Magic iCub: a humanoid robot autonomously catching your lies in a card game

Dario Pasquali
RBCS & ICT
Istituto Italiano di
Tecnologia & DIBRIS,
Università di Genova
Genova, Italy
dario.pasquali@iit.it

Jonas Gonzalez-Billandon
RBCS
Istituto Italiano di
Tecnologia & DIBRIS,
Università di Genova
Genova, Italy
jonas.gonzalez@iit.it

Francesco Rea
RBCS
Istituto Italiano di
Tecnologia
Genova, Italy
francesco.rea@iit.it

Giulio Sandini
RBCS
Istituto Italiano di
Tecnologia
Genova, Italy
giulio.sandini@iit.it

Alessandra Sciutti
CONTACT
Istituto Italiano di
Tecnologia
Genova, Italy
alessandra.sciutti@iit.it

ABSTRACT

Games are often used to foster human partners' engagement and natural behavior, even when they are played with or against robots. Therefore, beyond their entertainment value, games represent ideal interaction paradigms where to investigate natural human-robot interaction and to foster robots' diffusion in the society. However, most of the state-of-the-art games involving robots, are driven with a Wizard of Oz approach. To address this limitation, we present an end-to-end (E2E) architecture to enable the iCub robotic platform to autonomously lead an entertaining magic card trick with human partners. We demonstrate that with this architecture a robot is capable of autonomously directing the game from beginning to end. In particular, the robot could detect in real-time when the players lied in the description of one card in their hands (the *secret card)*. In a validation experiment the robot achieved an accuracy of 88.2% (against a chance level of 16.6%) in detecting the *secret card* while the social interaction naturally unfolded. The results demonstrate the feasibility of our approach and its effectiveness in entertaining the players and maintaining their engagement. Additionally, we provide evidence on the possibility to detect important measures of the human partner`s inner state such as cognitive load related to lie creation with pupillometry in a short and ecological game-like interaction with a robot.

CCS CONCEPTS

• Human-centered Computing~Interaction design process and methods

KEYWORDS

Entertainment, magic, human-robot interaction, pupillometry, cognitive load

1 Introduction

Historically, robots always fascinated the public, entertaining the audience. Indeed, the first recorded example of humanoid robot was a robotic musical band meant to entertain the guests of an Arabian king [1]. Nowadays, robots can have a role not only in task-oriented research or industrial applications, but also in the field of entertainment. The concept of Entertainment Robotics refers to any robotic platform and application not directly useful for a specific task, but rather meant to entertain and amuse humans. Recently several entertainment robotic platforms [2]–[9], frameworks [10]–[13] and applications have been developed. Amusement and theme parks are one of the main application fields. Here, robots are meant to be observed, providing an entertaining show without any interaction. For instance the Disney World Company employs robots to act on stages [14], to perform acrobatic actions [15], [16], or to freely roam in the theme parks [17]. The latter can perform a finite set of human-robot interactions meant to handle approaching crowds. Rather than just being watched, a few robots socially engage the users. For instance, Sophia [18], [19] and Geminoid robots [20]–[22] can handle a dialogue with a human partner. However, despite the complexity of the interaction, in most cases everything is scripted and relies on a Wizard of Oz control configuration. Other robots interact physically with the human partner; for example, they play ping-pong [23], soccer [24], [25], table hockey [26] and ball catching [8]. Robot companions, like PARO [27], [28], AIBO [3], [4], [29] or Keepon [30], are a special branch of entertainment robot platforms, usually employed in education [3], [6], [31], [32]

and therapy [9], [33]. Such robots, usually resemble the appearance of animals or cute creatures, providing a limited set of predefined animations and reactive interactions. Recently, researchers started exploiting games and entertaining tasks as ecological and realistic scenarios to investigate human-robot interaction (HRI). Competitions like the IEEE Human application Challenge: Robot Magic and Music [34]–[37] and the IEEE RoboCup [24], [38], pushed researchers' interest toward entertaining experiments and applications. For instance: Ahmadi et al. [39] and Ahn et al. [40] played rock-paper-scissor trying to predict playmate's gestures; Michalowski et al. [41] studied how rhythm affects attention and intent in a dance game; Gori et al. [42] made the iCub robotic platform playing mime with a human partner; Leite et al. [43] studied the effect of non-verbal communication on user engagement with storytelling robots (see also [44], [45]); Aroyo et al. [46] studied the compliance of players to iCub's hints in a treasure hunt; Palinko et al. [47] studied mutual gaze with Androids in a gaze-based social game.

Entertainment applications demonstrated to be an effective way to introduce naïve users to robots and foster robots' diffusion in the society. However, most of the presented robotic platforms and applications lack on autonomy, which limits their diffusion beyond specific contexts: robots depend on a Wizard of Oz [48] control configuration – and an expert handler – or follow a predefined script. To overcome these limitations, a robot should show autonomy, sensing its human partner and taking decisions accordingly, at least within the framework of a closed-world scenario as a game.

Inspired by the television show *Box of Lies* [49], we explored whether the iCub humanoid robot could lead an entertaining magic trick in an autonomous way by detecting its human partner's lies. In the game, iCub has to detect in real-time the player's *secret card* – the card about which the human is lying – from a set of six random cards, during a quick and ecological social interaction. The approach is inspired by previous findings on lie detection in HRI [50], [51] based on cognitive load assessment [52], [53] via pupillometric features [54], [55]. We propose an autonomous end-to-end (E2E) architecture which integrates the cognitive load assessment, the decision making and the robot control, enabling iCub to lead the magic trick with no need of a Wizard of Oz control configuration. Based on our system, iCub successfully detected players' *secret cards* with an accuracy of 88.2% (N=34, against a chance level of 16.6%). We further report post hoc analysis of the participants' strategies and pupillometry features and discuss whether the approach adopted could be improved and could effectively detect the cognitive load associated to lie creation in a short and game-like interaction.

2 Magic Trick Interaction Design

We designed a human-robot interaction where the iCub robotic platform plays a magic trick with a human player. The players describe six cards in front of iCub and have to lie about one of them (the *secret card*). iCub autonomously detects which is the

fake description among the six. During the game, the players sit in front of iCub with a table (covered with a black cloth) between them. On the table, lie six green rectangular marks, a deck of 84 gaming cards with blue back, a keyboard and a Tobii Pro Glasses 2 eye-tracker [56] (Figure 1).

Figure 1: Participant describing a card to iCub from Logitech Brio 4k webcam point of view.

As the game starts, iCub asks the players to shuffle the deck, draw out six random cards without looking at them and put the deck aside. Then, iCub instructs the player to shuffle the six cards again, draw out one of them (iCub calls it *secret card*), memorize it and put it back on the table. Then, iCub asks to look at all the cards, one by one, shuffle them and put them covered on the six green marks. iCub says that, to perform a magic trick, it is going to point each card one by one and instructs the player to take the pointed card, describe it and then put it back on the green mark. It says: "*The trick is this: if the card you take is your secret card, you should describe it in a deceitful and creative way. Otherwise, describe just what you see*". Finally, iCub asked the player to wear the Tobii Pro Glasses 2 eye-tracker, take a deep breath and relax. After that, iCub starts pointing the cards one by one, providing a verbal feedback at the end of each description. After the last description, iCub guesses the *secret card* and asks the player a confirmation: the player has to remove all the cards from the table to confirm the guess or show the *secret card* to reject it.

3 Computational Architecture

iCub could autonomously lead the whole magic trick thanks to the E2E architecture in Figure 2.

3.1 Secret Card Detector

During the magic trick iCub guesses players' *secret card* by detecting the Task Evoked Pupillary Responses (TEPRs) [54], [57] related to lying [53]. The variation of cognitive load [58], [59] during a task has been proved to reflect on pupillometry features [60], in particular on pupil dilation. The fabrication and maintenance of a credible and consistent fake card description triggers a cognitive load peak in players' mind [52], [61], [62]

which reflects on their pupils. During the magic trick, iCub measures in real-time the player's pupil dilations through the Tobii Pro Glasses 2 eye-tracker. The Secret Card Detector implements the algorithm that allows iCub to detect the *secret card* among the six cards, based on an heuristic approach [50]. At the end of the game, iCub selects as *secret card* the one related to the highest mean pupil dilation among the six. More precisely, for each card, it computes an average value from the moment the players take the card from the marker to the moment they put it back (we refer to these intervals as *player's turns*). Before evaluating which is the *secret card*, each pupil dilation datapoint is normalized with respect to the average pupil dilation during the 5 seconds before the first pointing (when iCub asks the player to take a deep breath and relax) [50], [63].
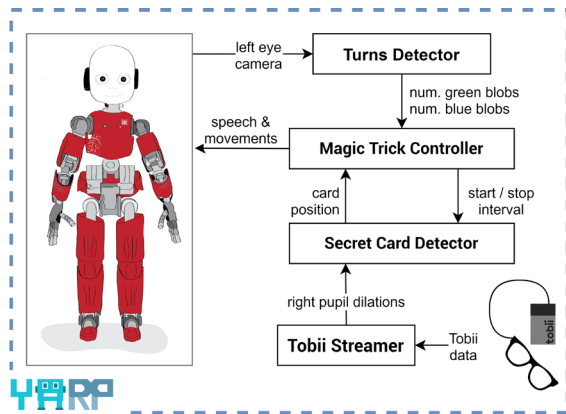

Figure 2: end-to-end computational architecture to play the Magic Trick with the iCub humanoid robot

3.2 Tobii Streamer

The Tobii Pro Glasses 2 is a wearable eye-tracker meant to collect pupillometric features and post-hoc analyze them [56]. We developed the Tobii Streamer (extending the Tobii Glasses Py Controller python module [64]) in order to stream the right eye pupil dilation in real-time over a YARP robotic platform [65]. Even if the magic trick is based only on the right eye pupil dilation, the features for both eyes are logged on YARP for further analysis. We decided to focus on right-eye features since prior findings on lie detection based on pupillometric features [51] and Tobii documentation [66] reported no significant difference between the two eyes. We also decided to skip the Tobii Pro Glasses 2 eye-tracker calibration to not impact the informality of the interaction. Tobii documentation reports how the calibration is only relevant for the gaze features and does not impact the pupil dilation measurement [66]. Finally, the system uses the Tobii REST APIs to record the full set of pupillometric features, exposed by the proprietary software, for future analysis.

3.3 Turns Detector

The Turns Detector allows iCub to autonomously handle the turn-taking during the game. Indeed, iCub needs to know when the players take a card to start the pupil dilation aggregations and

when they put it back on the table to store the collected data and present the new item. The Turns Detector implements a simple HSV color thresholding that detects the number of blue (cards) and green (marks) rectangular blobs in the scene. During the game, iCub double checks the number of visible marks and cards blobs to robustly understand the game phases. For instance, it detects when the players put the cards on the table for the first time by detecting zero green marks; this triggers the rules explanation. When the players take a card to describe it, it tracks five blue and one green blobs until the player put it back on the table.

3.4 Magic Trick Controller

The Magic Trick Controller handles iCub speech and movements and coordinates the other components. iCub's pointings were performed in a human-like manner: first gazing to the card and then pointing, by moving both the arm and the body. In order to increase players' engagement and provide a more social interaction, iCub acknowledges the end of each description with a simple feedback sentence (e.g., *"ok", "mh mh", "I see"*). The Magic Trick Controller autonomously commands the Secret Card Detector to segment the pupil dilation timeseries, based on the card tracking of the Turns Detector. At the end of the game, it autonomously handles the validation of the detected *secret card* description. Moreover, it annotates, through YARP, the timestamps related to the beginning and end of each pointing and description, along with the position of the *secret card* for further post-hoc analysis.

4 Methodology

To validate our computational architecture, we tested it in real interactions with several participants. The main objective is to demonstrate the effectiveness of our proposed architecture to make iCub autonomously lead an entertaining and effective human-robot interaction, based on the real-time reading of a biometrical feature from the players.

4.1 Participants

We asked 39 participants to play the magic trick with iCub. They were 14 males and 25 females with an average age of 28 years (SD=8) and a broad educational background; they received a monetary compensation of 10 € to participate in the experiment. Participants signed an informed consent form approved by the ethical committee of the Regione Liguria (Italy) where it was stated that cameras and microphones could record their performance and agreed on the use of their data for scientific purposes. Even if all participants completed the experiment, we discarded 5 interactions from the analysis (see Sec. 5.1.1), leading to a sample of N=34 participants (12 males, 22 females).

4.2 Setup

For the experiment, the experimental room was arranged to replicate an informal interaction scenario between a human and a robot (Figure 3). The participant sat on a chair in front of iCub.

Between the participant and the robot, we set a table covered with a back cloth. On the table lied: a deck of 84 playing cards with blue back, six green rectangular marks, a keyboard, and a Tobii Pro Glasses 2 eye-tracker. On participant's left, there was a little drawer while, on the right, a black curtain hid the experimenter from participant's sight. Behind iCub, a 47 inches television showed iCub speech during the interaction (to avoid any misunderstanding of the robot's speech). The Tobii Pro Glasses 2 streamed and recorded pupil dilations with a frequency of 100 Hz. A Logitech Brio 4k webcam [67], placed on the television, recorded the participant during the whole interaction (Figure 1). The windows blinders were closed, and the room was lit with artificial light to ensure a stable light condition for all the participants during different times of the day.

4.3 Materials

A Dixit Journey gaming cards deck has been modified by coloring the back of each card in blue. These 80x120 mm cards present 84 different toon-styled drawings meant to stimulate creativity and creative thinking [68] (Figure 3, Right). Six green 95x70 mm marks with a white border have been glued to the black cloth. The iCub humanoid robot [69] played the role of magician. The experimenter, hidden behind the black curtain, monitored the scene through iCub's eyes to ensure the safety of the players.

4.4 Procedure

At least one day before the experiment, the participants filled in the Big Five personality traits questionnaire (extroversion, agreeableness, conscientiousness, neuroticism, openness) [70], the Brief Histrionic Personality Disorder (BHPD) [71] questionnaire and the Short Dark Triad (SD3, machiavellianism, narcissism, and psychopathy) [72], meant to assess their personality.

After signing the informed consent, the experimenter led the participants in the experimental room. The experimenter asked the participants to sit on the chair in front of iCub, stated that iCub was going to explain everything and closed the black curtain, hiding himself from players' sight. iCub led the experiment as described in Sec. 2. During the initial rule explanation, iCub instructed the participant to press a key on the keyboard to move to the next task (i.e., after shuffling the cards deck or after memorizing the *secret card*). No time limit was given to memorize the *secret card*, neither to describe the cards. After the magic trick, the participants performed a second card game with iCub lasting on average 8 minutes (SD=2).

At the end of the game, the experimenter led the participants in the initial room and asked them to fill in a post-questionnaire. The questionnaire includes the NASA-TLX [73] and a set of questions meant to evaluate players' experience during the game: (i) experienced fun (5-points Likert scale), (ii) effort on fabricating a deceitful and creative *secret card* description (5-point Likert scale), (iii) deceptive strategy adopted (open question; e.g., premeditating the card description while iCub was explaining the rules; or being vague); and (iv) perceived strategy adopted by iCub in the

detection (open question). Also, we asked whether players had previous experience on improvising and acting and if they knew the Dixit card game. Finally, the participants were deeply debriefed, and they had time to ask questions about the experiment before receiving the monetary compensation.

Figure 3: (Left) Experimental room setup with iCub, the participant, six green marks, six blue cards, a Tobii Pro Glasses 2 and a keyboard on the table; (Right) Dixit Journey gaming cards (author: Jean-Louis Roubira, designer: Xavier Collette, publisher: Libellud). Top right card described as *"a blue shark riding a bike"*

4.5 Measurements

4.5.1 Card Segmentation. The card segmentation is autonomously performed by the architecture: the pupil dilations and the beginning and end events of each pointing and card description are logged on YARP which ensures the synchronization of the timestamps. For each card, we identified 3 temporal intervals: (i) *robot's turn*: from the moment iCub starts the pointing to the moment the players take the card from the table; (ii) *player's turn*: from the moment the players take the card to the moment they put it back on the table; (iii) *card trial*: the combination of both previous turns: from the beginning of iCub's pointing to the moment the players put the card back on the marker.

4.5.2 Gaze Measurements. Participants' pupillometry features were recorded using the Tobii Pro Glasses 2 eye-tracker at a frequency of 100 Hz. Recorded features include right and left pupil dilation in mm, gaze point 2D, gaze point 3D and fixation and saccades events. Since we decided to not perform the Tobii Pro Glasses 2 calibration, only pupil dilation features are reliable. For each of the 3 temporal intervals, (robot turn, player turn and card trial) we computed 5 features: duration, maximum, minimum, mean and standard deviation of pupil dilation leading to a final feature set composed of 15 features.

4.5.3 Data Preparation. To post-hoc analyze the collected data (see Section 5.4), we preprocessed the pupil dilation features. We applied a low pass filter at 10 Hz, a median filter, and a rolling window filter to clean the pupil dilation time series. Before segmenting the intervals, we corrected each time series subtracting a baseline average value for each participant [63]. We computed the baseline, by averaging the pupil dilation, for each

eye separately, during the five seconds before the first pointing – when iCub asks the player to take a deep breath and relax. In this reference system, a positive value represents a dilation, while a negative value represents a contraction with respect to the baseline.

5 Results

5.1 In-game Analysis

The Magic Tricks lasted 8 minutes (SD=2) on average, from when iCub started explaining the rules to the final confirmation of the detection. ICub successfully detected players' *secret card* with an accuracy of 88.2% (against a chance level of 16.6% and considering the N=34 interactions not affected by technical issues or rule misunderstanding; see below).

5.1.1 Discarded Interactions. Although all participants completed the game, we had to exclude 5 of them from further analysis. Two of them failed to follow the rules of the game: one misunderstood the instructions and fabricated a deceitful and creative description for all the cards; one misunderstood iCub's pointing gesture and ended the game without describing the *secret card.* Another participant took very long to describe each card concluding the game after 26 minutes (vs. an average of 8 min for all other participants). For the last two participants we had technical issues: for one, a problem with the blinders did not allow to maintain a stable light condition during the game; for the other, even if the secret card detection was successful, we had a problem with the storage server that prevented the data saving.

5.1.2 Detection Failures. Considering the 4 participants (out of 34) in which iCub failed to detect the *secret card*, we had two particularly interesting cases. One participant produced an incomplete description for the first card because the experimenter interrupted it by mistake. Looking at the pupil dilation timeseries of that player, it experienced a cognitive load peak probably due to the novelty of the game. We speculate that the card description was interrupted too early to allow a mitigation of such cognitive load (and hence the pupil dilation), resulting in a higher mean pupil dilation; indeed, iCub detected that card as the *secret card.* For the second participant we noticed a pupillary pattern opposite with respect to the others: the *secret card* was the one related to the lowest mean pupil dilation among the six. Regarding the other two: one reported, during the debriefing, to be used to creative thinking; the second one described the card vaguely and by omitting details rather than creating a novel one. Both failures could be explained by the need for a lower cognitive effort to fabricate a creative description because of the adopted strategies.

5.1.3 Experimenter's Interventions. In general, the game unfolded properly, and we encountered a few issues that required human intervention. More precisely, considering all the interactions, the experimenter had to intervene 3 times verbally mainly to remind the player to put the deck aside to prevent interferences with the cards and marks tracking. Additionally, some technical issues occurred: some major (N=5) where the experimenter had to stop and restart the game and two minors where the experimenter needed to intervene (i.e., asking to move the cards deck). The major issues were related to either the malfunction of the Tobii Pro Glasses 2 that prevented the streaming of pupil dilations (N=4) or to robot malfunction that needed the restart of the robotic platform (N=1). After restarting the devices, the game went flawlessly for those participants. Finally, for 2 participants one of the card description was erroneously interrupted. In one case, the Turns Detector failed to track the cards due to a misplacement over the marks; in the other case the interruption was due to a human error, as mentioned above, that did not hinder the completion of the game.

5.2 Questionnaire Analysis

With the questionnaire analysis, we mainly wanted to understand: (i) how much the game was able to entertain the players; (ii) if a bad performance during the game (due to misdetections and/or game failures) had an impact on players' fun; (iii) how much effort was required to play the game.

5.2.1 Experienced Fun. Considering the whole sample (N=39), participants reported a high average fun of M=4.4 (SD=0.82). We then compared the fun for those for which iCub failed to detect the *secret card* (N=8, M=3.75, SD=1.28) and for the others (N=31, M=4.63, SD=0.56). A Wilcoxon rank-sum test showed no significant statistical differences between the two samples (Z=1.74, p=0.082). Moreover, we supposed that the presence of failures during the interaction could impact the experienced fun. We compared the reported fun of the games which proceeded without any (even minor) technical issues and the iCub successfully guessed the *secret card* (N=26, M=4.68, SD=0.56), against the others (N=13, M=4.0, SD=1.08). The Wilcoxon rank-sum test revealed no significant statistical difference, although there was a trend to find more entertaining the flawless games (Z=1.9, p=0.056).

5.2.2 Creative Effort and Task Load. On average, participants reported a creative effort of M=3.6 (SD=0.97), considering only the individuals who followed the game rules and for which there was no severe technical issue or outlier behavior (N=34). The participants for which iCub failed the *secret card* detection reported an average creative effort of M=3.0 (SD=1.41, N=4), while the others reported an average creative effort of M=3.87 (SD=0.73, N=30), with no significant difference between the two groups (Wilcoxon rank-sum test, Z=1.12, p=0.26). Considering task load in general, the Task Load indeX (TLX), computed from the NASA-TLX questionnaire was not high on average. Participants reported an average TLX of M=3.7 (out of 10, SD=1.03).

A Wilcoxon rank-sum tests showed that both Fun (Z=378.0, p<0.001) and creative effort (Z=341.5, p<0.001) are significantly higher than the "neutral" median value (3). Also, we found that the higher was the reported effort in creating a lie, the higher was the experienced fun (Spearman correlation: rs(28) = 0.53, p<0.001). Considering the relation with the personality traits of the

participants from the pre-questionnaire, the creative effort was linearly (negatively) correlated with the openness to experience (t(28)=-3.96, p<0.001, Adj. R^2=0.62), and (positively) correlated with the conscientiousness (t(28)=5.99, p<0.001, Adj. R^2=0.62), whereas no other element of the Big 5 showed a significant correlation with it. Regarding the Dark Triad, we found a positive linear correlation between the machiavellianism (t(28)=3.49, p=0.0019, Adj. R^2=0.271) and the mental effort component of the NASA-TLX. We found no effect from the histrionic questionnaire.

5.2.3 Deceptive Strategies. The players exploited a variety of strategies to fabricate the creative and deceitful description for the *secret card.* We manually translated the qualitative reports of the participants, integrated with experimenter's notes during the experiment in a finite set of strategies with intersection. The question was not mandatory, hence just 24 participants reported a qualitative strategy. Most of the players (N=8) reported the usage of memory recall, related to a previous card or a past event; 3 players swapped the roles of the characters in the cards and just 3 players reported the creation of a brand-new image; 3 participants focused on adding details while 3 tried to be vague and generic about the description; finally, 3 participants focused on the credibility and consistency. We also identified two classes related to the timing of fabrication of the creative description: 8 participants reported they premeditated the description as iCub presented game rules; other 8 participants instead, improvised the description on the fly. We did not find any statistical difference between the samples on predicting fun or creative effort.

We applied a similar preprocessing on the perceived methods used by iCub to detect the *secret cards.* Although the eye-tracker was the only evident sensing device in the interaction, 27 of the 39 players (69%) did not mention gaze or pupil when describing the strategy used by the robot to guess the *secret card.* 8 participants assumed iCub was able to detect a variation on the description, including both prosodic features and number of details; 3 participants assumed iCub detected the presence or absence of keywords in their descriptions; only a participant thought about facial and postural features. Interestingly, 6 participants assumed iCub knew all the 84 cards and hence it could understand the card description and match it (or not) with one of the cards. Few of them (N=3) also assumed iCub could see the card from its reflection on the glasses and pair the image with the description.

Finally, as a qualitative report, all the participants were surprised when the experimenter presented iCub and stated that it was going to lead the experiment. At the end of the experiment, they all reported they had fun, even the ones that experienced failures. They were also extremely surprised to learn the effect of cognitive load on pupil dilation.

5.3 Post-hoc Analysis

We analyzed the collected pupillometric features to provide statistical support to the results of the validation experiment and assess whether the heuristic method can be further improved. A Saphiro-Wilk [74] and D'Agostino K-squared [75] normality tests showed that the data were normally distributed, justifying the use of a parametric analysis.

5.3.1 Robot and Player turns comparison. First, we ran a paired t-test comparing the average of mean pupil dilation for right and left eyes. Results showed no significant difference (t(33)=1.58, p=0.123), hence we focused on the right players' eye as in the real-time Magic Trick. We compared the mean pupil dilation for the *secret card* against the average of the other cards during the different turns of the game. We performed a two-way repeated measures ANOVA on players' mean pupil dilations with factor "card label" (two levels: Real, Fake) and factor "turn" (two levels: Robot, Player). The test shows a highly significant difference in players' pupil dilation as a function of the card label (F(1, 33)=44.17, p<0.001, η_P^2=0.57), no significance of the turn factor (F(1,33)=2.69, p=0.11, η_P^2=0.08), but a highly significant interaction (F(1,33)=58.01, p<0.001, η_P^2=0.64). Hence, mean pupil dilation is overall different between real and untruth card descriptions, but this difference is significantly larger in the player turn, i.e., while the description was performed. More precisely, post-hoc analysis (Bonferroni corrected) showed that the mean pupil dilation for the *secret card* description was significantly higher than the mean pupil dilation for the average of the other cards during the *player's turn* (t(33)=9.87, p<0.001) but not in the *robot's turn* (t(33)=0.16, p=0.33). The effect is also visible in Figure 4. For the *player's turn,* we also analyzed whether other features (maximum, minimum and standard deviation of pupil dilation) differed significantly between the *secret card* and the others. Paired t-test tests showed that both minimum pupil dilation (t(33)=7.18, p<001) and maximum pupil dilation (t(33)=7.87, p<0.001) were significantly higher during the false description than during the truthful ones.
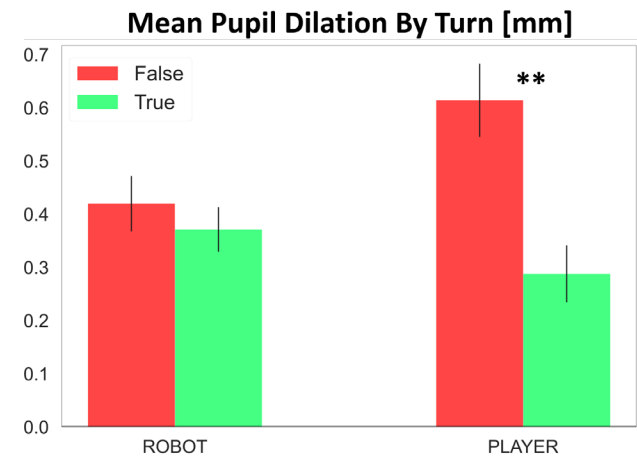
Figure 4: Average right mean pupil dilation for the *secret card* (red) and averaged other cards (green) during the magic trick *turns.* Error bars represent standard errors of the mean. Stars represent statistical difference (p<0.001)**

5.3.2 Card trials analysis. As an exploratory analysis, we investigated whether it is possible to further simplify the

interaction by removing the turn segmentation. Figure 5 represents the right mean pupil dilation during the whole card trial for *secret card* and average of the other cards for each participant. Except for two participants, all the others lie above the identity line, showing larger mean pupil dilation on the s*ecret card*. We ran a paired t-test comparing the mean pupil dilation on the *secret card* with the average of the others during the whole card trials. The abovementioned affect is still present since the mean pupil dilation (t(33)=9.14, p<0.001), maximum pupil dilation (t(33)=6.91, p<0.001) and minimum pupil dilation (t(33)=6.37, p<0.001) are significantly higher during *secret cards* descriptions. If the heuristic to detect the *secret card* had been based on the whole card trial interval, the robot would have guessed the right card with an accuracy of 85.3% (against a chance level of 16.6%). This simulation result proves that it is possible in the future to further simplify the interaction and the Secret Card Detector, by analyzing online the whole interval from the instantiation of one pointing to the beginning of the next, without the need of segmenting exactly the time in which the participant takes the card from the table.
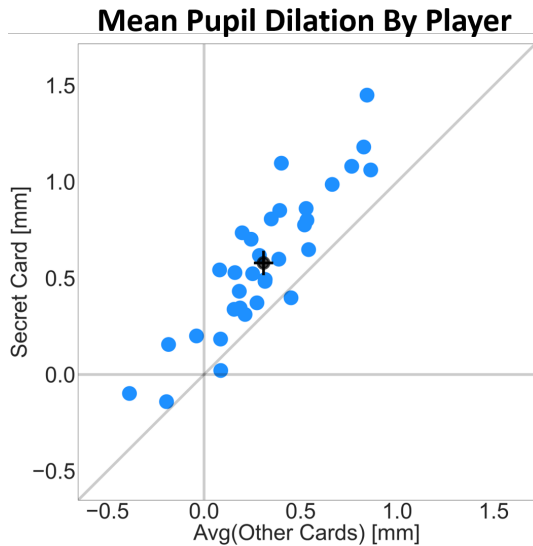
Figure 5: Right mean pupil dilation during the whole card interval for *secret card* and average of the other cards for each participant. The black dot represents the sample mean with standard error.

5.3.3. A more robust lie detector. The heuristic function enables iCub to autonomously lead the proposed game; however, it is still affected by two limitations: (i) it is unreliable in case of light changes during the game; and (ii) it does not consider unexpected behaviors from the players (e.g., lying on multiple cards). To address these limitations, in the post-hoc analysis we corrected each pupil dilation datapoint by subtracting the average pupil dilation during the five seconds before each *card trial*. This kind of baseline should compensate for potential fluctuations of both environmental light and players' cognitive load during the game. Then, we trained a machine learning model able to classify a

generic description as true or false, independently from the number of items or lies. Assuming lying is a rare behavior with respect to a normal truth telling, we analyzed the problem as an anomaly detection; this technique also avoided us to oversample the dataset to tackle its unbalancing. We considered the whole feature set and included data from both right and left players' eyes, discriminated by a proper categorical feature. We trained a one-class support vector machine (OCSVM) [76] on the resulting dataset (405 datapoints x 16 features). OCSVMs are semi-supervised models meant to train only on normal data (true card descriptions), learning to discriminate what is abnormal (false card descriptions). We considered 75% of the true card descriptions as train set and the remaining (both true and false) as validation and test sets. A grid-search cross validation shows that the best model has an AUCROC of 0.61, an F1 score of 79.6%, a precision of 77.4% and a recall of 81.8%.

6 Discussion

In this study we show how a humanoid robot can successfully guide a prolonged and entertaining activity with a human partner based on a real-time measurement of players' pupil dilation. Our innovative approach shows how the autonomous end-to-end (E2E) architecture successfully promotes an enjoyable activity with a robot. At the same time the architecture allows for the extraction of important information about the inner state of the human partner (i.e., cognitive load related to lying). Players' lies can be recognized with a good accuracy level of 88.2% (N=34, against a chance level of 16.6%) during a short interaction (8 min) without leveraging on a priori knowledge of individual attitudes. The measures of fun and task load, reported after the game, confirm how the magic trick is entertaining, even if iCub failed to detect the *secret card* or malfunctions happened during the interaction. Also, the reported creative effort and task load suggest how the human-robot interaction does not require any significant effort to be played.

The current architecture implementation and setup still presents two main issues, as the employed pupillometry measure is sensitive to illumination changes during the interaction and the approach is not robust against unexpected behaviors from the players (e.g., multiple lies). Considering the sensitivity to illumination, it mostly represents a limitation for outdoor environments. Since our solution does not require a specific illumination, but rather a constant one, this requirement can be easily met in most indoor contexts. For what concerns the dependency of the system on a fixed number of lies or items, the different preprocessing and the one-class support vector machine tested in the post-hoc analysis show promising expectations that also these limitations could be overcome. However, further research must be performed to improve the reliability of the system.

Although the validation experiment was conducted with the humanoid robot iCub, the architecture is highly modular and portable. The relatively limited sensing and acting abilities needed

along with the decomposition between sensing and robot control make the architecture easily adaptable to different robotic platforms. The pointing actions could be replaced by different ways to show the cards, and the detection of the robot and player turns could be performed with ad hoc sensing. The architecture is also extremely light weighted: it does not require excessive computational power or a network connection. This makes it easily deployable directly on other robotic platforms' boards. We did not explore the effect of robot appearance on game entertainment; however, we speculate that the childish appearance of the iCub humanoid robot contributed to engage the players, making the game more entertaining. Further research must be performed to address the impact of robot appearance on the proposed game.

The interaction, and hence the autonomous architecture, could be further simplified. We demonstrated with a post-hoc analysis that even considering the whole interval of time in which a single card is shown and described, the heuristic would work well (accuracy: 85.3% against a chance level of 16.6%). Hence, the Turn Detector could be simplified by detecting just the end of the description to know when to present the next stimulus. Indeed, the Turn Detector implementation, based on the HSV color thresholding of cards and marks, is a limitation of the current architecture. It depends on light conditions and camera calibration and it is prone to potential false positives due to other colored objects in the scene. We decided to implement such simple approach thinking about the potential deployment of our entertainment architecture in other fields. For instance, amusement and theme parks are crowded and loud places, hence it would not be feasible to use speech-based algorithms (i.e., a voice activity detector algorithm) to detect players' descriptions. We also decided to avoid any computer-readable marks (i.e., QR codes) to avoid that the player would assume that iCub could recognize the cards by their backs. Thinking about a future deployment of the architecture, it will be mandatory to improve our card tracking method. For instance, we could track players' gestures or the original Dixit gaming card back with a feature-based object localization algorithm. This way, it would also be possible to remove the green marks on the table, further simplifying the setup of the game.

The elimination of the green marks would reduce the required materials to just the eye-tracker. Even if the interaction unfolded naturally, we recognize that the use of a head mounted eye-tracker, though lightweight, reduced the naturalness of the task. We partially reduce its impact on the informality of the interaction by removing the calibration phase, since it is not strictly required to measure pupil dilation. Moreover, 27 of the 39 participants did not mention the eye-tracker (or any eye-related feature) as the method used by iCub to detect their *secret card*. Hence, we speculate that the eye-tracker did not compromise players' fun during the game, nor induced them to be self-aware of their own gaze behavior. However, to port the application to a real-world scenario, the ideal solution would be measuring the player's pupil dilation from the RGB cameras embedded on the robotic platform. Recent research developments have shown the feasibility of using RGB cameras to assess pupillometric features [77]. Hence, we believe that in the future it will be possible to also remove the eye-tracker requirement.

Beside the applications in a real-world entertaining scenario (i.e., amusement parks), the system could represent a natural way to introduce robots in the society by allowing naïve users to experience a quick, pleasant, and interactive game with a real robot. Additionally, this system could become a novel tool to measure pupillometric modulations associated to creativity in a pleasant and non-invasive way (e.g., appropriate for children). Also, this work demonstrates that a robot can effectively monitor the variations in cognitive load during a natural interaction. The generality of cognitive load detection is supported by the high variability of the items employed (84 different cards). Hence, the measure should not be limited to a specific set of items. This is novel with respect to the state-of-the-art cognitive load assessment methods based on long, tedious and strictly constrained tasks [51], [78], [79] and cumbersome sensing devices. Hence, it represents a step toward those applications where robots could take benefit from evaluating the human partner's internal state and change their behavior accordingly (e.g., by providing a less challenging task). Moreover, this evaluation is performed preserving the informality of the human-robot interaction, an important factor in fields like teaching or caretaking.

In the future we plan to improve the architecture as both (i) an entertaining and autonomous game with a humanoid robot and (ii) and an effective and quick method to assess human partners' cognitive load in real-time. We aim to adapt iCub's behavior based on the measured cognitive load.

7 Conclusion

Thanks to the autonomous architecture proposed in the manuscript we provide evidence that robots can, at the same time, (i) autonomously guide a human-robot interaction in and ecological magic trick (detecting players' *secret cards* with an accuracy of 88.2%, against a chance level of 16.6%) and (ii) promote, through the proactive interaction, the online acquisition of important insights on the human counterpart's inner state. The future implications of such approach are activities that are beneficial or entertaining for the human partners and, at the same time, allow the robots to adapt their behavior to the specific inner state of the participant in real-time. This will be a key factor for robots that aim to act in fields related to tutoring, caregiving, and security. Finally, we hope that the development of more accessible and portable entertaining applications could foster the diffusion of robots in the world as enjoyable playmates, thus paving the way toward their acceptance in the society.

REFERENCES

[1] "The Musical Boat for a Drinking Party – The Book of Knowledge of Ingenious Mechanical Devices." [Online]. Available: https://aljazaribook.com/en/2019/08/07/the-musical-boat-en/. [Accessed: 01-Oct-2020].

[2] H. Kitano, "Development of an Autonomous Quadruped Robot," *Science (80-.).*, vol. 18, pp. 7–18, 1998.

[3] K. Sabe, "Development of entertainment robot and its future," *IEEE Symp. VLSI Circuits, Dig. Tech. Pap.*, vol. 2005, pp. 2–5, 2005.

[4] M. Fujita, "Digital creatures for future entertainment robotics," *Proc. - IEEE Int. Conf. Robot. Autom.*, vol. 1, no. April, pp. 801–806, 2000.

[5] M. Fujita, Y. Kuroki, T. Ishida, and T. T. Doi, "A small humanoid robot SDR-4X for entertainment applications," *IEEE/ASME Int. Conf. Adv. Intell. Mechatronics, AIM*, vol. 2, no. Aim, pp. 938–943, 2003.

[6] A. Billard, "Robota: Clever toy and educational tool," *Rob. Auton. Syst.*, vol. 42, no. 3–4, pp. 259–269, 2003.

[7] C. Y. Lin, C. K. Tseng, and P. C. Jo, "A multi-functional entertaining and educational robot," *J. Intell. Robot. Syst. Theory Appl.*, vol. 53, no. 4, pp. 299–330, 2008.

[8] T. H. and U. F. Tim Laue1, Oliver Birbach1, "An Entertainment Robot for Playing Interactive Ball Games," *Lect. Notes Comput. Sci.*, no. January 2014, pp. 592–599, 2014.

[9] H. Kozima, M. P. Michalowski, and C. Nakagawa, "Keepon: A playful robot for research, therapy, and entertainment," *Int. J. Soc. Robot.*, vol. 1, no. 1, pp. 3–18, 2009.

[10] M. Fujita and K. Kageyama, "Open architecture for robot entertainment," *Proc. Int. Conf. Auton. Agents*, no. August, pp. 435–442, 1997.

[11] B. Robins, K. Dautenhahn, R. Te Boekhorst, and C. L. Nehaniv, "Behaviour delay and robot expressiveness in child-robot interactions: A user study on interaction kinesics," *HRI 2008 - Proc. 3rd ACM/IEEE Int. Conf. Human-Robot Interact. Living with Robot.*, no. January, pp. 17–24, 2008.

[12] X. Huang, "A virtual entertainment robot based on harmonic of emotion and intelligence," *Chinese J. Electron.*, vol. 19, no. 4, pp. 667–670, 2010.

[13] Y. C. Kim, H. T. Kwon, W. C. Yoon, and J. C. Kim, "Designing emotional and interactive behaviors for an entertainment robot," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2009, vol. 5611 LNCS, no. PART 2, pp. 321–330.

[14] C. Causer, "innovation," 2019.

[15] "Disney's high-flying acrobatic robots will floor you." [Online]. Available: https://edition.cnn.com/videos/cnnmoney/2018/07/04/disney-robots-stuntronics-animatronics-orig.cnnmoney. [Accessed: 27-Sep-2020].

[16] K. Bassett, M. Hammond, and L. Smoot, "A fluid-suspension, electromagnetically driven eye with video capability for animatronic applications," *ACM SIGGRAPH 2010 Emerg. Technol. SIGGRAPH '10*, pp. 40–46, 2010.

[17] "Disney World is About To Be Invaded By Robots...And You're Going To Love It! | the disney food blog." [Online]. Available: https://www.disneyfoodblog.com/2020/04/10/disney-world-is-about-to-be-invaded-by-robots-and-youre-going-to-love-it/. [Accessed: 27-Sep-2020].

[18] H. Kaur Kalra and R. Chadha, "A Review Study on Humanoid Robot SOPHIA based on Artificial Intelligence," *Int. J. Technol. Comput.*, vol. 4, no. 3, pp. 31–33, 2017.

[19] "Sophia 2020 - Hanson Robotics." [Online]. Available: https://www.hansonrobotics.com/sophia-2020/. [Accessed: 05-Oct-2020].

[20] "HIL." [Online]. Available: http://www.geminoid.jp/en/index.html. [Accessed: 05-Oct-2020].

[21] A. M. Aroyo *et al.*, "Will People Morally Crack under the Authority of a Famous Wicked Robot?," in *RO-MAN 2018 - 27th IEEE International Symposium on Robot and Human Interactive Communication*, 2018, pp. 35–42.

[22] H. Ishiguro and F. Dalla Libera, *Geminoid studies: science and technologies for humanlike teleoperated androids.* .

[23] "FORPHEUS | Our technology | Technology | OMRON Global." [Online]. Available: https://www.omron.com/global/en/technology/information/forpheus/. [Accessed: 27-Sep-2020].

[24] S. Behnke, J. Müller, and M. Schreiber, "Playing soccer with RoboSapien," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2006, vol. 4020 LNAI, pp. 36–48.

[25] S. Behnke, J. Müller, and M. Schreiber, "Toni: A soccer playing humanoid robot," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2006, vol. 4020 LNAI, pp. 59–70.

[26] D. C. Bentivegna, A. Ude, C. G. Atkeson, and G. Cheng, "Humanoid robot learning and game playing using PC-based vision," *IEEE Int. Conf. Intell.*

[27] *Robot. Syst.*, vol. 3, pp. 2449–2454, 2002.

[27] L. Hung *et al.*, "The benefits of and barriers to using a social robot PARO in care settings: A scoping review," *BMC Geriatrics*, vol. 19, no. 1. BioMed Central Ltd., p. 232, 23-Aug-2019.

[28] R. Yu *et al.*, "Use of a Therapeutic, Socially Assistive Pet Robot (PARO) in Improving Mood and Stimulating Social Interaction and Communication for People With Dementia: Study Protocol for a Randomized Controlled Trial," *JMIR Res. Protoc.*, vol. 4, no. 2, p. e45, May 2015.

[29] S. Paepcke and L. Takayama, "Judging a bot by its cover: An experiment on expectation setting for personal robots," *5th ACM/IEEE Int. Conf. Human-Robot Interact. HRI 2010*, pp. 45–52, 2010.

[30] H. Kozima, M. P. Michalowski, and C. Nakagawa, "Keepon: A playful robot for research, therapy, and entertainment," *Int. J. Soc. Robot.*, vol. 1, no. 1, pp. 3–18, 2009.

[31] C. Griffith, "Make it Move," *Real-World Flash Game Dev.*, no. August, pp. 85–95, 2010.

[32] N. Akalin, P. Uluer, H. Kose, and G. Ince, "Humanoid robots communication with participants using sign language: An interaction based sign language game," *Proc. IEEE Work. Adv. Robot. its Soc. Impacts, ARSO*, pp. 181–186, 2013.

[33] B. Robins *et al.*, "Human-centred design methods: Developing scenarios for robot assisted play informed by user panels and field trials," *Int. J. Hum. Comput. Stud.*, vol. 68, no. 12, pp. 873–898, Dec. 2010.

[34] "IEEE Human Application Challenge: Robot Magic and Music."

[35] K. J. Morris, V. Samonin, J. Anderson, M. C. Lau, and J. Baltes, *Robot Magic: A Robust Interactive Humanoid Entertainment Robot*, vol. 1, no. January. Springer International Publishing, 2018.

[36] K. J. Morris, V. Samonin, J. Baltes, J. Anderson, and M. C. Lau, "A robust interactive entertainment robot for robot magic performances," *Appl. Intell.*, vol. 49, no. 11, pp. 3834–3844, 2019.

[37] A. Redacted, "Interaction and Learning in a Humanoid Robot Magic Performance," *Proc. AAAI Spring Symp. Integr. Represent. Reason. Learn. Robot.*, no. Wooldridge 2009, p. 6, 2018.

[38] "RoboCup - IEEE Robotics and Automation Society - IEEE Robotics and Automation Society." [Online]. Available: https://www.ieee-ras.org/robocup. [Accessed: 27-Sep-2020].

[39] E. Ahmadi, A. G. Pour, A. Siamy, A. Taheri, and A. Meghdari, *Playing Rock-Paper-Scissors with RASA: A Case Study on Intention Prediction in Human-Robot Interactive Games*, no. November. Springer International Publishing, 2019.

[40] H. S. Ahn, I. K. Sa, D. W. Lee, and D. Choi, "A playmate robot system for playing the rock-paper-scissors game with humans," *Artif. Life Robot.*, vol. 16, no. 2, pp. 142–146, 2011.

[41] M. P. Michalowski, S. Sabanovic, and P. Michel, "Roillo: Creating a social robot for playrooms," *Proc. - IEEE Int. Work. Robot Hum. Interact. Commun.*, no. October 2006, pp. 587–592, 2006.

[42] I. Gori, S. R. Fanello, G. Metta, and F. Odone, "All gestures you can: A memory game against a humanoid robot," *IEEE-RAS Int. Conf. Humanoid Robot.*, pp. 330–336, 2012.

[43] I. Leite, M. McCoy, D. Ullman, N. Salomons, and B. Scassellati, "Comparing Models of Disengagement in Individual and Group Interactions," *ACM/IEEE Int. Conf. Human-Robot Interact.*, vol. 2015-March, no. March, pp. 99–105, 2015.

[44] J. Ham, M. van Esch, Y. Limpens, J. de Pee, J.-J. Cabibihan, and S. S. Ge, "The Automaticity of Social Behavior towards Robots: The Influence of Cognitive Load on Interpersonal Distance to Approachable versus Less Approachable Robots," Springer, Berlin, Heidelberg, 2012, pp. 15–25.

[45] B. Mutlu, J. Forlizzi, and J. Hodgins, "A Storytelling Robot: Modeling and Evaluation of Human-like Gaze Behavior," 2006.

[46] A. M. Aroyo, F. Rea, G. Sandini, and A. Sciutti, "Trust and Social Engineering in Human Robot Interaction: Will a Robot Make You Disclose Sensitive Information, Conform to Its Recommendations or Gamble?," *IEEE Robot. Autom. Lett.*, vol. 3, no. 4, pp. 3701–3708, 2018.

[47] O. Palinko, A. Sciutti, Y. Wakita, Y. Matsumoto, and G. Sandini, "If looks could kill: Humanoid robots play a gaze-based social game with humans," *IEEE-RAS Int. Conf. Humanoid Robot.*, pp. 905–910, 2016.

[48] L. Riek, "Wizard of Oz Studies in HRI: A Systematic Review and New Reporting Guidelines," *J. Human-Robot Interact.*, vol. 1, no. 1, pp. 119–136, 2012.

[49] J. Fallon, "Box of Lies." [Online]. Available: https://www.youtube.com/watch?v=Md4QnipNYqM.

[50] D. Pasquali, A. M. Aroyo, J. Gonzalez-billandon, F. Rea, G. Sandini, and A. Sciutti, "Your Eyes Never Lie: A Robot Magician Can Tell if You Are Lying," in *In Proceedings' of HRI (HRI '20) Cambridge conference*, 2020.

[51] J. Gonzalez-Billandon *et al.*, "Can a Robot Catch You Lying? A Machine Learning System to Detect Lies During Interactions," *Front. Robot. AI*, vol. 6, Jul. 2019.

[52] B. M. DePaulo, J. J. Lindsay, B. E. Malone, L. Muhlenbruck, K. Charlton, and H. Cooper, "Cues to deception.," *Psychol. Bull.*, vol. 129, no. 1, pp. 74–

118, 2003.

[53] D. P. Dionisio, E. Granholm, W. A. Hillix, and W. F. Perrine, "Differentiation of deception using pupillary responses as an index of cognitive processing.," *Psychophysiology*, vol. 38, no. 2, pp. 205–11, Mar. 2001.

[54] J. Beatty, "Task-Evoked Pupillary Responses, Processing Load, and the Structure of Processing Resources," 1982.

[55] M. Nakayama and Y. Shimizu, "Frequency analysis of task evoked pupillary response and eye-movement," in *Proceedings of the Eye tracking research & applications symposium on Eye tracking research & applications - ETRA'2004*, 2004, pp. 71–76.

[56] "Tobii Pro Glasses 2."

[57] J. Beatty and B. Lucero-Wagoner, "The pupillary system.," *Handb. Psychophysiol. 2*, 2000.

[58] J. Sweller, P. Ayres, and S. Kalyuga, *Cognitive Load Theory*. Psychology of learning and motivation, volume 55. Elsevier., 2011.

[59] J. Leppink, "Cognitive load theory: Practical implications and an important challenge," *J. Taibah Univ. Med. Sci.*, vol. 12, no. 5, pp. 385–391, 2017.

[60] S. Mathôt, "Pupillometry: Psychology, Physiology, and Function," *J. Cogn.*, vol. 1, no. 1, Feb. 2018.

[61] A. K. Webb, C. R. Honts, J. C. Kircher, P. Bernhardt, and A. E. Cook, "Effectiveness of pupil diameter in a probable-lie comparison question test for deception," *Leg. Criminol. Psychol.*, vol. 14, no. 2, pp. 279–292, Sep. 2009.

[62] S. M. Kassin, "On the Psychology of Confessions: Does Innocence Put Innocents at Risk?," *Am. Psychol.*, vol. 60, no. 3, pp. 215–228, Apr. 2005.

[63] S. Mathôt, J. Fabius, E. Van Heusden, and S. Van der Stigchel, "Safe and sensible preprocessing and baseline correction of pupil-size data," *Behav. Res. Methods*, vol. 50, no. 1, pp. 94–106, 2018.

[64] D. De Tommaso and A. Wykowska, "TobiiglassespySuite: An open-source suite for using the Tobii Pro Glasses 2 in eye-tracking studies," in *Eye Tracking Research and Applications Symposium (ETRA)*, 2019.

[65] P. Fitzpatrick, G. Metta, and L. Natale, "Towards long-lived robot genes," *Rob. Auton. Syst.*, vol. 56, no. 1, pp. 29–45, Jan. 2008.

[66] Tobii Pro, "Quick Tech Webinar - Secrets of the Pupil." [Online]. Available: https://www.youtube.com/watch?v=I3T9Ak2F2bc&feature=emb_title.

[67] "Webcam Logitech BRIO con video Ultra HD 4K e tecnologia RightLight 3 con HDR." [Online]. Available: https://www.logitech.com/it-it/product/brio. [Accessed: 04-Oct-2020].

[68] "Dixit 3: Journey | Board Game | BoardGameGeek." [Online]. Available: https://boardgamegeek.com/boardgame/119657/dixit-3-journey. [Accessed: 27-Sep-2020].

[69] G. Metta, G. Sandini, D. Vernon, L. Natale, and F. Nori, "The iCub humanoid robot," in *Proceedings of the 8th Workshop on Performance Metrics for Intelligent Systems - PerMIS '08*, 2008, p. 50.

[70] G. B. (Universita di M.-B. Flebus, "Versione italiana dei big five markers di goldberg," 2006.

[71] C. J. Ferguson and C. Negy, "Development of a brief screening questionnaire for histrionic personality symptoms," *Pers. Individ. Dif.*, vol. 66, pp. 124–127, 2014.

[72] D. N. Jones and D. L. Paulhus, "Introducing the Short Dark Triad (SD3): A Brief Measure of Dark Personality Traits," *Assessment*, vol. 21, no. 1, pp. 28–41, 2014.

[73] F. Bracco and C. Chiorri, "Versione Italiana del NASA-TLX."

[74] S. Shapiro and M. Wilk, "An analysis of variance test for normality (complete samples)," *Z. Bibliothekswes. Bibliogr.*, vol. 52, no. 6, pp. 311–319, 2005.

[75] R. B. D'Agostino, A. Belanger, and R. B. D'Agostino, "A Suggestion for Using Powerful and Informative Tests of Normality," *Am. Stat.*, vol. 44, no. 4, p. 316, Nov. 1990.

[76] "One-Class Support Vector Machines." .

[77] C. Wangwiwattana, X. Ding, and E. C. Larson, "PupilNet, Measuring Task Evoked Pupillary Response using Commodity RGB Tablet Cameras," *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 1, no. 4, pp. 1–26, Jan. 2018.

[78] J. Klingner, "Measuring Cognitive Load During Visual Task by Combining Pupillometry and Eye Tracking," *Ph.D. Diss.*, no. May, 2010.

[79] A. M. Aroyo *et al.*, "Can a Humanoid Robot Spot a Liar?," *IEEE-RAS 18th Int. Conf. Humanoid Robot.*, 2018.