Sketch it for the robot! How child-like robots' joint attention affects humans' drawing strategies

Carlo Mazzola*
CONTACT Unit
Italian Institute of Technology
Genoa, Italy
0000-0002-9282-9873

Lorenzo Morocutti*
CONTACT Unit
Italian Institute of Technology
Genoa, Italy
0009-0009-4156-8800

Hillary Pedrizzi
Dep. of Applied Inf., CONTACT U.
Comenius University, IIT
Bratislava, Slovakia; Genoa, Italy
0009-0009-6132-1071

Xenia Daniela Poslon
Dep. of Applied Informatics
Comenius University
Bratislava, Slovakia
0000-0001-7476-2890

Andrej Lucny
Dep. of Applied Informatics
Comenius University
Bratislava, Slovakia
0000-0001-6042-7434

Sarah Marie Wingert
Dep. of Applied Informatics
Comenius University
Bratislava, Slovakia
0000-0001-5698-3809

Yukie Nagai
IRCN
The University of Tokyo
Tokyo, Japan
0000-0003-4794-0940

Francesco Rea
CONTACT Unit
Italian Institute of Technology (IIT)
Genoa, Italy
0000-0001-8535-223X

Alessandra Sciutti
CONTACT Unit
Italian Institute of Technology (IIT)
Genoa, Italy
0000-0002-1056-3398

***Abstract*—The work proposes investigating drawing activities in interactive contexts to shed light on the links between socio-cognitive and visual representation mechanisms. In our study, 53 adult participants were instructed to draw some object categories (e.g., a duck, an ambulance, ...) for a child-like robot, which first was just shown in a picture (individual condition) and then was co-present and engaged in joint attention behaviors with them (robot condition). The data collection was carried out in two sites, in Italy (N=26) and Slovakia (N=27), with two different robots, iCub and Nico. Participants significantly changed their drawing strategy in the presence of the robots by enlarging their sketches while speeding up their drawing. The phenomenon was more evident the more the individuals perceived the robot as closer to them, according to the IOS scale. The results were highly consistent between the two sites and showed that participants put more effort into drawing understandably when a robot actively attends to their behavior. Higher clarity is obtained with increased figure size rather than simplifying the drawing. Counter-intuitively, participants did not slow down their tracing to be more comprehensible. Instead, they became faster in front of the robot, potentially induced by the pressure of being observed by it. These findings are discussed in the framework of the motionese literature.**

***Index Terms*—Visual Representation, Motionese, Joint Attention, Human-Robot Interaction, Drawing**

* Carlo Mazzola and Lorenzo Morocutti equally contributed to this work.
The research leading to these results has received funding from the project titled TERAIS in the framework of the program Horizon-Widera-2021 of the European Union under the Grant agreement number 101079338 and by the Slovak Research and Development Agency under contract No. APVV-21-0105.

I. Introduction

Motionese phenomena are action style modifications, such as slowing down one's movements, introducing more segmentation, and standing closer when demonstrating a desired behavior, that occurs naturally when human caregivers interact with infants [3]. Several studies support the idea that humans use a similar movement style when interacting with robots [11], [12], [15], [19]. In this context, the learner's behavior can also influence how a tutor demonstrates an action. Nagai et al. [11] showed that humans not only modify their behavior when demonstrating action to a robot with exaggeration in space and synchronization in time, but the robot's bottom-up attention also influenced motions used by human teachers. Joint attention, the co-orientation established in triadic interactions among the self, another agent, and a third element, played a crucial role in human-robot interaction. The gazing behavior of the robot effectively gained the partners' interest and drew them to establish joint attention (for a review on robots and joint attention, see [4]). In the current study, we wanted to assess whether a similar phenomenon also extended to the activity of drawing for a (child) robot observer and if its actual physical co-presence influenced it.

Graphic illustration via sketches can be considered a gateway to the internal, private world of mental representations. The inherent complexity of drawing activity is connected to human cognition through perception, memory, motor, and social mechanisms [5], [14]. It has been shown that different representation styles follow if humans are asked to draw a

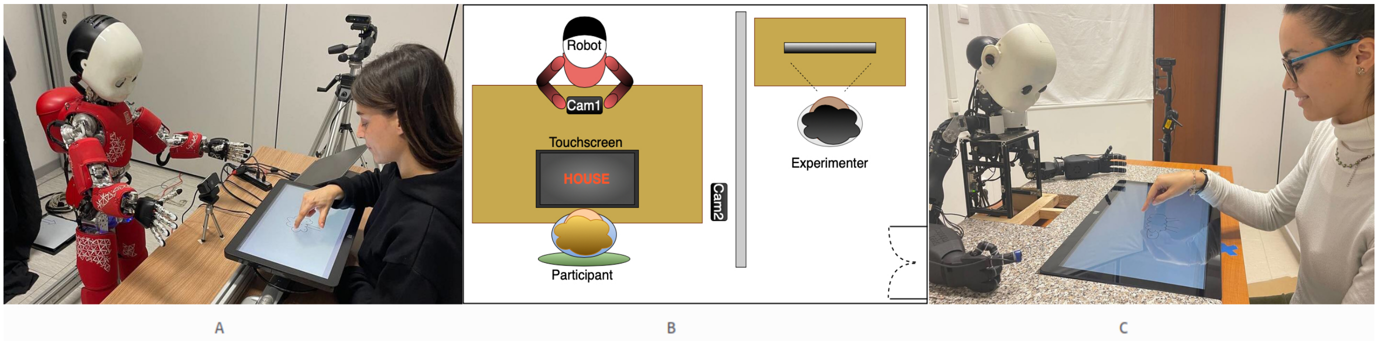

Fig. 1. Pictures and Schema of the experimental setup for the robot condition: A at IIT (Italian site), B schematical representation of the setup, C at UKBA (Slovak site).

specific exemplar or the general category of a given object [20]. Graphic representations also evolve and are gradually modified if the same concept is represented more times within the same social relationship [7].

Yet, whether and how one would modify its sketching behavior when asked to draw for an observer is still unknown. For this reason, we investigated for the first time, if and how these changes occurred in front of and for a social (child) robot observing drawers in their task.

Humanoid robots can provide the embodied context to investigate human representational mechanisms and strategies within social interaction in a controllable and repeatable way [17]. We therefore carried out a two-site experiment by using two child-like robots (iCub and Nico) to investigate the influence of the robot's presence and joint attention behavior on participants who were asked to draw for them. In particular, we aimed at studying mechanisms related to motionese in the context of drawing to evaluate whether and how the human drawing style is modified while graphically demonstrating object categories to a robot observer.

II. Methods

A. Participants

53 human participants (25 M, 25 W, 3 NB) interacted with a social child-like robot in a drawing test at two different sites, at the Comenius University of Bratislava (UKBA, SK, N=27) and the Italian Institute of Technology (IIT, IT, N=26). Only mother-tongue participants from the two countries were recruited to ensure a perfect comprehension of the task. Consequently, the experiment was conducted in Slovak and Italian. All participants provided written informed consent before their participation. At IIT, the study received approval from the regional ethical committee, Comitato Etico Regione Liguria; and at UKBA, from the ethical committee of the Faculty of Mathematics, Physics and Informatics. Participants were compensated with a sum of 10 € for their time.

B. Design

Participants were asked to draw object categories on a touchscreen with their index finger, first alone (individual condition) and then in the presence of the robot (robot condition).

*1) **Setup**:* In both experimental sites, the rooms were divided into two compartments to limit, as much as possible, the pressure of the experimenter on the participant (see Fig.1). The experimenter could monitor the experiment through two cameras to have a frontal and lateral view of the participant. The experimental compartment was arranged with a chair, a desk, and an LCD Touchscreen Monitor placed on it. The models adopted for the experiment were the ELO 2002L at IIT (436.9x240.7 mm, 1920x1080 px, 60 Hz), and ELO 2202L at UKBA (476.06x267.79 mm, 1920x1080 px, 60 Hz). A cross-colored tape marked the starting position of the participants' hands before the beginning of the drawing session. In the robot condition, additionally, the robot was positioned in front of the participant on the opposite side of the touchscreen.

*2) **Experimental sessions**:* To investigate the differences in the cognitive mechanisms and strategies elicited by the robot's presence and behavior, the experiment consisted of two sessions (conditions). A pause lasting about 5 minutes separated the two. The order of conditions was always the same: the first condition was planned to be without the robot to avoid any bias caused by the encountering of the robot.

Individual Condition. Participants were shown a picture of the robot and instructed to draw 12 different object categories (e.g., Computer, Duck, see table I for the complete list) with the following goal: "Make the robot in the picture understand your drawings." Three categories were requested twice to check for repetition effects. At the end of each drawing, participants rated the difficulty of the task on a Likert scale (ranging from 1 to 7).

Robot Condition. In the following condition, the robot stood in front of the participants, welcoming them by looking them in the face when they entered the room and before each drawing and looking at the screen during the drawing completion. Participants were instructed to draw 12 object categories, out of which six were already present in the individual condition. After participants finished drawing, the robot gave neutral vocal feedback: "Ok!". This choice follows the idea that a lack of feedback could be interpreted as anti-social behavior, while more expressive or differentiated positive feedbacks could be an additional factor influencing participants and leading them to verbally interact with the robot (something we preferred

TABLE I
OBJECT CATEGORIES DIVIDED FOR EACH CONDITION.

INDividual 15 drawings (12 categories)		**ROBot 12 drawings (12 categories)**
object categories	repeated	object categories
Bee, Bus, Sheep	Bee, Bus, Sheep	Bee, Bus, Sheep
Duck, Face Computer		Duck, Face Computer
Alarm Clock, Ambulance, Ant Crab, Drums, Penguin		Map, Mosquito, Pig, Pizza, Sea Turtle, Teddy Bear

to avoid to maintain control on the experiment). Additionally, we also reduced non-verbal backchanneling (except for mutual gaze) by removing facial expressions to avoid differences between the robots (Nico was not endowed with them. After each drawing, while participants scored the task's difficulty, the robot looked away to avoid exerting pressure on them.

*3) **Stimuli (object categories)**:* 18 object categories were chosen from the ones included in the Google Quick Draw Dataset [8]. Table I shows the selected categories according to the condition(s) they were presented in. The categories were sequentially presented to participants on the touchscreen as a visual stimulus (red writing on a black window) lasting 4 seconds. Categories were randomized to avoid any bias. After the category presentation, a white canvas opened on the touchscreen as a virtual sheet participants had to sketch on.

C. The Robots

The use of humanoid robots in this study aimed to exploit the controllability and repeatability of robots' actions and their human-like, social appearance and behavior to study human cognitive mechanisms in an interactive embodied scenario. We used an iCub robot (IIT) [10] and a Nico robot (UKBA) [9] that was inspired by the iCub in its design (see Fig. 1 for a picture of the two robots).

*1) **iCub**:* iCub is a complex robotic platform developed to study human cognition with computational and Human-Robot Interaction approaches. It has been designed with the shape of a 5-year-old human child and can engage in social interactions given its motor-cognitive and social abilities. The sensors used in this experiment are the two cameras installed in its eye cavities. The three DoFs of the neck and the three for the eyes allowed the robot to perform head-gaze movements generated with the iKinGazeCtrl module [16]. Specifically, the robot looked at the touchscreen during the drawing activity, at some random points in the room away from the participant during the question time after the drawing, and it tracked the participant's face at the beginning of the experiment and before every drawing. The LEDs to generate facial expressions were switched off to make it appear more similar to the version of Nico at UKBA. The various modules communicated through YARP middleware (v3.8)[1,2].

*2) **Nico**:* Nico was designed and built taking inspiration from the iCub to produce a cheaper version of it. Nico's neck two DoFs were controlled to allow the same gaze behaviors as iCub. Based on the iCub head, it has 2 overlapping cameras in its eye cavities. As for the iCub, we used the cameras for participants' Face-Tracking. Nico's version at UKBA is not endowed with legs, which, in any case, were not needed for the interaction. Thus, it was placed directly on the table. Nico's code was written in python3.8[1,2].

[1]DOI of codes with git links: https://doi.org/10.5281/zenodo.10944480
[2]DOI of data: https://doi.org/10.5281/zenodo.10943977

D. Questionnaires and scales

Participants were asked to fill out several surveys before the beginning of the experiment after observing a picture of the robot. The same questionnaires were also submitted at the end of the experiment, i.e. after the interaction.

Inclusion of Other in the Self (IOS): This single-item scale [1] measures the level of closeness the respondent experiences towards another agent or a group.

Godspeed Questionnaire: The participants filled out the scales Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety [2] immediately after the IOS, in Italian and Slovak. The Godspeed Questionnaire was preferred to other questionnaires used in HRI because of its extensive use and availability in many languages. Since no additional HRI questionnaire was present in Slovak, it was easier to translate it with a double check from the English and the Czech versions.

Additionally, after each drawing, we asked participants to self-evaluate the difficulty they faced in drawing that specific sketch with a scroll bar ranging from 1 to 7 (*Task Difficulty Survey*).

E. Data Analysis

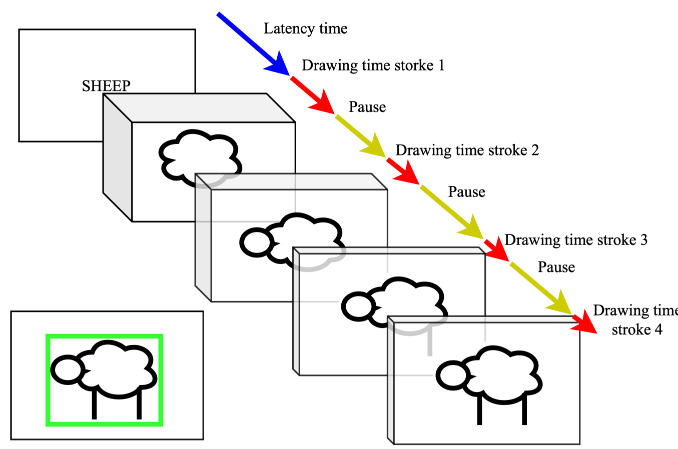

Fig. 2. Graphic representation of the drawing activity with some of the extracted drawing parameters illustrated. After the window with the object category is shown to participants, Latency Time (blue arrow) is computed until the first stroke is sketched. Drawing Time (red arrows) and Pauses (yellow arrows) are calculated as shown in the picture. The green rectangle shows the bounding box for this specific drawing.

*1) **Feature Extraction**:* To test our hypothesis, we based our analysis on quantitative measures extracted from the drawing activity (see Fig. 2). The basic components of drawings are strokes, which are drawing traits produced by a continuous touch of the finger on the screen. Strokes are defined by triplets of data (*x*, *y*, *t*), representing the Cartesian coordinates of the trait, *x* and *y*, measured in pixels, for each timestamp, *t*, measured in nanoseconds. Pixels coordinates were measured by taking the top left corner of the screen as a reference and ranging according to the screen resolution (1920x1080). Data were collected with the Python library Tkinter. Through strokes, spatial and temporal information of the drawing became available. Linking such information with the timestamp of the stimulus window used to show the object category to participants, we could extract all the parameters needed to evaluate different features of the drawing as follows.

Latency time. It is the time a participant takes, after being instructed on the category to draw, to reach the screen and start drawing. It is computed as the difference between the timestamp taken at the opening of the canvas and the timestamp of the first "touch" of the screen. It can be interpreted as the time needed to visually recall the object category.

Avg. Pause. Pauses are thought of as the difference in time between the first timestamp of one stroke t_{start_i} and the last timestamp of the previous one $t_{end_{i-1}}$. It can be interpreted as the reflection time while drawing, when trying to fill the gap between one's mental and graphic representation.

$$Avg.Pause = \frac{1}{n}\sum_{i=2}^{n}(t_{start_i} - t_{end_{i-1}}) \quad (1)$$

Drawing Time. It is the total time of actual drawing activity, computed as the total time drawers use to sketch the strokes. It is the difference between the final timestamp t_{end_i} of a stroke and its initial one t_{start_i}.

$$DrawingTime = \sum_{i=1}^{n}(t_{end_i} - t_{start_i}) \quad (2)$$

Stroke Number. It is the total number of strokes used to complete the drawing. Every time participants lifted their finger and touched the screen again, a new stroke was created. It can indicate the *amount of details* inserted in the drawing.

For what concerns the spatial features, considering each stroke formed by *m* micro-traits, the *Total Stroke Length* (*Tot.StrkL.*) computed on *n* strokes can be calculated by summing the micro-traits length of each stroke *j*, considered as the difference between the coordinates of two subsequent pixels (*i* and *i*-1):

$$Tot.StrkL. = \sum_{j=1}^{n}\sum_{i=2}^{m}\sqrt{(x_{j,i} - x_{j,i-1})^2 + (y_{j,i} - y_{j,i-1})^2} \quad (3)$$

Average Length (*Avg.Len.*) can be derived from 3 dividing the total stroke lengths for the number of strokes:

$$Avg.Len. = \frac{1}{n}\sum_{j=1}^{n}\sum_{i=2}^{m}\sqrt{(x_{j,i} - x_{j,i-1})^2 + (y_{j,i} - y_{j,i-1})^2} \quad (4)$$

TABLE II
DESCRIPTIVES OF DRAWING FEATURES IN TERMS OF AVERAGES AND STANDARD DEVIATIONS

Drawing features	**All Object Categories**			
	INDIVIDUAL		ROBOT	
	mean	std	mean	std
Latency Time (s)	3.57	3.68	2.84	1.74
Avg. Pause (s)	1.17	0.5	0.97	0.37
Drawing Time (s)	17.3	8.4	17.0	7.89
Stroke Number (#)	23.5	12.5	22.2	11.5
Tot. Length (cm)	132.0	91.7	158.0	98.9
Avg. Length (cm)	6.61	5.01	8.13	5.08
Bounding Box (cm^2)	175.0	139.0	238.0	173.0
Avg. Velocity (cm/s)	7.97	4.01	9.78	5.0

Bounding Box. The dimension of the imaginary rectangle area enclosing the final drawing, computed from the greatest and smallest pixel coordinates of the drawing

$$BoundingBox = (x_{MAX} - x_{min}) \times (y_{MAX} - y_{min}) \quad (5)$$

The Tot.StrkL., the Avg.Len. and the BoundingBox can be used to indicate the amplitude of the drawing.

Avg. Velocity. Computed the length of the strokes and considering the contributions of the strokes' *Drawing Time*, it is immediate to find the velocity of each stroke as $\frac{length}{time}$

*2) **Hypotheses and Statistical Analysis**:* The idea underlying this study was that the robot's embodied presence and joint attention to the drawing activity would have increased participants' effort to be understood, leading to changes in their drawing strategy. More specifically,

HP1: given the child-like shape of our robots and following suggestions from the literature on motionese concerning the human caregivers' tendency to exaggerate movements to interact with children [3], we expected an increased drawing amplitude (i.e., increased Tot. Length, Avg. Length and Bounding Box parameters) in the robot condition.

HP2: based on motionese literature (e.g., [15]), we also hypothesized a lower Avg. Velocity when sketching strokes.

HP3: we expected an increased Drawing Time in the robot condition as a parameter connected to the higher level of engagement.

We did not have strong expectations for the Stroke Number, the Latency time, or the Avg. Pause.

To extract the abovementioned parameters, we used Python 3.8, whereas we used Jamovi 2.4.11 for statistical analysis. More specifically, we used Linear Mixed Models from [6].

III. RESULTS

To assess whether and how the robot's presence and behavior affected participants' drawing activity and graphic representations, we compared the two main conditions of our experiment – individual vs robot – using several parameters for a quantitative evaluation.

We statistically evaluated the Condition effect (i.e., the effect produced by the robot's presence and joint attention

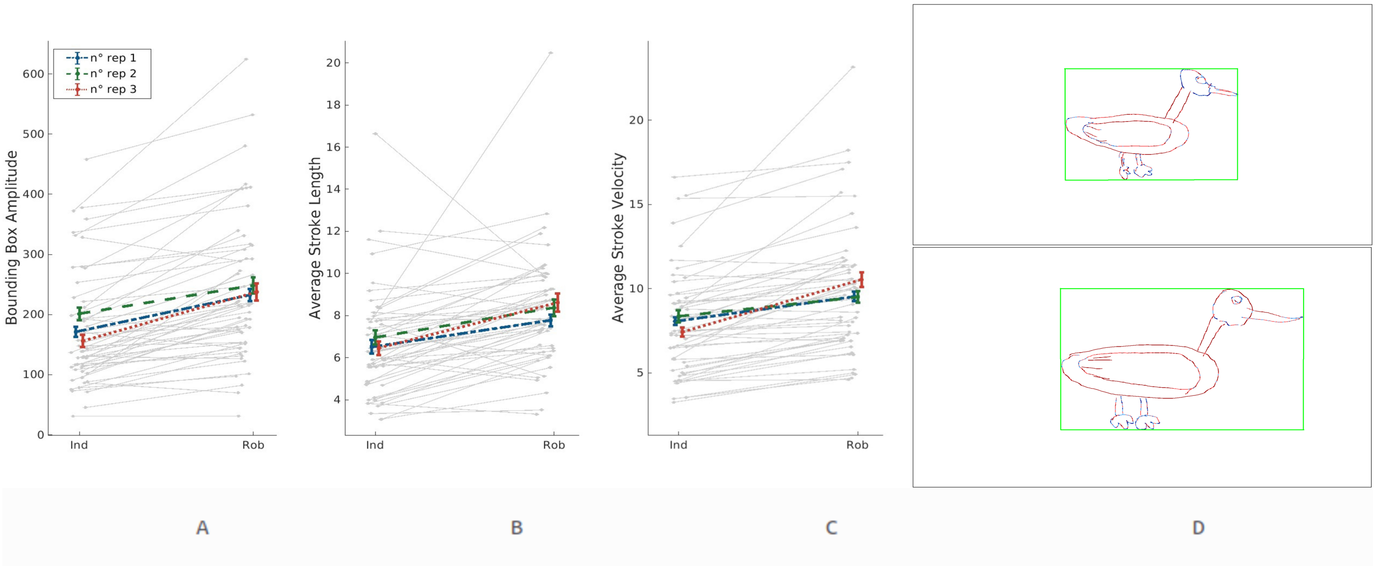

Fig. 3. Plots representing the mean and standard error (error bars) values for Bounding Box (A), Avg. Length (B) and Avg. Velocity (C) for every participant, both in INDividual and ROBot conditions. Global means of these features are shown for the object categories represented 1, 2, and 3 times. Grey dots and lines show each participant's means for the two conditions. Figure D shows the two drawings for 'Duck' of one representative participant, one sketched in the Individual (upper one), the other in the Robot condition (bottom one). The two drawings show the larger Bounding Box, longer Avg. Length, and faster Avg. Velocity for the robot condition. The two Ducks are colored using two colors, two shades per color: red for faster micro-traits (the darker, the faster), and blue for slower micro-traits (the darker, the slower).

behavior), considering it a predictor for each of these features (dependent variables), by fitting 8 different Linear Mixed Models (LMM), one for each drawing parameter described in Section II-E (averages and standard deviations in Table II). In this way, it was possible to use the same model to compute the effect of other possible predictors – Number of Representations (i.e., the number of times an object category was represented by a participant, thus assessing the repetition effect) and Experimental Site – and incorporate the random effect of other factors – participants and object categories. The random effect of participants was applied to adjust for each participant's baseline and model the intra-subject correlation of repeated measurements. The random effect at the object categories level served to model inter-stimulus variability in the error parameters. Random effects were submitted to the model in this order.

We computed the shift of the robot condition from the individual one, considered as a baseline. We found a significant decrease in the Latency Time (ROB – IND: B=-0.801, t=-4.188, p<0.001) and the Avg. Pause (ROB – IND: B=-0.209, t=-7.197, p<0.001) and a significant increase in the Tot. Length (ROB – IND: B=27.760, t=3.448, p=0.002), the Avg. Length (Ind – Rob: B=1.61, t=4.563, p<0.001), the Avg. Velocity (ROB – IND: B=1.921, t=6.508, p<0.001), and the Bounding Box (ROB – IND: B=63.25, t=4.93, p<0.001) (Fig. 3 graphically shows the Tot. Length, Avg. Velocity, and Bounding Box results). After applying the Bonferroni correction to our tests, post hoc results confirmed the significance of the effects. No significant effect of Condition was found for the Stroke Number and the Drawing Time.

Moreover, we analyzed the questionnaires that participants filled out before the experimental phase and after completing all the tasks. The measures used in questionnaires were used to examine whether the differences in participants' drawing representation between conditions could be explained by their perception of the robot. Therefore, we analyzed participants' evaluation of the robot (measured using IOS and Godspeed scales) by assessing differences between their pre-experiment expectations and post-experiment rating and tested if such differences correlated with how participants' drawing features varied from the individual to the robot condition. The only statistically significant difference reported by paired T-Tests (or Wilcoxon signed-rank tests when required by non-normal distributions) was in Godspeed's Anthropomorphism scale with a decrease from the pre-experiment (*M*=14.5 *Std*=3.69) to the post-experiment rating (*M*=11.2 *Std*=4.21). For all other scales we did not find any significance between the tests submitted before and after the interaction with the robot (all ps > 0.05).

Concerning the correlation analysis, Spearman's rank correlation revealed a significant positive correlation of delta-IOS (difference between post-experiment and pre-experiment), with delta-Bounding Box (difference between the robot and the individual condition): r(51)=0.361, p=0.045, see Fig. 4. The same positive correlation was found between delta IOS and Delta-Drawing Time (r(51)=0.320, p=0.020). Delta-Bounding Box and Delta-Drawing Time were computed by averaging all the categories for each participant.

To check the effect of repetition with respect to the robot effect, we looked more deeply into the results of the 8 above-mentioned LMM. No statistical significance was found concerning the effect of the Number of Representations and its interaction with the Condition. The only exceptions were the Avg. Pause, the Tot. Length, and the Avg. Velocity. In particular, for the Avg. Pause, a significant effect of interaction between the Number of Representations and the Condition was evident between categories drawn 3 times and those drawn 2 or 1 times, with a significantly greater decrease of the Avg. Pause in the robot condition for the categories represented 3 times (ROB – IND * 2 – 3: B=0.166, t=2.855, p=0.004, and ROB – IND * 1 – 3: B=0.168, t=2.182, p=0.037). For the Tot. Length and the Avg. Velocity, the interaction effect goes in the opposite direction between categories drawn 3 and 2 times, with a significantly greater increase of the two parameters in the robot condition for the categories represented 3 times (Tot. Length: ROB – IND * 2 – 3: B=-32.159, t=-2.862, p=0.004, Avg. Velocity: ROB – IND * 2 – 3: B=-1.966, t=-4.199, p<0.001). Moreover, we checked if we had the same effect also considering only categories represented 1 time (in which no repetition/habit effect is present). We found a significant difference in both features (Avg. Velocity: ROB - IND: t=5.575, p<0.001, Tot. Length: ROB - IND: t=3.463, p<0.001).

We analyzed whether the different experimental sites and types of robots influenced the drawing parameters and the above-mentioned robot effects. The only drawing feature that was different between different experimental sites was the Avg. Length, which was significantly higher for the drawing completed in Slovak (SK – IT: B=1.426, t=2.372, p<0.021). No interaction effect of the Experimental Site with the Condition (individual/robot) was found.

Eventually, to assess if any relationship occurred between participants' self-reported Drawing Difficulty and the above-mentioned drawing features, we ran a Spearman's rank correlation by averaging the results for each category. A positive correlation was found when checking the self-reported Drawing Difficulty with the Latency Time (r(10)=0.783, p=0.004) and with the Avg. Pause (r(10)=0.708, p=0.010) but not with the other drawing features. There was also a positive correlation between Latency Time and Avg. Pause: r(10)=0.620, p=0.032.

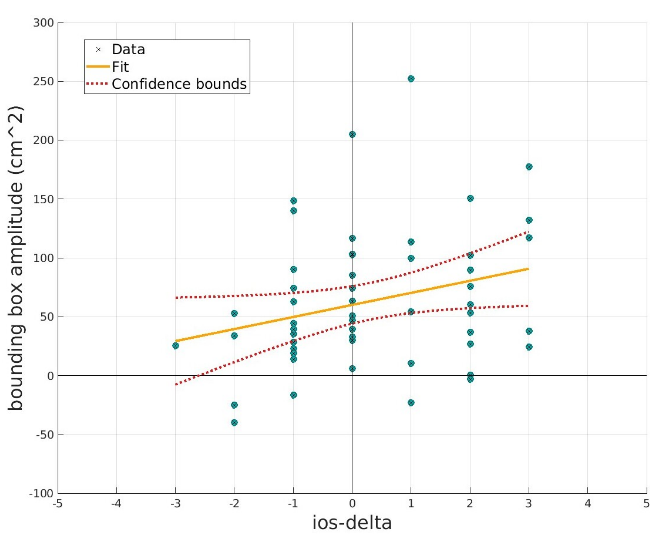

Fig. 4. Scatter plot of the shift in Bounding Box (Delta-Bounding Box: computed as robot-individual) plotted on the shift in IOS (Delta-IOS: computed as post-pre experiment). The positive linear correlation is highlighted by the yellow linear fit of data.

IV. Discussion

A. *The co-presence and joint attention effect on drawings.*

This study is based on the hypothesis that, when asked to sketch pictures for another agent, humans' graphic activity and outcomes are affected by the agent's presence and joint attention behavior during the drawing completion. Statistical analysis performed with LMM revealed the robot's joint attention behavior affected most of the drawing features: in some cases confirming our hypothesis, in others falsifying them or shedding light on a different interpretation of the effect.

1) ***Enlarging drawings (HP1 confirmed) rather than changing the details number:*** Participants drew larger sketches in front of the robot, as evidenced by the higher Tot. Length, Avg. Length, and Bounding Box. The effect was evident independently from the number of times the object categories were drawn, evidencing that the robot's presence and joint attention led drawers to extend strokes and produce a larger - and supposedly more comprehensible - drawing. These findings are coherent with what is suggested by motionese theories [3]. When demonstrating action to a child, adults tend to perform an increased range of motion than in front of adults to enhance clarity. It is possible, therefore, that the child-like shape of iCub and Nico might have pushed the drawers' tendency to enlarge their sketches when in front of the robot.

This result is even more interesting if we consider that the increased drawing amplitude is connected with the degree of robots' inclusion in the participants' self. The shift of IOS from before to after the experiment positively correlated with the shift in the Bounding Box from the individual to the robot condition, which suggests that the closer to the robot participants reported feeling, the more they enlarged their drawing: a result that seems to confirm our theory about the drawing amplitude enlarged to help the robot understand the drawings.

No difference was found in the Stroke Number, showing that the robot's presence did not influence the drawers' strategy in adding or removing details. Previous studies demonstrated that the Number of Strokes depends on the drawer's goal in representing an object's specific exemplar or its general category [20], but in our study, this request was not specified. Another study proved this drawing feature to depend on the number of times the same concept is repeated within a social interaction [7], but this is not the case in our experiment, where categories were repeated only once with the robot, and even focusing only on those that were first drawn in the individual condition and then repeated with the robot, no decrease in the number of strokes was evident.

2) ***Drawing faster in the presence of an observer (HP2 disconfirmed):*** Contrary to our second hypothesis, we found that Avg. Velocity increased with the robot. In front of the robot observing their activity, participants drew significantly faster regardless of the number of repetitions. While categories repeated 2-3 times showed an (anticipated) habituation effect, attributed to familiarity, it's noteworthy that this effect persisted even for categories drawn only in one condition, even if to a lesser extent. This suggests that the increase in speed cannot be solely attributable to habituation. A possible explanation is that velocity could be connected with the increased amplitudes of drawing in the robot condition: longer traits are drawn faster. However, the increased velocity might also be associated with some social pressure exerted by the robot observing participants, similar to previous research, where the robot's eye gaze impacted human decision-making, fastening their response time [18].

In any case, the velocity of drawing completion could depend on the goal of the social interaction. Our study investigated a drawing activity performed to let the robot understand the object category represented in the sketch. Different might be the case of a drawing demonstration finalized to teach children (or robots, as suggested by Nishide et al. [13]) to draw pictures or simple shapes. In this case, motionese effects might impact the drawing style, slowing the demonstrator's movements, as hypothesized by HP2.

*3) **Drawing time as a function of closeness to the robot (HP3 revised)**:* Considering HP3, the results of the Drawing Time seemed to contradict our expectations. From the LMM analysis, the Drawing Time was not found to be longer in the robot condition. However, it is interesting to note that the feeling of closeness to the robot reported by the IOS test after the experiment, concerning the one reported before interacting with the robot, was associated with the drawers' activity. The correlation between this shift (computed as the difference between pre and post-experiment) and the shift of the Drawing Time between the Individual and the Robot condition revealed that the more participants felt closer to the robot, the more they spent time on the drawing, which made us reconsider HP3. In this case, this positive correlation with Drawing Time may lead us to interpret the amount of time spent drawing as a measure of participants' engagement with the robot.

*4) **The robot effect on the other parameters connected with completion time**:* Two other parameters investigated in this study are linked to the temporal dimension of the drawing: the Latency Time and Avg. Pause. In the robot condition, LMM analysis highlighted a diminished amount of time spent by participants before starting to draw (Latency Time) as well as between one stroke and another (Avg. Pause). On the one hand, these two parameters may be evaluated to interpret respectively 1) the complexity in visual recall of the object category (the more complex recall, the longer time interval before starting the drawing) and 2) the difficulty experienced in filling the gap between the mental and the graphic representation (the more difficult, the longer the pauses between one stroke and another). The positive correlation found between these parameters and the Perceived Difficulty scored by participants after each drawing suggests that the representational complexity is connected with such drawing features. On the other hand, the decreased amount of time in the robot condition could be explained similarly to the results of Avg. Velocity. Observing participants while sketching the pictures, the robot might have exerted social pressure on them to complete the task faster than they would have done alone. Without more cues, the social explanation seems preferable, but future research is needed to investigate this phenomenon further.

*5) **Overview on effects elicited by the robot**:* Considering all these findings together, drawers appeared to have opted for enlarging the final drawing (i.e., the outcome of the activity) to represent and demonstrate the object category as clearly as possible to the robot in front of them. Coherently with their task (i.e., producing a comprehensible drawing), the strategy used by participants was focused on the final outcome of the activity more than on demonstrating each stroke separately. These results may contribute to comprehending motionese phenomena in drawing tasks. Specifically, if the task is to create comprehensible sketches - in our case, for a child-like robot - the strategy to clearly represent the object category results in enlarging the final sketch rather than slowing movements performed to draw strokes.

The robot's observing behavior is also tied to other effects than widening the sketches. With the robot, drawers start their task faster, take shorter pauses between strokes, and sketch strokes faster: all elements that could be interpreted with some social pressure in speeding up the task completion exerted by the robot's gaze. All these findings are coherent among the number of times object categories were represented. In three cases, the difference between conditions was larger for categories repeated three times, particularly for the Avg. Pause, evidencing that repetition could affect the completion time and be an additional factor in fastening the pauses between strokes.

The influence provoked by the robot's observing behavior is even more relevant if we consider the weak manipulation designed between the two main conditions and demonstrated by the Godspeed questionnaire results before and after the interaction. The robot's behavior was not engineered to strongly promote engagement and social interaction. We aimed to study the clear-cut effect of the robot's co-presence and joint attention. It is true that, albeit minimal, the robot provided feedback to participants. A lack of that could be considered an anti-social behavior rather than a neutral one or raise doubt the robot was monitoring the interaction. Besides that, the robot only looked at the participants' faces before the drawing started and at the touchscreen during the drawing. It is possible that a stronger manipulation of the robot's behavior toward sociality intensifies the effects revealed by this study, as suggested by the positive correlation of IOS results with the Drawing Time and the Bounding Box data.

V. Conclusions

A. Core findings

The core findings of this research aligned with most of our initial expectations but also unearthed additional insights about graphic representations in front of a social (child-like) robot.

In this scenario, individuals altered their drawing strategies, aiming at clarity by enlarging their drawings rather than adding details. Notably, these alteration were more pronounced the closer they felt to the robot. Moreover, the decrease in Latency time and Avg. Pause in the presence of the robot underscores a possible social pressure to complete the task efficiently.

B. Reproducibility of experimental design

This study was conducted in two sites, with two different robots, although similar in shape. Theoretically, differences consisted of different cultures (Italian and Slovak), different robots (iCub and Nico), and different, albeit slightly, touchscreens (same model, the Slovak 2" larger). The different touchscreens might explain the higher Avg. Length found in the Slovak population. The lack of interaction effect between the Country and the Condition suggests that the study's main findings have been replicated across the two experimental sites.

One of the objectives of this study was to propose an experimental setup to study human cognition with drawing tasks in interactive scenarios. Drawings can indeed be an insight into human cognition and mental representations, also in the context of communication [5]. Given the controllability and reproducibility of robots' actions, Human-Robot Interaction may be an effective solution to study social cognition through embodied interactions [17]. In this context, our experimental design proved to be reproducible in different laboratories and with different robotic platforms and may be further used to deepen new aspects of cognition. Moreover, the several parameters we used to investigate different drawing features might be an asset to further explore the links between mental and graphic representations.

C. Shortcomings and future directions

A limitation of this study is visible in the robots' social behaviors. These had been intentionally design to be minimal to gain greater control on the interaction but some participants suggested that the repetitive movements made the robot seem like a puppet more than a sentient agent. After the baseline traced by the present study, future ones may be focused on the effect of more complex social cues. Future research may also involve qualitative analysis of drawings, an aspect not considered in our study that could bring interesting insights. Moreover, this setup can be used to investigate other social and developmental cognition aspects. This study left some open questions, such as how a different goal (e.g., teaching to draw) or the impact of an active collaborative behavior (with respect to a passive observer role) affects the drawing activity. However, we believe that evaluating the influence of the other's presence and joint attention behavior, as we did in this study, is the first crucial step to understanding the socio-cognitive mechanisms and strategies we use to represent things to others and achieve a shared representation with them.

ACKNOWLEDGMENT

We thank Hana Hornáčková, Marvin Steven Marz, and Teodora Savić for their assistance with the data collection and Fabio Vannucci for his contribution to the iCub code.

REFERENCES

[1] Arthur Aron, Elaine N. Aron, and Danny Smollan. Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, 63(4):596–612, October 1992.

[2] Christoph Bartneck. *Godspeed Questionnaire Series: Translations and Usage*, pages 1–35. Springer International Publishing, 2023.

[3] Rebecca J. Brand, Dare A. Baldwin, and Leslie A. Ashburn. Evidence for 'motionese': modifications in mothers' infant-directed action. *Developmental Science*, 5(1):72–83, March 2002.

[4] Pauline Chevalier, Kyveli Kompatsiari, Francesca Ciardo, and Agnieszka Wykowska. Examining joint attention with the use of humanoid robots-a new approach to study fundamental mechanisms of social cognition. *Psychonomic Bulletin & Review*, 27(2):217–236, 2020.

[5] Judith E. Fan, Wilma A. Bainbridge, Rebecca Chamberlain, and Jeffrey D. Wammes. Drawing as a versatile cognitive tool. *Nature Reviews Psychology*, 2(9):556–568, July 2023.

[6] Marcello Gallucci. Gamlj: General analyses for linear models. *Jamovi Module)*, 2019.

[7] Robert D. Hawkins, Megumi Sano, Noah D. Goodman, and Judith E. Fan. Visual resemblance and interaction history jointly constrain pictorial meaning. *Nature Communications*, 14(1), April 2023.

[8] Jonas Jongejan, Henry Rowley, Takashi Kawashima, Jongmin Kim, and Nick Fox-Gieg. The quick, draw!-ai experiment. *Mount View, CA, accessed Feb*, 17(2018):4, 2016.

[9] Matthias Kerzel, Erik Strahl, Sven Magg, Nicolás Navarro-Guerrero, Stefan Heinrich, and Stefan Wermter. Nico—neuro-inspired companion: A developmental humanoid robot platform for multimodal interaction. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 113–120. IEEE, 2017.

[10] Giorgio Metta, Lorenzo Natale, Francesco Nori, Giulio Sandini, David Vernon, Luciano Fadiga, Claes Von Hofsten, Kerstin Rosander, Manuel Lopes, José Santos-Victor, et al. The icub humanoid robot: An open-systems platform for research in cognitive development. *Neural networks*, 23(8-9):1125–1134, 2010.

[11] Y. Nagai, A. Nakatani, and M. Asada. How a robot's attention shapes the way people teach. In *in Proc. 10th Int. Conf. Epigenetic Robot*, pages 81–88, 2010.

[12] Y. Nagai and K.J. Rohlfing. Computational analysis of motionese toward scaffolding robot action learning. *IEEE Transactions on Autonomous Mental Development*, 1(1):44–54, May 2009.

[13] Shun Nishide, Keita Mochizuki, Hiroshi G. Okuno, and Tetsuya Ogata. Insertion of pause in drawing from babbling for robot's developmental imitation learning. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2014.

[14] Anja Philippsen, Sho Tsuji, and Yukie Nagai. Quantifying developmental and individual differences in spontaneous drawing completion among children. *Frontiers in Psychology*, 13:783446, 2022.

[15] Karola Pitsch, Anna-Lisa Vollmer, and Manuel Mühlig. Robot feedback shapes the tutor's presentation: How a robot's online gaze strategies lead to micro-adaptation of the human's conduct. *Asymmetry and adaptation in social interaction*, 14(2):268–296, July 2013.

[16] Alessandro Roncone, Ugo Pattacini, Giorgio Metta, and Lorenzo Natale. A cartesian 6-dof gaze controller for humanoid robots. In *Robotics: science and systems*, volume 2016, 2016.

[17] Alessandra Sciutti and Giulio Sandini. Interacting with robots to investigate the bases of social interaction. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(12):2295–2304, 2017.

[18] Christopher Stanton and Catherine J. Stevens. Robot pressure: The impact of robot eye gaze and lifelike bodily movements upon decision-making and trust. In Michael Beetz, Benjamin Johnston, and Mary-Anne Williams, editors, *Social Robotics*, pages 330–339, Cham, 2014. Springer International Publishing.

[19] A. L. Vollmer, K. S. Lohan, K. Fischer, Y. Nagai, K. Pitsch, J. Fritsch, K. J. Rohlfing, and B. Wredek. People modify their tutoring behavior in robot-directed interaction for action learning. In *IEEE 8th International Conference on Development and Learning*, pages 1–6, 2009.

[20] Justin Yang and Judith Fan. Visual communication of object concepts at different levels of abstraction. *Journal of Vision*, 21(9):2951, September 2021.