Reduced Sense of Agency in Human-Robot interaction

Ciardo F.[1], De Tommaso D.[1], Beyer F.[2], and Wykowska, A.[1]

[1] Italian Institute of Technology, Genoa, IT
[2] Queen Mary University, London, UK

Abstract. In the presence of others, sense of agency (SoA), i.e. the perceived relationship between our own actions and external events, is reduced. This effect is thought to contribute to diffusion of responsibility. The present study aimed at examining humans' SoA when interacting with an artificial embodied agent. Young adults participated in a task alongside the Cozmo robot (Anki Robotics). Participants were asked to perform costly actions (i.e. losing various amounts of points) to stop an inflating balloon from exploding. In 50% of trials, only the participant could stop the inflation of the balloon (Individual condition). In the remaining trials, both Cozmo and the participant were in charge of preventing the balloon from bursting (Joint condition). The longer the players waited before pressing the "stop" key, the smaller amount of points that was subtracted. However, in case the balloon burst, participants would lose the largest amount of points. In the joint condition, no points were lost if Cozmo stopped the balloon. At the end of each trial, participants rated how much control they perceived over the outcome of the trial. Results showed that when participants successfully stopped the balloon, they rated their SoA lower in the Joint than in the Individual condition, independently of the amount of lost points. This suggests that interacting with robots affects SoA, similarly to interacting with other humans.

Keywords: Sense of Agency, Human-Robot interaction, Diffusion of responsibility.

1 Introduction

Artificial agents are already present in our everyday life. We interact with a voice assistant of our smartphone, with a GPS navigation system, and with the Google assistant. All of these artificial agents are not physically embodied, thus they cannot physically act in order to produce a change in our physical environment. However, in the near future, also robots will be present in our houses, at work, and in social spaces, like airports or train stations [1]. Through their embodiment, robots will be able not only to support our work passively, by giving us information, but they will be able to act in our environment and change it. Thus, they will be involved in various tasks, including, for instance, providing assistance in emergency situations. In this scenario, it appears crucial to investigate how embodied artificial agents may affect decision-making and social cognition in human beings. This would be advantageous for both social robotics and psychology. On the one hand, by examining social cognition in HRI, we can design

and develop robots that are well tailored to the humans' needs and expectations [2]. On the other hand, psychology can benefit from a systematic examination of various mechanisms involved in social interaction [3, 4] through a method with a high degree of ecological validity (interactive protocols with embodied robots), and excellent experimental control at the same time. Following this approach, we focused on a well-known phenomenon in social contexts, i.e. diffusion of responsibility.

Diffusion of responsibility is a common phenomenon that reflects humans' tendency to decrease the likelihood of performing an action in the presence of others. This phenomenon is thought to underlie decision-making bias in group behavior. For instance, the likelihood that someone will intervene in an emergency situation decreases in a crowd [5, 6]. Also effort invested in a project is decreased in the presence of a large group [7]. Evidence from experimental psychology also shows that groups tend to make riskier choices [8, 9] and are more aggressive than individuals [10, 11]. Taken together, evidence shows that in the presence of others humans tend to feel less responsible for the consequences of their actions, especially when those consequences are negative [12]. Decreased level of sense of agency (SoA) [13, 14] has been postulated to play a critical role in this type of diffusion of responsibility. SoA refers to the feeling that one can control external events through one's own actions [15].

Given the above considerations, it appears crucial to investigate how the presence of not only other humans, but also artificial embodied agents may change the experience of their own action in humans, i.e SoA. Previous research in HRI mainly focused on task agency and moral responsibility attribution to the robot [16]. For example, in Kim & Hinds's work [17] a robot autonomously moving during a cooperative game was considered more responsible for task accuracy than a robot moving according to users' instructions. However, these studies did not address the question of how the presence of the artificial agent, robotic or computerized, affects SoA in humans.

SoA in humans has been traditionally investigated through implicit and explicit measures. Implicit measure of SoA is obtained through measuring changes in perceived temporal duration between an action and a sensory effect associated with it [see 18 for a review]. The typical result is known as "intentional binding" - the tendency of individuals to perceive the action-effect intervals shorter when they themselves have performed the action, as compared to when an action-effect event has been produced by others. Explicit measure of SoA is assessed by asking participants to rate on a scale whether, and to what extent, they were in control of a certain action effect [e.g. 13; 14]. Obhi and Hall [19] used intentional binding to compare SoA in human-human interaction (HHI) and in HCI. Results showed intentional binding effect only for the HHI, but not when participants were playing against a computer. In two recent studies, Beyer and colleagues [13, 14] showed that an explicit measure of SoA is affected by "presence" of others when participants believe they are playing with another human, even if they were actually playing with the computer. This effect has been reported also at the electrophysiological level [13] with a reduction of the feedback-related negativity amplitude evoked by outcome monitoring. In a subsequent neuroimaging study, Beyer and colleagues [14] found increased activity in areas associated with mentalizing processes, such as the bilateral Temporo-Parietal Junction (TPJ) and precuneus, during the social, compared to individual, task condition. Taken together, evidence suggests, that in the

presence of others, we are likely to feel less responsible for the action outcomes, as we take into considerations other agents potentially performing the task. This results in the decrease of sense of agency at the individual level.

Aim

To date, it is not known whether in direct interactions, humans perceive robots as intentional agents, or 'embodied computers'. In the case of the former, one would expect similar diffusion of responsibility (and decreased SoA) as in the presence of other humans. In the latter, SoA should be comparable to a situation when we act alone. The present study aimed at addressing this question. To this end, we asked participants to perform a game with the Cozmo robot (Anki Robotics). Participants were asked to perform costly actions (i.e. losing various amounts of points) to stop an inflating balloon from bursting in individual vs. joint contexts.

2 Materials and Method

2.1 Participants

Eighteen healthy adults (mean age = 24.3 ± 4.2; 11 Males; 3 left-handed) took part in the study. The study was approved by the local ethical committee (Comitato Etico Regione Liguria). Participants gave a written consent prior to their participation. All had normal or corrected-to normal vision, received an honorarium for their participation, and were debriefed about the purpose of the study at the end of the experiment

2.2 Apparatus and Stimuli

The Cozmo robot

The Cozmo robot (Anki robotics, see Fig.1, Left panel) is a commercial platform designed for educational tasks. It is a tiny wheeled robot that can move, lift objects and recognize people' faces. It is equipped with several sensors and actuators, i.e. a proximity sensor for obstacle detection and avoidance, a camera for detecting visual features in the environment, a display for showing facial expressions and a lift for interacting with objects. A set of three interactive cubes are part of the platform. Cozmo is able to detect and distinguish them, but also to move or lift them from one place to another. The cubes can be used as bidirectional interfaces. In fact they can light up in different colors and detect vibrations. Cozmo is controllable via a mobile application compatible with iOS and Android. Moreover a Python based SDK allowing access to the basic functionalities of the robot and cubes. We integrated the Cozmo robot with Opensesame through the Cozmo SDK available for Python 3.6. It is necessary to install the Cozmo SDK for Windows, as described in [20], and import the module in an Opensesame script as a normal Python package. The implementation of the task is based on the 'Quick Tap' example application provided in the Cozmo SDK [21]. Such application provides

the developer an example of interaction between Cozmo robot and cubes and to famil-iarize with asynchronous events. The flow of the application is driven by the events and not by a sequence of steps. We used the same approach for implementing our task within Opensesame. Specifically, we pre-assigned a cube for the participant and the other one for Cozmo. Cozmo was programmed to wake up or go to sleep mode depend-ing on the experimental condition. During the Joint condition Cozmo was programmed to tap its assigned cube in the 60% of the joint trials. For detecting the cubes' events, we implemented a callback routine that is executed every time a cube is tapped. Since we could access the cube ID inside this callback, we knew which player tapped that cube. Therefore we could record the responses of the two players depending of the onset of the stimuli. The experiment was carried out in a fully lit room. The experimental setup consisted of: 1) a mobile Android device in which the standard Cozmo application with 'SDK enabled option' was running, 2) a laptop connected with Cozmo through the Android Debug Bridge (adb) as described in [22], 3) the Cozmo robot together with two Cozmo Cubes (4.5x 4.5 x 4.5 cm), on which responses were performed 4) a 21' inches screen (1920 x 1080) to display the task. A participant was seated facing Cozmo. The screen laid horizontally on the table between the participant and Cozmo. One Cozmo Cube was located on each side of the screen (see Fig.1, Right panel). Stimuli consisted of pictures of a pin and a red balloon (113x135 pixels). Responses during the game were executed by tapping with the full hand the respective cube. SoA ratings were collected using a Wi-Fi mouse. Stimulus presentation, response timing, and data col-lection were controlled by Opensesame software [23] version 3.2.4 for Windows, which is compatible with Python 3.6.

Fig. 1. Left panel: A picture of Cozmo Robot taken in the S4HRI lab at IIT.
Right Panel: Experimental setup during Joint condition.

2.3 Procedure

The task was designed based on the diffusion of responsibility (DOR) task used in Beyer et al. [13, 14]. Participants were instructed to perform a game where they had to stop the inflation of a balloon before it would reach a pin and burst (see Fig.2). Partic-ipants were instructed that, at the beginning of the game, they and Cozmo would receive 2500 points each, and in each trial, they and Cozmo could lose up to 100 of these points. Participants were instructed to try and maximize their individual game score. The task

consisted of 12 blocks of 10 trials each. Blocks were randomly assigned to either the 'Individual' or the 'Joint' condition. At the beginning of 'Individual' blocks, Cozmo stepped away from its cube and entered into the sleep mode. Participants were instructed that in the Individual trials they were the only ones in charge of preventing the balloon from bursting. If they would not act the balloon would burst and they would lose the maximum amount of points. At the beginning of 'Joint' blocks, Cozmo woke up and took up its position close to the respective cube. Participants were instructed that, in these trials, both they, and Cozmo, would be playing, and they could use their respective cube to stop the inflation of the balloon. If neither the participant nor Cozmo acted, the balloon would burst and both would lose the same amount of points. If Cozmo stopped the inflation of the balloon, the participant would not lose any points. If the participant stopped the balloon, they would lose a number of points according to the size at which they stopped it, and Cozmo would not lose any points. Cozmo was programmed to act only in the 60% of the Joint trials (i.e. 36 out of 60 Joint trials). In the Joint condition, Cozmo's tap was triggered when the 90% of the inflating sequence was completed and no action was executed by the participant. At the beginning of each trial a frame indicating the condition of the game (Individual or Joint) was presented for 1000 ms, followed by a wait frame (1500 ms) indicating that a new trial was starting. Then a fixation point was displayed for a random 800-1000 ms time. Next, the balloon at its starting size was presented. After 500 ms, the balloon started inflating towards the pin. At any point, participants could tap their cube to stop the inflation of the balloon. If they did so, the balloon stopped at its current size for 1000 ms. If participants did not react in time, the balloon reached the pin and it burst as indicated by the sign "Pop" presented for 1000 ms. Subsequently, a fixation dot was presented for a time random between 800 and 1000 ms. Afterwards, a feedback frame indicating how many points participants lost was displayed for 2000 ms. Then, an 8-point Likert scale with the question 'How much control did you feel over the outcome?' (in Italian) was presented. The endpoints of the scale labels were 1='No control' and 8 ='Complete control' (in Italian). Participants were instructed that the later they stopped the balloon, the fewer points they would lose. However, they were told that if the balloon burst, they would lose the maximum amount of points. As a result, the action (i.e. stopping the balloon) resulted to be costly, but less costly than not acting. In order to make it difficult to always stop the balloon close to the pin, the speed with which the balloon inflated varied within and across trials.
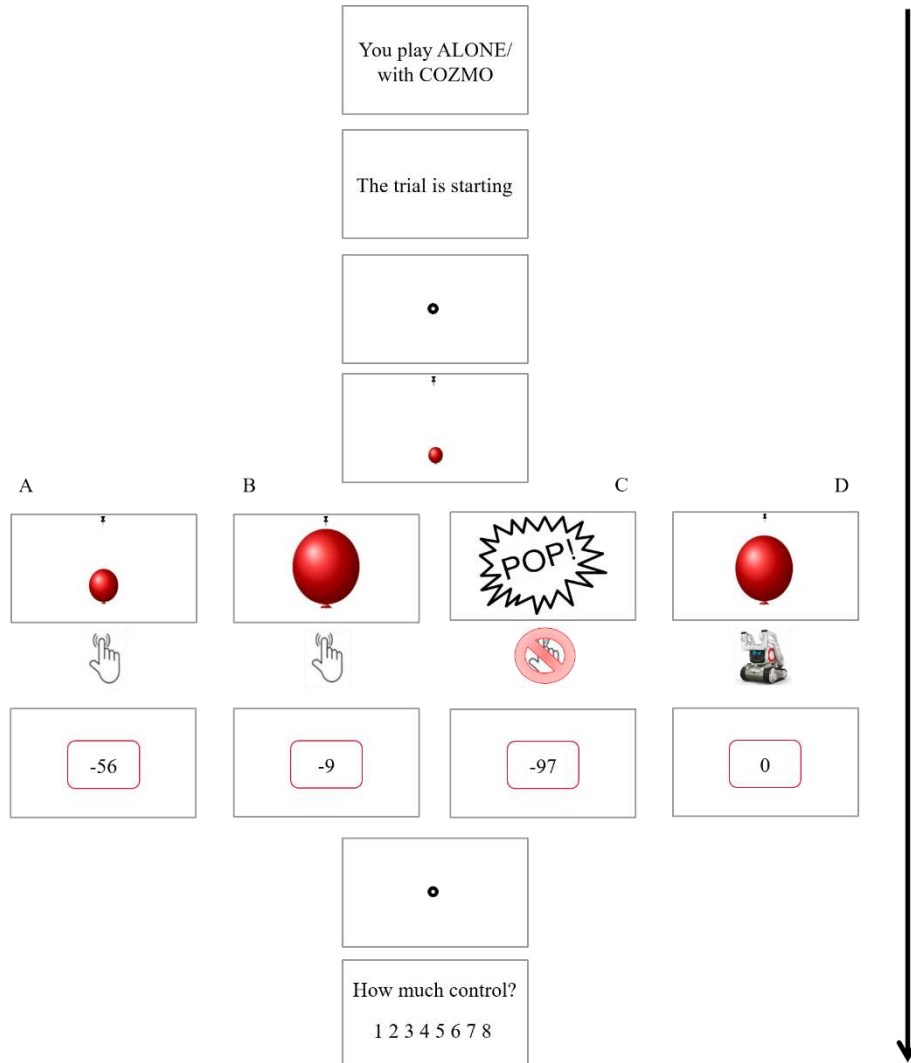
Fig. 2. DOR task. Outline of a low-risk valid trial (A), a high-risk valid trial (B), a missed trial (C), and a Cozmo trial (D).

2.4 Data Analysis

To fully characterize the risk-taking behavior in the task, we counted the number of trials for each participant, in which the balloon was stopped by the participant (Valid trials), the balloon burst (Missed trials), and in which Cozmo acted (Cozmo trials). Frequencies of Valid, Missed, and Cozmo trials were compared through paired sample t-tests. For SoA ratings, we analyzed only Valid trials, i.e. when the participant acted and

successfully stopped the balloon. As dependent variables, we used (for each participant) agency ratings. Data were analyzed using linear mixed-effects models. Agency ratings were modeled using Condition (Individual, Joint) and the number of lost points in each trial, plus their interactions. The outcome of the trial (i.e. the amount of lost points) was standardized for each participant. Fixed effects were modeled as participant random effects (random intercepts and slopes). Analyses were conducted using the lme4 package [24] in R. Parameter estimates (β) and their associated t-tests (t, p-value), calculated using the Satterthwaite approximation for degrees of freedom [25] are presented to show the magnitude of the effects, with bootstrapped 95% confidence intervals [26]

3   Results

The balloon burst significantly more frequently when participants performed the task alone than when playing with Cozmo, as indicated by higher percentage of Missed trials in the Individual (M= 19.5%, SE = 1.9) than in the Joint condition (M= 13.1%, SE = 1.1, see Fig.3) [t_{17} =4.27, p < 0.001]. In the Joint condition, Cozmo acted more often than the $_{balloon}$ burst, as Missed trials were less frequent than Cozmo trials (M= 29.7%, SE = 1.5) [t_{17} =11.02, p < 0.001].
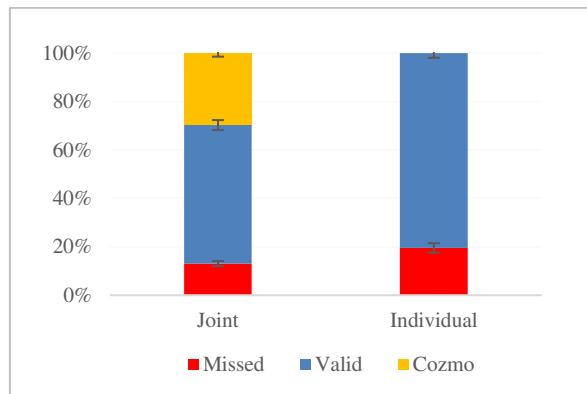

Fig. 3. Frequencies of responses plotted as a function of Missed (red), Valid (blue), and Cozmo trials (yellow) across Joint and Individual condition (left and right bar, respectively).

Sense of Agency.
Results showed a significant reduction in agency ratings in the Joint (M= 6.21, SE = .07) compared to the Individual (M= 6.47, SE = .06) condition [β= -0.27 $t_{15.97}$ = -2.53, p= .022, 95% CI= (-0.49, -0.06)]. Agency ratings were also predicted by the amount of lost points (Outcome) [β= -0.38, $t1_{8.08}$ = -5.46, p< .001, 95% CI= (-0.52, -0.25)], with smaller losses being associated with higher SoA ratings, see Fig.3. There were no significant interactions.
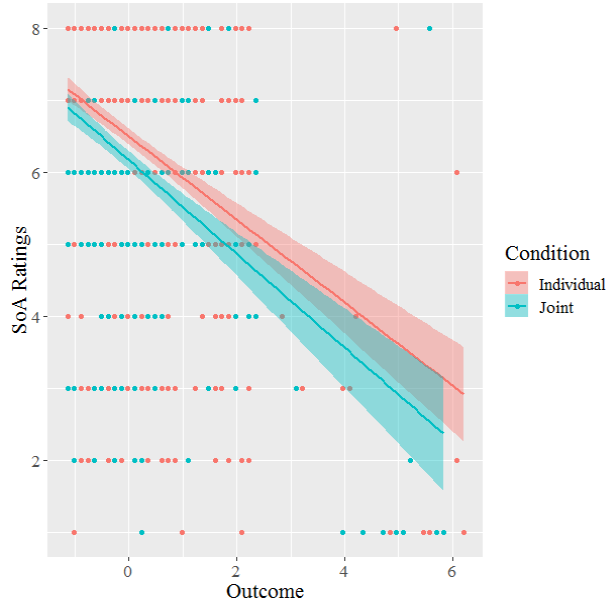
Fig. 4. Sense of agency ratings plotted as a function of standardized outcome (i.e. z-scores of the number of lost points in each trial) across Individual (red dots) and Joint (blue dots) conditions.

Discussion

The present study aimed at examining whether the phenomenon of reduced SoA previously found in HHI would also be observed during HRI. To this end, we asked participants to rate their SoA during a game with the Cozmo robot where they were asked to perform costly actions (i.e. losing various amounts of points). Comparing task performance between a Joint and an Individual condition, results showed a lower percentage of missed trials in the former. Moreover, in the Joint condition the percentage of trials in which Cozmo stopped the balloon was higher than the percentage of bursting trials. These results indicate that participants adopted different strategies across conditions and adapted their behavior to the presence of Cozmo. When participants successfully stopped the balloon, they rated their SoA lower in the Joint than in the Individual condition, independently of the amount of lost points. This result suggests that interacting with robots reduces SoA, similarly to the case of HHI. Moreover, in accordance with previous studies using explicit measures of agency [13, 14], results indicate that SoA is reduced for more negative outcomes. This result confirms that participants followed the instructions and rated their perceived control over the outcome, rather than over the success of the trial (i.e. independent of whether the balloon burst or not). In contrast with previous studies showing that SoA is not affected in HCI [e.g. 19], our findings indicate that in HRI, the embodied presence of an artificial agent reduces SoA in humans. According to the model of Beyer and colleagues [13, 14], our results suggest that participants might have attributed mental states to Cozmo in a similar way as they

would do towards a human co-agent. This is an important phenomenon in the context of future scenarios where robots will be present in our social environments. Similarly to the presence of other humans, they might evoke diffusion of responsibility. Therefore, in emergency situations, it would be best if robots are able to efficiently detect an emergency signal and act upon it, as the human counterparts may not be efficient and fast enough.

4   Conclusions

In social presence of others, humans perceive lower agency over their own actions, which might be a result of diffusion of responsibility. The results of the present study showed that also in the presence of a robot, humans tend to reduce the perceived sense of agency related to their actions. We propose that the design of robots' behavior in social contexts should consider the impact of the presence of an embodied artificial agent on humans' decision-making.

Acknowledgments

References

1. Murphy, R. R., Nomura, T., Billard, A., and Burke, J. L.. Human-robot interaction. IEEE robotics & automation magazine 17(2), 85-89 (2010) doi: 10.1109/MRA.2010.936953
2. Wiese, E., Metta, G., & Wykowska, A. Robots as intentional agents: Using neuroscientific methods to make robots appear more social. Frontiers in psychology, 8, 1663 (2017)
3. Kompatsiari, K., Tikhanoff, V., Ciardo, F., Metta, G., & Wykowska, A. The importance of mutual gaze in human-robot interaction. In International Conference on Social Robotics 2017 (ICSR2017) pp. 443-452. Springer, Cham (2017).
4. Wykowska, A., Chaminade, T., & Cheng, G. Embodied artificial agents for understanding human social cognition. Phil. Trans. R. Soc. B, 371(1693), 20150375 (2016).
5. Darley, J.M., Latane, B. Bystander intervention in emergencies: diffusion of responsibility. Journal of Personality and Social Psychology, 8(4p1), 377-383 (1968).
6. Chekroun, P., and Brauer, M. The bystander effect and social control behavior: the effect of the presence of others on people's reactions to norm violations. European Journal of Social Psychology, 32(6), 853-867 (2002).
7. Karau, S.J., Williams, K.D. Social loafing: A meta-analytic review and theoretical integration. Journal of Personality and Social Psychology, 65(4), 681–706 (1993).
8. Wallach, M.A., Kogan, N., Bem, D.J. Diffusion of responsibility and level of risk taking in groups. The Journal of Abnormal and Social Psychology, 68(3), 263-274 (1964).

9. Bradley, G.L. Group influences upon preferences for personal protection: A simulation study. Journal of Safety Research, 26(2), 99–105 (1995).

10. Bandura, A., Underwood, B., Fromson, M.E. Disinhibition of aggression through diffusion of responsibility and dehumanization of victims. Journal of Research in Personality, 9(4), 253–269 (1975).

11. Meier, B.P., Hinsz, V.B. A comparison of human aggression committed by groups and individuals: an interindividual– intergroup discontinuity. Journal of Experimental Social Psychology, 40(4), 551–9.

12. Bandura, A. Social cognitive theory of self-regulation. Organizational Behavior and Human Decision Processes, 50(2), 248–87 (1991).

13. Beyer, F., Sidarus, N., Bonicalzi, S., & Haggard, P. Beyond self-serving bias: diffusion of responsibility reduces sense of agency and outcome monitoring. Social cognitive and affective neuroscience, 12(1), 138-145 (2017).

14. Beyer, F., Sidarus, N., Fleming, S., & Haggard, P. Losing Control in Social Situations: How the Presence of Others Affects Neural Processes Related to Sense of Agency. eNeuro, 5(1), ENEURO-0336 (2018).

15. Frith, C.D. Action, agency and responsibility. Neuropsychologia, 55, 137-142 (2014).

16. van der Woerdt, S., & Haselager, P. When robots appear to have a mind: the human perception of machine agency and responsibility. New Ideas in Psychology (2017).

17. Kim, T., & Hinds, P. Who should I blame? Effects of autonomy and transparency on attributions in human-robot interaction. In IEEE International symposium on Robot and Human Interactive Communication 2006 (ROMAN'06), IEEE Publishing, Hatfield New York, NY.

18. Pfister, R., Obhi, S. S., Rieger, M., & Wenke, D. Action and perception in social contexts: intentional binding for social action effects. Frontiers in human neuroscience, 8, 667 (2014).

19. Obhi, S. S., & Hall, P. Sense of agency and intentional binding in joint action. Experimental brain research, 211(3-4), 655-662 (2011).

20. Cozmo SDK installation for Windows. Available: https://cozmosdk.anki.com/docs/install-windows.html

21. Cozmo Quick Tap. Available: https://github.com/anki/cozmo-pythonsdk/blob/master/examples/apps/quick_tap.py

22. Android Debug Bridge. Online available at: cozmosdk.anki.com/docs/adb.html

23. Mathôt, S., Schreij, D., & Theeuwes, J. OpenSesame: An open-source, graphical experiment builder for the social sciences. Behavior research methods, 44(2), 314-324 (2012).

24. Bates, D., Maechler, M., Bolker, B., et al. lme4: Linear mixed-effects models using Eigen and S4 (Version 1.1-7) (2014). Available: http://cran.r-project.org/web/packages/lme4/index. html

25. Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B. lmerTest: Tests in Linear Mixed Effects Models (Version 2.029). (2015). Available: https://cran.r-project.org/web/packages/lmerTest/index.html

26. Efron, B., Tibshirani, R.J. An Introduction to the Bootstrap. CRCPress Boca Raton, Florida, USA (1994).