Do beliefs about a robot's capabilities influence alignment to its actions?

Anna-Lisa Vollmer*, Britta Wrede†, Katharina J. Rohlfing† and Angelo Cangelosi*
*School of Computing and Mathematics, Plymouth University, Plymouth, UK
†CITEC Center of Excellence Cognitive Interaction Technology, Bielefeld University, Bielefeld, Germany

Abstract—**Interlocutors in a dialog align on many aspects of behavior (word choice, speech rate, syntactic structure, gestures, facial expressions, etc.). Such alignment has been proposed to be the basis for communicating successfully. We believe alignment could be beneficial for smooth human-robot interaction and facilitate robot action learning from demonstration. Recent research put forward a mediated communicative design account of alignment according to which interlocutors align stronger when they believe it will lead to communicative success. Branigan et al. showed that when interacting with an artificial system, participants aligned their lexical choices more to an artificial system they believed to be basic than to one they believed to be advanced. Our work extends these results in two ways: First, instead of an artificial computer dialog system, participants interact with a humanoid robot, the iCub robot. Second, instead of lexical choice, our work investigates alignment in the domain of manual actions. In an action demonstration and matching game, we examine the extent to which participants who believe that they are playing with a basic version or an advanced version of the iCub robot adapt the way they execute actions to what their robot partner has previously shown to them. Our results confirm that alignment also takes place in action demonstration. We were not able to replicate Branigan et al.'s results in general in this setup, but in line with their findings, participants with a low questionnaire score on neuroticism and participants who are familiar with robots aligned their actions more to a robot they believed to be basic than to one they believed to be advanced.**

I. INTRODUCTION

Interlocutors in a dialog adapt and converge on many aspects of linguistic behavior. Such convergence or *alignment* has been proposed to be the basis for communicative success [1]. Alignment has been shown to take place not only in spoken dialog, but also in written dialog, and non-verbal actions (cf. [2], [3], [4]). In dialog, interlocutors align for example accents, speech rates, phonetic realizations, syntactic structure, speech rhythms, and choice of words. Alignment of choice of words is also known as lexical entrainment [5] (meaning that in a conversation, people tend to use the same words as their dialog partner) and seems to be partner-specific [6]. With regard to non-verbal behavior, interlocutors align their gestures, co-speech gestures, posture, timing of posture changes, body movements (e.g. foot shaking), mannerism, and facial expressions. Research revealed correspondence of gestures in handshape, motion, location, representation, technique, handedness, and palm orientation [7], [8]. Alignment of non-verbal behavior is often referred to as mimicry [9], [10].

With the long-term goal to develop robots that are capable of interacting and of learning in natural social interaction with a human user (while not relying on pre-programmed interaction patterns), we believe that alignment of actions in human-robot interaction could increase the smoothness of cooperative interaction of a robot with a human user and might also facilitate robot action learning from a tutor's demonstration, especially when the user/tutor is naive to the robot's internal functioning and learning mechanisms. So far, related work has investigated alignment in human-robot interaction in speech [11] and qualitatively in gestures ([12], children adapting their behavior to a robot learner's behavior [13]). Alignment towards computers or computational dialog systems has been identified as similar and even stronger alignment than towards human interlocutors [14], [15], [16]. Recent research proposed a mediated *communicative design* account of alignment according to which interlocutors extent of alignment is mediated by their judgment and beliefs about each other ([14], [17], [18]). Pearson et al. conducted a study in which participants played a picture labeling and matching game with an artificial dialog system. The system appeared to be either basic (i.e. old-fashioned and unsophisticated) or advanced (i.e. new and sophisticated). Two pictures were presented to the participants. One of them was textually labeled by the system and the participant had to match this label to the corresponding picture. After that, the participant had to type a label for one of the pictures in a new pair. The system then matched this label to the respective picture. The system's responses were actually scripted. For experimental trials, pictures to be labeled had two different labels; both were equally acceptable, but one was common and the other uncommon. The system systematically gave the uncommon label for these pictures and participants had to label the same picture in a later trial. The results revealed that participants aligned their lexical choices more (i.e. repeated the uncommon label more often) when they believed the artificial system to be basic than when they believed it to be advanced, even though the system's performance was equally good in both cases. This is in line with findings of previous work, suggesting that people will modify their behavior according to their interaction partner's needs and understanding (in human-robot interaction and adult-child interaction [19], [20], [21]). Human alignment in human-robot dialog clearly appears to be beneficial. Consider an example of alignment of choice of words in a conversation with a robotic or artificial interlocutor that has a limited vocabulary: In an office building, the robot asks about the directions to a certain office, "Do I have to continue down this corridor and then turn left?" The human interlocutor now could answer: "No, you need to take the hallway to your right and then turn left." The robot does not know the word 'hallway' or what it means, so it has to ask a follow-up question and communicative success is delayed. Or the human interlocutor aligns to the robot and uses the same word the robot used: "No, you need to take the corridor to your right and then turn left." In this case, the robot knows and understands the word 'corridor' and communicative success is reached right away.

This work focuses on the questions "Do untrained users align their actions to a robot in order to successfully communicate?" and "Do they adapt the way they demonstrate actions to the robot's capabilities and understanding?" Aligning to a robot's *actions* would entail picking up what is already known and understood by the robot. To give a specific example, demonstrating an action "like the robot" by for example restricting one's movements to a certain range, if the robot's possible range of motion is limited, or grasping in a certain manner, for instance with only two fingers, if the robot is equipped with only two finger-like manipulators on its end effector or hand could facilitate mapping of the human's movements onto the robot's own body despite potential differences in human and robot embodiment. Thus, alignment might make understanding and learning easier by lowering the complexity of communication. Our work extends Pearson et al.'s results in two ways: First, instead of an artificial computer dialog system, participants interact with a humanoid robot, the iCub robot. Second, instead of lexical choice, our work investigates alignment in the domain of manual actions. In an action demonstration and matching game, we examine the extent in which participants who believed that they were playing with a basic version or an advanced version of the iCub robot adapt the way they execute actions to what their robot partner has previously shown to them.

II. METHOD

The action demonstration and matching game was designed after Pearson et al.'s picture naming and matching game [17]. A picture in their setup here corresponds to a text box containing the description of an action. Equivalently, the naming of a picture thus corresponds to the demonstration of the respective action described in the text box. Participants were told that they would play a game with a robot in a different room. The robot's responses were really pre-recorded videos of action demonstrations and the robot's choice for matching turns (i.e. the robot's feedback) was scripted choosing always the correct answer in both conditions. We were interested in whether participants would adapt their action demonstrations to the robot's way of demonstrating the action. On experimental trials, the robot demonstrated an action that had one common way and one uncommon way of execution. After two filler trials, participants were asked to demonstrate the same action. We were interested if participants would demonstrate the action in the same way the robot had previously demonstrated it (i.e. if they would align). Analogously to Pearson et al.'s work, with beliefs about an interaction partner's capabilities influencing participants' behavior, we expected alignment stronger in the basic robot condition than in the advanced robot condition.

A. Participants

In all pretests and the experiment, the participants were English native speakers and members of community of Plymouth University. No participant took part in more than one pretest or experiment. The experiment used 31 participants (13 f, 18 m) aged 18–50 years ($M = 23.45, SD = 7.74$) who were paid to participate.

B. Robot

The robot platform with which the experiment was carried out was the iCub humanoid robot [22]. For the experiment, only the robot's upper body was used (head, eyes, torso, arms and hands).

C. Items

We constructed 9 *experimental items* that consisted of the descriptions of a *prime action* and a *distractor action* in text boxes, together with a common and an uncommon way of execution for the prime action, and the descriptions of a *target action* (identical to the prime action) and another distractor action in text boxes. As most actions involved an object, the prime/target action and distractor actions had to involve the same object because the actions should not be recognized by object only. For example, one item comprised the prime action "how to pet a dog" with the distractor action "how to give a dog treats", the common way of execution to pet the dog by doing several long strokes across its back and the uncommon way by giving the dog a couple of pats on its head, the target action and the distractor action "how to pull a dog's tail". Actions that did not involve an object (e.g. "how to wave") were paired with another action that did not involve any objects either. Additionally, we constructed 22 pairs of *filler actions*. One action in each pair was executed, the other was a distractor action. Each of the executed actions was paired with a distractor action, again either involving the same (or no) object as the filler action with which it was paired. A *filler item* consisted of two different, independent pairs of filler actions in text form. One of the described actions in one pair was executed by the robot and one action of the remaining pair was executed by the human participant. One action in each of 11 of the 22 pairs of filler actions was demonstrated by the robot and equivalently one action in each of the other 11 pairs was demonstrated by each participant.

The iCub robot as well as the setup imposed constraints on the actions in the experimental and filler items. As the experiment was set in a sitting position at a table (see Section II-D1), the actions should only involve movement of the upper body which could be carried out on a table top in a sitting position. Additionally, the robot's hardware only permitted actions with light objects, such as soft toys, and a low degree of finicky accuracy. Before the experiment, we conducted three separate pretests to build a dataset of actions and their executions and to find appropriate experimental items. Pretest 1 revealed actions with more than one way of execution that were compatible with the setup of the experiment. In pretest 2, we singled out those actions obtained from pretest 1, which had two equally acceptable ways of execution. In pretest 3, these actions were tested to determine which way of an action's execution participants would favor. From the pretests resulted the actions on which the experimental items were based.

1) Pretest 1: 24 participants (9 f, 15 m) were asked to perform 34 actions (e.g. how to stamp) sitting at a table. They were instructed to give only one unambiguous demonstration for each action description and perform the first one that came to their minds. The action demonstrations were recorded. Of the actions which were demonstrated by participants, we chose those actions which appeared to have more than one type of execution (24 actions) for the next pretest.

2) Pretest 2: 12 participants (6 f, 6 m) were presented with video clips of an actor acting out the ways of execution for the 24 actions obtained from pretest 1 and rated the acceptability of each demonstration for the corresponding action description on a scale from 1 [completely unacceptable] to 7 [completely acceptable]. We then selected 16 of the actions which had two different types of execution both with a rating of more than 5 (*M*=6.18, *SD*=0.61).

3) Pretest 3: In a forced-choice task, 14 participants (7 f, 7 m) named which one of the two demonstrations for each action they would perform. For 10 actions one demonstration was favored by more than 80% of participants (*M*=92.86%, *SD*=5.83%). 9 of these 10 actions constitute the experimental items for the experiment. One item was excluded from the experiment because the robot was not able to perform either way of execution of the action as the objects involved were too small for grasping and the task was too intricate. Summarizing, the actions resulting from the pretests can be demonstrated – also by the robot – in two different ways which are equally acceptable for the task, but of which only one is common and favored and the other is uncommon and unfavored (for an overview, see Table I).

Experimental actions	common way	uncommon way
"how to ring a bell"	with the flat hand	with the thumb
"how to brush (the table)"	long strokes to the side	circular motion
"how to nest cups into each other"	smaller cup into next bigger one	biggest cup into goal cup, smallest into next bigger, then into goal cup
"how to roll dice"	shaking before rolling	direct rolling
"how to pet a dog"	long strokes along back	several pats on head
"how to cut fruit"	one cut	sawing movement
"how to point to the die"	straight	like a pistol
"how to use a saltshaker"	several shakes	tapping with the index finger
"how to swim"	breast stroke	butterfly stroke

TABLE I. EXPERIMENTAL ACTIONS WITH COMMON AND UNCOMMON WAY OF EXECUTION.

4) Robot action demonstration: The experiment required videos of the iCub robot carrying out the 9 actions for the 9 experimental items in two different ways (one common, one uncommon) and the 11 actions of the filler items in a preferably natural manner. For this, we applied slow kinesthetic teaching and replaying of the movements taught using the Aquila framework [23]. This method was not suitable for actions involving large movements of both arms or large shoulder movements. For these actions, joint angles of separate discrete positions were recorded and then interpolated for subsequent replay. Additionally, the robot exhibited a simple gaze behavior: It gazed in direction of the camera at the beginning of each video and action execution, then it gazed down to the object or hand carrying out the action and at the end of the execution, it gazed back to the camera. A front view of the robot's actions was video-recorded and the video clips then sped up partially to compensate for artifacts from the methods used and achieve more natural robot action executions which are more similar to the human action executions observed in pretest 1.

D. Procedure

1) Setup and Course of the Experiment: Participants sat at a table with a monitor on the right in front of them and a camera video-recording their action demonstrations from the front (see Fig. 1). The experimenter was seated behind the monitor so that her face was hidden from the participant. She had a laptop to observe the course of the experiment and to know which objects to get. The participant monitor showed a purpose-built Skype-like application with a big window in which the robot's video was shown and a smaller window in which the participant's video overlaid the lower right corner of the big window (see screenshots in Fig. 2). No parts of the robot's actions were occluded by the participant window. While the windows did not show an action demonstration, they displayed "no video input" to give the impression that the network video connection was running even though no video was shown. Below the windows two text boxes were displayed side by side which contained a pair of action descriptions as part of an experimental or a filler item. The participant and the robot took turns in demonstrating an action and matching it to one of two action descriptions in text boxes. The robot's and participant's videos were only displayed during the action demonstrations and thus robot and participant videos were never shown at the same time. The robot began with an action demonstration turn which was followed by a participant matching turn (see Fig. 2). Then succeeded the participant's demonstrating turn with a subsequent robot matching turn and so on. From the participant's view, the game began with two text boxes appearing beneath the video-windows reading "no video input". The robot's demonstration turn was indicated by the token "A" in the upper left corner of the screen. Then (after a delay of $10s$ for experimental items and a randomly chosen delay of $2s$ to $10s$ for filler items), a countdown from 3 to 1 initiated the transmission of the robot video showing its action demonstration in the big window. After the robot's demonstration, the window went back to the initial "no video input" display and the black border around the text boxes changed to a light gray, prompting the participant to match the robot's demonstration to one of the action descriptions. Choosing an action description was done with the arrow keys and shown with a yellow border around the respective text box and the choice confirmed with the enter key and shown with a red border around the respective text box. All keys of the keyboard other than the ones used at a certain point in time were disabled. In the next step, the participant's demonstration turn began. The token "B" appeared in the corner of the screen and two new text boxes with black borders appeared. One of them was subsequently (after $2s$)

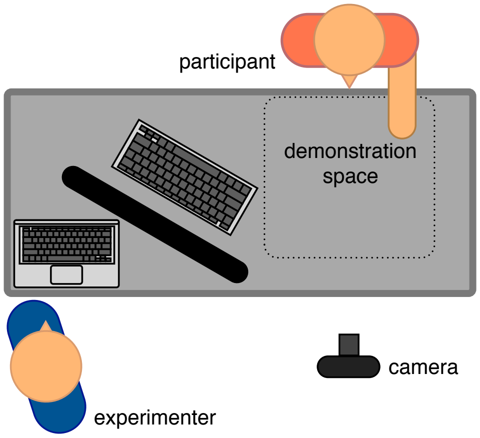

Fig. 1. Experimental setup.

highlighted with a yellow border indicating that the contained action should be demonstrated. If the action involved an object, the experimenter gave it to the participant and when he/she was ready to demonstrate the action, he/she pressed the space bar to start the transmission of the video. Again a countdown was shown, here, in the small participant window, before the video of the webcam was visible. After the action demonstration, the participant pressed the space bar again to stop the video and the window reverted to the "no video input" display. The space bar signal was used not only to show the videos, but also to store them for later analysis. The robot then matched the action demonstration to an action description in the text boxes. After a short delay (randomly chosen between $2s$ to $10s$), its match was shown with a red border around the chosen text box. The game continued with another robot demonstration and participant matching turn and so on until all items had been demonstrated.

The game began with two filler items before the first experimental item was presented. Each experimental item was presented as an *experimental trial* consisting of a robot demonstration turn and participant matching turn for the prime action of the experimental item, followed by two *filler trials*: the participant demonstration and corresponding robot matching turn for a pair of actions of a filler item and robot demonstration and corresponding participant matching turn for the other pair of filler actions of the filler item. Then the target action (same as prime action) was demonstrated by the participant and subsequently matched by the robot. Thus, there was always one participant demonstration turn and one participant matching turn in-between the matching and naming of an experimental action. All 9 experimental trials were presented in succession. At the end of the experiment, participants filled out a questionnaire, which included a small personality questionnaire based on [24] (two items for each factor of the Big Five personality traits [25]; e.g. for neuroticism "I see myself as someone who ...gets nervous easily" and scored in reverse order "is relaxed, handles stress well" on a 5-point Likert scale) and questions about how much experience they had with robots, if they had known the iCub robot before participating, which other robots they knew, what they thought the study was investigating and whether they believed there was a real robot on the other side.

2) Randomizations: Participants were randomly assigned to the basic robot or the advanced robot condition. We constructed two lists of experimental items. In the one list, 4 (in the other list the remaining 5) experimental actions were presented in the uncommon way and 5 (resp. 4) were presented in the common way. Thus, one version of each item appeared in each list. For the filler items we defined a fixed randomized order. The order of the experimental items was randomized individually for each participant. The action text box containing the matching action was displayed on the right or left side for half of the participant matching turns. Accordingly, the text box containing the target action appeared as well on the right or left for half of the participant demonstration turns.

3) Instructions: The experimenter gave verbal instructions to the participants. These comprised two parts. In the first part, the experimenter explained the game and setup. Participants were told that iCub was in a different room and that communication took place over a network connection. They were instructed that they would be referred to as player B and the robot as player A. Then the network application and the course of the game was shown and explained to them in detail including the participant and robot turns and controls for matching an action and starting the video for demonstration. They were told that on the robot screen, their action demonstration was shown in the big window and the robot's demonstration in the small window. The second part consisted of an introduction to the iCub robot according to the condition the respective participant belonged to: *basic* or *advanced*. In the basic condition the experimenter stated that the iCub they would play with was a very basic version and explained that this meant it was restricted in hardware and software and was not equipped with the newest sensors. Additionally, it had not received much training in action recognition or execution. In the advanced condition, the iCub was presented as an advanced version which received all updates and was equipped with the newest sensors. The experimenter told the participants in this condition that the iCub had received much training in action recognition and execution.

4) Scoring: Five raters (3 f, 2 m) were consecutively presented with pairs of videos of all experimental trials. Each pair of videos consisted of the robot's demonstration of the experimental action in one particular experimental trial and the respective demonstration of the same action by the participant in that trial. The order in which the pairs of videos were presented was randomized for each rater individually. The similarity of the two action demonstrations shown in the pairs of videos was rated independently on a scale from 1 [very similar] to 7 [not similar at all]. For the ratings, handedness was not taken into account as a factor for similarity. Raters were told to consider the ways in which each action could be carried out, but were not instructed which factors to take into account. An interrater reliability analysis was performed to determine consistency among the raters. We computed the absolute agreement for the average scores of the raters via a two-way random effects model of intraclass correlation. It revealed a strong agreement: $ICC = 0.881$ with $p < 0.001$, $95\%\ CI(0.857, 0.902)$. The mean of the ratings was taken as a measure for the extent of alignment of the participants to the robot's demonstrations. For the analysis we considered only 8 of the experimental items excluding one experimental item, "how to swim" (common way: breast stroke, uncommon way: butterfly stroke), because during the course of the study, it became evident that most participants were unfamiliar with the butterfly stroke and did not feel confident enough to perform it. For each participant, the ratings for all common experimental items (i.e. actions which were presented by the robot in a common way) and respectively also the ratings for all uncommon experimental items (i.e. actions which were presented by the robot in an uncommon way) were averaged.

III. RESULTS

A two-way mixed ANOVA with robot demonstration condition (common, uncommon) as within-subject factor and the belief about robot capabilities (basic, advanced) as between-subjects factor for the remaining 8 experimental items revealed a statistically significant difference in the extent of alignment based on the robot's way of demonstrating an action (common: $M = 2.39, SD = 0.6$, uncommon: $M = 3.68, SD = 1.04$) , $F(1, 29) = 52.72$, $p < 0.001$; *Wilks'* $\Lambda = 0.355$, *partial* $\eta^2 = 0.645$. Participants rarely demonstrated an action in

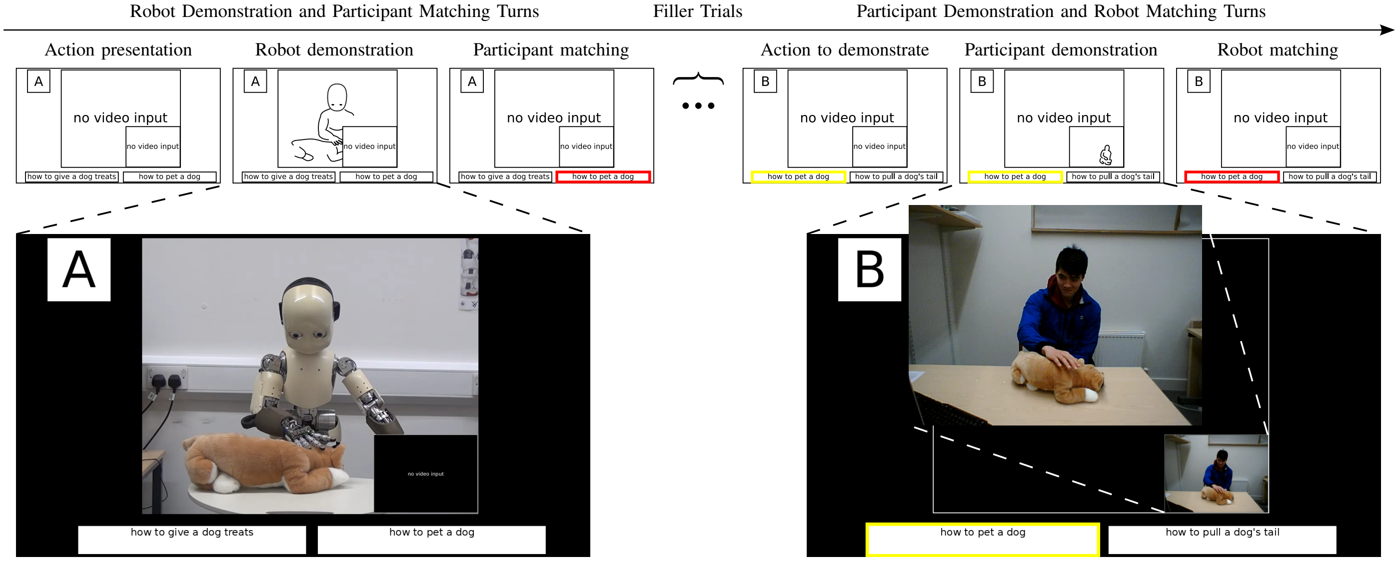

Fig. 2. Schematic view of the course of the experiment showing robot demonstration / participant matching turns and subsequent participant demonstration / robot matching turns with action demonstration screenshots.

a different way, when the robot previously demonstrated the action in the common way (i.e. after a common prime). We expected this main effect, as the common way of demonstrating an action was chosen because it is preferred over the uncommon way by over 80% of people. We were not able to find any significant main effect for a participant's prior beliefs about the robot's capabilities or interaction effect for within- and between-subject factors. However, including the participants' personality traits and their familiarity with robots (i.e. experience with robots, number of robots known, and knew iCub) in the analysis showed significant interaction effects. We split the participants into two groups at the median of factors assessed in the questionnaire (ties were assigned to the higher score except for the participants' experience with robots where the median was 1 corresponding to 'no experience' and the split simply seperated no experience from some experience). We obtained different splits of participants for the Big Five personality traits (openness, conscientiousness, extraversion, agreeableness, neuroticism) according to the Five Factor model [25], for the participants' experience with robots, for the number of robots they knew, and for if they knew iCub before the experiment. Separate two-way mixed ANOVAs for each subset revealed a significant interaction effect for the subsamples of participants with a low score on neuroticism ($Median = 2.5$, on a scale from 1 [lowest value] to 5 [highest value], observed range: 1.5–4.5) and participants that had some experience with robots ($Median = 1$, on a scale from 1 [no experience] to 5 [much experience], observed range: 1–5) and a trend for participants that knew more robots ($Median = 5$, observed range: 0–13), and participants that knew iCub before (for an overview of descriptive statistics and ANOVA results, see Tables II and III). Follow-up one-tailed independent t-tests were conducted to test our hypothesis that scores in the basic condition are lower than in the advanced condition for uncommon robot demonstrations (see Table IV). Results show a trend (low score on neuroticism $p < 0.1$, experienced with robots $p < 0.1$, know more robots $p < 0.1$, and knew iCub $p = 0.12$) that participants in these subsamples were more likely to align their actions to the uncommon way the robot previously demonstrated, when they believed the robot to be basic, then when they believed it to be advanced.

According to the post-experimental questionnaire, no participants had suspicions about the nature of the interaction or knew what the study was investigating. All participants believed they were interacting with the robot in real-time.

Participant subsample	within-subj.	between-subj.	M	SD	N
low score neuroticism	common	basic	2.52	0.57	10
		advanced	2.39	0.5	9
	uncommon	basic	3.45	0.83	10
		advanced	4.04	0.91	9
experienced	common	basic	2.44	0.72	6
		advanced	2.07	0.65	5
	uncommon	basic	3.05	0.68	6
		advanced	3.86	1.1	5
know more robots	common	basic	2.11	0.63	7
		advanced	2.14	0.47	11
	uncommon	basic	3.13	1.01	7
		advanced	3.96	0.99	11
knew iCub	common	basic	2.26	0.68	6
		advanced	2.14	0.54	7
	uncommon	basic	3.17	0.86	6
		advanced	3.89	1.15	7

TABLE II. DESCRIPTIVE STATISTICS FOR SUBSAMPLES OF PARTICIPANTS

Participant subsample	Wilks' Lambda	F	Hyp. df	Error df	Sig.	Partial Eta Squared
low score neuroticism	0.79	4.53	1	17	0.048	0.21
experienced	0.38	14.65	1	9	0.004	0.62
know more robots	0.8	3.94	1	16	0.064	0.2
knew iCub	0.78	3.17	1	11	0.103	0.22

TABLE III. INTERACTION EFFECTS FROM TWO-WAY MIXED ANOVAS FOR SUBSAMPLES OF PARTICIPANTS

Participant subsample	t	df	p
low score neuroticism	-1.48	17	0.0785
experienced	-1.5	9	0.0835
know more robots	-1.716	16	0.0525
knew iCub	-1.26	11	0.1165

TABLE IV. RESULTS FROM INDEPENDENT SAMPLES T-TESTS FOR SUBSAMPLES OF PARTICIPANTS. ONE-TAILED P-VALUES ARE REPORTED.

IV. CONCLUSION

Our experiment found that people seem to also align their non-verbal actions to the ones of a robot interaction partner. We were not able to replicate Branigan et al.'s findings in general for this setup, but people with a less neurotic personality as well as people who are familiar with robots seem to align their actions to a greater extent when they believe the robot to be basic and less capable than when they believe it to be advanced and more capable. Thus, in these subsamples of participants, we were able to verify our hypothesis suggesting that also in the domain of actions, alignment can be mediated by beliefs about the interaction partner. One explanation for our findings could be that a less neurotic (or equivalently an emotionally stable) personality and greater familiarity with robots cause a positive attitude and less tension in the interaction with the robot. Thus, these factors might enable the participant to be less focused on her-/himself and be more responsive to the interaction partner and its needs in the communicative situation and equally to consider how to achieve successful communication. Therefore, participants with a low questionnaire score on neuroticism and participants with a greater familiarity with robots align more strongly to their basic interaction partner they believe to be untrained and less capable of recognizing the actions they show. The opposite subsamples of participants on the other hand might need more feedback from the robot which gives them cues about the robot's understanding to be more reactive and adapt their behavior accordingly. Branigan et al. suggested that the beliefs about a computer as interlocutor may be resistant to dynamic updates and changes induced by the feedback from the computer [14]. In the current study, the robot's feedback consisted of the robot's answers in the robot matching trials which always indicated perfect understanding and additionally, beliefs could possibly have been updated when seeing the robot perform the actions in a sophisticated manner. Investigating this aspect more closely for a robot as interaction partner will be subject of future work. We believe that a thorough understanding of alignment toward robot interaction partners would enable us to develop robots that can elicit these phenomena in interaction and benefit from a reduced interactional complexity for action learning and understanding.

ACKNOWLEDGMENT

This research has been supported by the EU project RobotDoC (235065) from the FP7 Marie Curie Actions ITN.

REFERENCES

[1] M. J. Pickering and S. Garrod, "Alignment as the basis for successful communication," *Research on Language and Computation*, vol. 4, no. 2-3, pp. 203–228, 2006.

[2] T. L. Chartrand, W. W. Maddux, and J. L. Lakin, "Beyond the perception-behavior link: The ubiquitous utility and motivational moderators of nonconscious mimicry," *The new unconscious*, pp. 334–361, 2005.

[3] S. Kopp, "Social resonance and embodied coordination in face-to-face conversation with artificial interlocutors," *Speech Communication*, vol. 52, no. 6, pp. 587–597, 2010.

[4] R. J. Hartsuiker, S. Bernolet, S. Schoonbaert, S. Speybroeck, and D. Vanderelst, "Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue," *Journal of Memory and Language*, vol. 58, no. 2, pp. 214–238, 2008.

[5] S. Garrod and A. Anderson, "Saying what you mean in dialogue: A study in conceptual and semantic co-ordination," *Cognition*, vol. 27, no. 2, pp. 181–218, 1987.

[6] S. E. Brennan and H. H. Clark, "Conceptual pacts and lexical choice in conversation." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 22, no. 6, p. 1482, 1996.

[7] I. Kimbara, "On gestural mimicry," *Gesture*, vol. 6, no. 1, pp. 39–61, 2006.

[8] K. Bergmann and S. Kopp, "Gestural alignment in natural dialogue," in *CogSci*, 2012.

[9] T. L. Chartrand and J. A. Bargh, "The chameleon effect: The perception–behavior link and social interaction." *Journal of personality and social psychology*, vol. 76, no. 6, p. 893, 1999.

[10] J. L. Lakin and T. L. Chartrand, "Using nonconscious behavioral mimicry to create affiliation and rapport," *Psychological Science*, vol. 14, no. 4, pp. 334–339, 2003.

[11] T. Iio, M. Shiomi, K. Shinozawa, T. Miyashita, T. Akimoto, and N. Hagita, "Lexical entrainment in human-robot interaction: can robots entrain human vocabulary?" in *IROS 2009*. IEEE, 2009, pp. 3727–3734.

[12] C. Breazeal, "Regulation and entrainment in human-robot interaction," *The International Journal of Robotics Research*, vol. 21, no. 10-11, pp. 883–902, 2002.

[13] M. Nalin, I. Baroni, I. Kruijff-Korbayová, L. Canamero, M. Lewis, A. Beck, H. Cuayáhuitl, and A. Sanna, "Children's adaptation in multi-session interaction with a humanoid robot," in *RO-MAN, 2012 IEEE*. IEEE, 2012, pp. 351–357.

[14] H. P. Branigan, M. J. Pickering, J. Pearson, J. F. McLean, and A. Brown, "The role of beliefs in lexical alignment: Evidence from dialogs with humans and computers," *Cognition*, vol. 121, no. 1, pp. 41–57, 2011.

[15] S. Stoyanchev and A. Stent, "Lexical and syntactic priming and their impact in deployed spoken dialog systems," in *Proceedings of Human Language Technologies*. Association for Computational Linguistics, 2009, pp. 189–192.

[16] J. Gustafson, A. Larsson, R. Carlson, and K. Hellman, "How do system questions influence lexical choices in user answers," in *Proceedings of EUROSPEECH*, vol. 97. Citeseer, 1997, pp. 2275–2278.

[17] J. Pearson, J. Hu, H. P. Branigan, M. J. Pickering, and C. I. Nass, "Adaptive language behavior in hci: how expectations and beliefs about a system affect users' word choice," in *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 2006, pp. 1177–1180.

[18] H. H. Clark and E. F. Schaefer, "Concealing one's meaning from overhearers," *Journal of memory and Language*, vol. 26, no. 2, pp. 209–225, 1987.

[19] K. Fischer, K. Foth, K. J. Rohlfing, and B. Wrede, "Mindful tutors: Linguistic choice and action demonstration in speech to infants and a simulated robot," *Interaction Studies*, vol. 12, no. 1, pp. 134–161, 2011.

[20] A.-L. Vollmer, K. S. Lohan, K. Fischer, Y. Nagai, K. Pitsch, J. Fritsch, K. J. Rohlfing, and B. Wrede, "People modify their tutoring behavior in robot-directed interaction for action learning," in *ICDL 2009*. IEEE, 2009, pp. 1–6.

[21] A.-L. Vollmer, K. S. Lohan, J. Fritsch, B. Wrede, and K. J. Rohlfing, "Which motionese parameters change with childrens age?" Poster presented at the biennial meeting of the Cognitive Development Society, San Antonio, Texas, USA, 2009.

[22] G. Metta, G. Sandini, D. Vernon, L. Natale, and F. Nori, "The icub humanoid robot: an open platform for research in embodied cognition," in *Proceedings of the 8th workshop on performance metrics for intelligent systems*. ACM, 2008, pp. 50–56.

[23] M. Peniak, A. Morse, C. Larcombe, S. Ramirez-Contla, and A. Cangelosi, "Aquila: An open-source gpu-accelerated toolkit for cognitive robotics research," in *International Joint Conference on Neural Networks (IJCNN)*, 2011.

[24] B. Rammstedt and O. P. John, "Measuring personality in one minute or less: A 10-item short version of the big five inventory in english and german," *Journal of Research in Personality*, vol. 41, no. 1, pp. 203–212, 2007.

[25] P. Costa Jr and R. R. McCrae, "Neo personality inventory–revised (neo-pi-r) and neo five-factor inventory (neo-ffi) professional manual," *Odessa, FL: Psychological Assessment Resources*, 1992.