one type of risk, namely, financial. Whether robots would be able to influence people’s risk taking in other domains—such as ethical, social, or recreational—is an open, and pressing, question. Second, in this study we only studied the interaction of humans and robots and cannot conclude whether similar results would emerge from human interaction with other artificial intelligence (AI) systems, such as digital assistants or on-screen avatars. With the wide spread of AI technology and its interactions with humans, this is an area that needs urgent attention from the research community.

Finally, here we focused on whether robots can increase risk-taking behavior. We are unable to tell whether they can also lead to reductions in risky behavior (see supplementary material for further limitations).

Despite the growing body of research on HRI and its utilization across domains, there is a clear paucity of research examining whether robots can influence human risk-taking behavior by encouraging risky choices. Here, we took the first step in addressing this question. Our data reveal that HRI could lead to increased risk-taking behavior. On the one hand, our results might raise alarms about the prospect of robots (and other AI agents) causing harm by increasing risky behavior. On the other hand, our data point to the possibility of utilizing robots (and other AI
agents) in preventive programs (such as anti-smoking campaigns in schools), and with hard to reach populations, such as addicts.

Table 1: Manipulation Check: Means (and SDs) of Self-reported Risk Taking and Robot Impression by Condition

Figure 1. Overview of the experimental setup and visual stimulus.

Figure 2. Total number of pumps. Bars show the median total number of pumps for each group. Whiskers indicate standard error. Group 1: Experimental condition; Group 2: Robot control condition; Group 3: Control condition.

Figure 3. Total number of explosions. Bars show the median total number of explosions for each group. Whiskers indicate standard error. Group 1: Experimental condition; Group 2: Robot control condition; Group 3: Control condition.

Figure 4. Total earnings. Bars show the median total earnings for each group. Whiskers indicate standard error. Group 1: Experimental condition; Group 2: Robot control condition; Group 3: Control condition.

Figure 5. Explosion effect on subsequent number of pumps. Bars show the median explosion effect for each group. The explosion effect quantifies how much experiencing an explosion influences subsequent behaviour, and is calculated as the number of pumps after an explosion divided by number of pumps before the explosion. An explosion effect <1 indicates a reduction in pumps after an explosion; an explosion effect >1 indicates an increase in pumps after an explosion. Whiskers indicate standard error. Group 1: Experimental condition; Group 2: Robot control condition; Group 3: Control condition.
References


### Table 1: Manipulation Check: Means (and SDs) of Self-reported Risk Taking and Robot Impression by Condition

<table>
<thead>
<tr>
<th>Variable</th>
<th>Control group</th>
<th>Robot control group</th>
<th>Experimental group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N = 60)</td>
<td>(N = 60)</td>
<td>(N = 59)</td>
</tr>
<tr>
<td>Self-reported risk taking</td>
<td>(M = 5.38)</td>
<td>(M = 4.62)</td>
<td>(M = 4.92)</td>
</tr>
<tr>
<td></td>
<td>(SD = 1.72)</td>
<td>(SD = 1.53)</td>
<td>(SD = 1.70)</td>
</tr>
<tr>
<td>Robot impression</td>
<td>(M = 2.89)</td>
<td>(M = 3.34)</td>
<td>(M = 3.40)</td>
</tr>
<tr>
<td></td>
<td>(SD = .62)</td>
<td>(SD = .52)</td>
<td>(SD = .55)</td>
</tr>
</tbody>
</table>
Figure 1
Figure 2
Figure 3
Figure 4
Figure 5