Visual appearance modulates motor control in social interactions

Stephan de la Rosa \(^{ab,⁎}\), Tobias Meilinger\(^d\), Stephan Streuber\(^d\), Aurelie Saulton\(^{ab}\), Laura Fademrechte\(^{ab}\), M. Alejandra Quiros-Ramirez\(^c\), Heinrich Bülthoff\(^d\), Isabelle Bülthoff\(^d\), Rouwen Cañal-Bruland\(^e\)

\(^a\) Department for Perception Cognition and Action, Max Planck Institute for Biological Cybernetics, Germany
\(^b\) IMPRS for Cognitive and Systems Neuroscience, Eberhard Karls Universität Tübingen, Germany
\(^c\) Department for the Psychology of Human Movement and Sport, Institute of Sport Science, Friedrich Schiller University Jena, Germany
\(^d\) Department of Computer and Information Science, Universität Konstanz, Germany
\(^e\) Department of Psychology, FOM University, Augsburg, Germany

**ABSTRACT**

The goal of new adaptive technologies is to allow humans to interact with technical devices, such as robots, in natural ways akin to human interaction. Essential for achieving this goal, is the understanding of the factors that support natural interaction. Here, we examined whether human motor control is linked to the visual appearance of the interaction partner. Motor control theories consider kinematic-related information but not visual appearance as important for the control of motor movements (Flash & Hogan, 1985; Harris & Wolpert, 1998; Viviani & Terzuolo, 1982). We investigated the sensitivity of motor control to visual appearance during the execution of a social interaction, i.e. a high-five. In a novel mixed reality setup participants executed a high-five with a three-dimensional life-size human- or a robot-looking avatar. Our results demonstrate that movement trajectories and adjustments to perturbations depended on the visual appearance of the avatar despite both avatars carrying out identical movements. Moreover, two well-known motor theories (minimum jerk, two-thirds power law) better predict robot than human interaction trajectories. The dependence of motor control on the human likeness of the interaction partner suggests that different motor control principles might be at work in object and human directed interactions.

1. Introduction

Shaking a friend's hand or picking up a glass for drinking are two examples of humans commonly interacting with their physical and social environment in everyday life. To date it remains largely unknown whether motor actions directed towards the physical (e.g. objects) and social (e.g., other people) environment rely on the same motor principles. On the one hand, researchers have suggested that object and social interactions might rely on common motor principles based on theoretical considerations (Wolpert et al., 2003; Wolpert & Flanagan, 2001). On the other hand, abundant empirical evidence suggests that humans tap into specialized cognitive processes when interacting with their confederates (Becchio et al., 2016; Gallese & Goldman, 1998; Jeannerod, 2001; Sebanz, 2006). For example, theories about social behavior (including motor behavior) stress the importance of cognitive factors in social interactions (Frith & Frith, 2012). Here, a prominent view is that one's own behavior is guided by the assumed mental states of another person (theory of mind) (Lieberman, 2007; Perner & Wimmer, 1985). This is in contrast to motor control theories, which typically explain motor control by sensory, proprioceptive, and kinematic information. It, therefore, remains largely unknown whether motor control differs for social and object interactions.

The few studies comparing social and object interactions showed that social and object directed actions (e.g. imitating a robot vs human or watching hand vs object movements) affect motor control in non-interactive observational scenarios, e.g. eye movements during action observation (Flanagan & Johansson, 2003) or during imitation (Kilner et al., 2003). In real life, however, the interaction with the environment entails the simultaneous observation of the environment and the execution of action. Little is known about motor control in such interactive scenarios, where participants' movements are not restricted. To this end, we examined whether motor control processes differ when humans interact (i.e. high-fived) with a social (i.e. human) and object (i.e. a robot) environment.

We used a novel mixed reality setup, in which participants high-fived a life-size three-dimensional avatar (Fig. 1) without constraining their movement. An optical motion capture system recorded participants' hand movements. This setup allowed us to manipulate the visual appearance of the interaction partner, and thereby the social nature of the interaction, independently of movement. Specifically, participants were interacting either with a human- or robot-looking avatar. Importantly both the human- and the robot-looking avatar were animated...
with an identical pre-recorded high-five movement sequence. As a result, human and robot-looking avatar conditions were identical except for the visual appearance of the interaction partner. We further ensured that human- and robot-looking avatar conditions were also identical with respect to task instructions by painting the human- and robot hand green and asking participants to hit the green area.

We assessed implicit motor behavior by using a task in which participants high-five an avatar. This high-five action required fast and spatio-temporal accurate response of participants and left them little time to employ domain-unspecific cognitive processes for the online control of movements. This task, therefore, minimizes the influence of higher cognitive processes on online motor and increases the chances of measuring implicit motor behavior. Differences in implicit motor behavior have previously been linked to different cognitive states and a-priori biases of human interaction (Becchio et al., 2008; Cavallo et al., 2016; Georgiou et al., 2007). The comparison of human and object interaction in the present study can, therefore, help in the identification of the sources that bias an interaction. This knowledge is useful in minimizing cognitive biases in human-robot interaction and enabling a more natural human-robot interaction.

Prominent motor control theories predict that changing the visual appearance while leaving the task and the movement pattern of the interaction partner unaltered does not affect motor control. This prediction arises from the reliance of motor control theories on information directly relevant to the motor task. For example the minimum jerk theory (Flash & Hogan, 1985; Hogan, 1984) proposes that human hand movements follow a spatial path that minimizes the jerk (i.e. the rate of change in acceleration) while the two-thirds power law (Viviani & Terzuolo, 1982) suggests a close relationship between tangential velocity and instantaneous curvature. Likewise, optimal control theories suggest that motor control is governed by cost functions that penalize higher effort and less accuracy of the movement (Harris & Wolpert, 1998; Körding & Wolpert, 2004; Scott, 2004; Wolpert et al., 2011; Wolpert & Landy, 2012). For example, it has been suggested that motor control processes optimize the end point accuracy (Körding & Wolpert, 2004). The tremendous success of these theories in explaining many object directed actions demonstrates the importance of motor task relevant information for guiding motor behavior. Because we changed the visual appearance of the interaction partner while leaving the movement information untouched, visual appearance does not affect motor task related information. Consequently, motor control theories predict that the visual appearance of the interaction partner (human vs robot) in our tasks should have little to no influence on motor control.

We tested the influence of visual appearance on motor control in an mixed reality experiment where participants interacted with a life-size avatar of different visual appearances in an unconstrained fashion. Specifically, we probed how movement trajectories were affected by the visual appearance of the avatar in two experimental conditions that mimicked two frequent everyday social interaction scenarios: one normal (unperturbed) social interaction condition and one in which the movement of the other person was perturbed. Perturbation of motor movements in everyday life is frequent e.g. when one gets accidentally pushed/shoved during a social interaction.

2. Methods

2.1. Participants

Twenty-eight participants were recruited from the local community in Tübingen. The sample size was determined a-priori based on our previous experience with action related mechanisms (de la Rosa et al., 2014; de la Rosa et al., 2016). In this research we found reliable experimental effects with about 14–16 participants. Fourteen participants were randomly assigned to the robot-looking (8 males; 6 females; 30.13 years) and another fourteen participants were assigned to the human-looking avatar condition (9 males; 5 females; 24.67 years). All participants had normal or corrected to normal vision. Participants gave their informed consent prior to the experiment. The experiment was approved by the local ethics committee of the University of Tübingen and all experiments were performed in accordance with relevant guidelines and regulations.

We used a paper-based questionnaire after the experiment to assess whether the two between-subject groups differed with respect to some basic features. Specifically, as part of this post-experiment assessment we first measured participants’ height and arm length. The questionnaire then contained questions regarding participants’ subjective experience during the experiment. Specifically, we asked how robot-like the movement felt, how human-like the movement felt, and how strong the feeling of interaction was. All these questions were answered on a discrete integer scale from 1 to 7, where 1 meant ‘not at all’ and 7 meant ‘completely’. Finally we also asked participants how many different final hand positions they had perceived, which they indicated by providing a number. We submitted all this information and age to a between-subjects MANOVA in order to determine whether any of the variables differed between the two avatar groups. The MANOVA showed a non-significant main effect of avatar, F(1,26) = 1.302, p = 0.302. Hence, the two between-subject groups did not differ significantly with respect to these measured features.

2.2. Methods

Stimuli were displayed on a large custom build stereo back-projection screen. The projection screen covered a field of view (FOV) of 94.4° horizontally and 78° vertically at a viewing distance of 1 m (projection screen was mounted 26.5 cm above the floor). A ChristieSX + stereoscopic video projector with a resolution of 1400 × 1050 pixels (refresh...
Participants task was to hit the green area of the avatar, which was always the end-effector of the hand. Participants were seated at a distance of 1 m in front of the screen. Participants rested their hand on a small table (see Fig. 1) whose triangular tabletop was located at a height of 99 cm. Participants received the following instructions. The center of the tabletop was marked with a red cross and participants were instructed to rest the center of their palm right in the middle of the cross. Only when participants were within a 5 cm radius of the center of the table for a period of 3 s a trial began. Participants saw a virtual white square hovering in mid-air, which instructed them to place their hand in the middle of the cross. If participants had positioned their hand in the correct position a countdown from 2 s was shown on the white square. During this time the avatar was shown in its resting position (standing with arms down). At time of zero, the avatar started to move. Then participants initiated their hitting movement. The use of a countdown and a predefined starting position were aimed at minimizing the movement variability between different trials, that would be otherwise introduced by variable movement starting positions and movement starting times. If participants initiated their movement earlier than the avatar, the instruction to position their hand onto the center of the table was shown. After hitting the green square, participants moved their hand back to the center of the table. Once there, the next trial started with the display of the countdown. These trials showed a normal (unperturbed) high-five action of the avatar. In contrast, on perturbed trials (offset different from 0 cm), the perturbation started once the participants had lifted her hand above 1.2 m. From this point on, the avatar hand position was presented with a spatial offset (± 0.08 m, ± 0.18 m). All offsets were clearly visible to the participants.

2.4. Design

 Avatar condition was a between subject factor. We chose the factor ‘avatar’ as a between-subject factor to rule out potential carry over effects between factor levels. Specifically, if participants had faced both avatar conditions, they would have been primed to visual appearance being the critical manipulation of the experiment, which itself may have induced demand characteristics (Orne, 1962). That is, participants might have tried to carry over their expected behavioral pattern from one avatar condition to the other one, for example, in order to behave as consistent as possible. For this reason, we decided to give as little information as possible about our main manipulation by choosing a between-subject design. As we had no prior expectations, whether any characteristics of participants (e.g. age) may co-vary with the results, we did not match between-subject groups with regards to specific properties but used random assignment to minimize the probability of systematic sampling bias. The factor offset manipulated the spatial offset that was added to the avatar's hand once the participant had lifted her hand above 1.2 m (± 0.08 m, ± 0.18 m; 0 m). The factor visibility manipulated whether the full avatar (full condition) or only the avatar's hand (partial condition) was shown. Offset and visibility were completely crossed within subject factors. While the testing order of offset was random, visibility was presented in blocked fashion with the testing order counterbalanced across participants. No perturbation (offset = 0 cm) occurred on 50% of the trials (120). On the remaining 120 trials, each of the four perturbation offsets (± 0.08 m, ± 0.18 m) was probed 30 times (12.5% of the trials). Hence, an experimental run consisted of 240 trials. Participants conducted twice 240 trials (once for the full and one for the partial visibility condition).

2.5. Analysis

The onset of the movement was defined as the time when participants had lifted their hand 1 cm above the table. The peak of the participants' movement was defined as the frame before participants started to retract their movement when their hand was above 1.2 m. For assessing the participants’ trajectories we time normalized the trajectory so that the start of the movement corresponded to frame 0 and the peak of the movement to frame 100. To calculate the confidence interval for motion trajectories (Fig. 2) we calculate separate three-way ANOVAs with offset, visibility, and avatar for each time frame separately. We used the error term of the three-way interaction to calculate a between subject 95% confidence interval (95% CI). This error term is used by the ANOVA to compare the individual conditions of all three manipulated factors. We chose this error term because we intend to visually compare the data across all three experimental manipulations (Fig. 2).

Radial error was defined as the spatial distance between the participants’ and avatars’ hand position when both were at their peak frame. The data will be available for download at https://nextcloud.banto.co/index.php/s/HxJ5Ty246KcQND

3. Results

3.1. Does visual appearance affect movement trajectories?

3.1.1. Normal (unperturbed) interactions

We first analyzed the normal high-five trials, in which the hand animation of the avatar was unperturbed. In contrast to the prediction that visual appearance does not influence motor control, we found that the arm movement depended on the visual appearance of the avatar along all spatial axes. Fig. 2A shows the mean trajectories along with the 95% confidence bands, which can be used for visual inspection of significant differences. Non-overlapping confidence bands indicate significant differences between two trajectories. Specifically, robot- and human-looking avatar conditions were markedly different with regard to the forward-backward movement of the hand (z-axis). Here participants exhibit a more pronounced retraction movement of their hand when interacting with the human- than the robot-looking avatar as indicated by the non-overlapping 95% confidence intervals in the rightmost panel of Fig. 2A. This retraction movement occurs shortly before participants move their hands forward to hit the green area of the avatar (i.e. the avatar’s palm).

3.1.2. Perturbed interactions (online motor control)

Another important aspect of motor control is the ability to adjust one’s movement to sudden changes in the target position (online motor
This type of movement control is particularly important in everyday social interaction. In such situations humans often do not know the final position of the interaction partner’s hand prior to or even during the execution of an action and need to adjust their movement to those of the interaction partner ‘on the fly’. To examine whether online motor control processes are affected by visual appearance the avatar adjusted its movement to a new target position that was randomly shifted medially or laterally (i.e. left and right) by 8 or 18 cm as soon as participants had raised their hand by 20 cm from the starting position.

Fig. 3 shows the average frame number when participants first deviated from their unperturbed performance in the perturbation conditions. We submitted this information as a dependent variable to a two-factorial mixed ANOVA with avatar as a between subject factor and offset as a within subject factor. For offsets to the medial side (positive offsets), we found that participants deviated earlier from their average trajectory in the unperturbed condition when the avatar looked like a robot compared to a human (significant interaction between avatar and offset: F (3,78) = 4.81, $\eta^2_{\text{partial}} = 0.16$, $p = 0.004$). Bonferroni-corrected post-hoc independent $t$-tests indicated that the start time between the human- and robot-looking avatar condition was only significantly faster in the 0.08 cm condition (Bonferroni-corrected $\alpha$-level = 0.0125; $-0.18$ cm offset: $t(26) = 0.752, p = 0.459$; $-0.08$ cm offset: $t(26) = -0.794, p = 0.436$; $0.08$ cm offset: $t(26) = 4.021, p > 0.001$; $0.18$ cm offset: $t(26) = 2.315, p = 0.029$). There were no such effects in the lateral perturbation conditions. The absence of this effect for lateral perturbations can be well understood by the visual presentation of the perturbation farther into the visual periphery. The earlier corrective movement in the robot condition could be explained by participants more closely attending object than social movements, thereby allowing for an earlier detection of deviation from the normal avatar movement.

---

**Fig. 2.** Average movement trajectories in (A) the full body avatar conditions and (B) partial body avatar. In both panels the results are shown for each spatial axis (left: x-axis (left-right movement); middle: y-axis (up-down movement); right: z-axis (forward-backward movement)). Gray areas indicate the 95% confidence band.
Support for this idea comes from eye movements studies, which show that saccadic eye movements differ markedly, when objects are moved by a social agent compared to when they move on their own (Flanagan & Johansson, 2003). Specifically, eye movements of passively moving objects follow the object’s movement trajectory more closely than eye movements directed at objects that are moved by a social agent. Overall, the results suggest that visual appearance also influences the human ability to initiate corrective movements.

3.1.3. Control analyses

3.1.3.1. Can different movement trajectories be explained by a speed-accuracy trade-off? Because participants traversed a longer way in space due to this larger retraction movement in the normal (unperturbed conditions), the speed in the human-looking avatar condition (v = 1.70 m/s) was significantly higher than in the robot-looking avatar condition (v = 1.42 m/s), t(20.53) = 3.143, p = 0.005. However, participants did not sacrifice accuracy for speed. Despite the higher speed in the human-looking avatar condition participants were similarly accurate in the human- (radial error = 14.9 cm) and robot-looking avatar (radial error = 16.54 cm) condition to hit the avatar’s hand, t(21.337) = −1.78, p = 0.089. In fact, participants had a tendency to be more accurate in the human-looking avatar condition despite having moved faster in this condition. Bayes Factor analysis showed no evidence in favor of the null hypothesis, r = 0.707, BF = 0.893. Overall, participants did not seem to have a tendency to sacrifice spatial end point accuracy for speed. The findings suggest that visual appearance does influence motor control. Humans interact equally accurate, but faster and with a markedly more pronounced retraction movement of the arm when interacting with a human-compared to a robot-looking avatar.

As for temporal end point accuracy the effect of a speed-accuracy trade-off is less clear. Movement synchrony, measured as the temporal difference between hitting point of the avatar and the participant, was not significantly different between the two avatar conditions, t (26) = 1.220, p = 0.2317. Yet, caution is needed when interpreting this result because there is only very weak evidence in favor of the null hypothesis, Bayes Factors BF = 1.620; r = 0.707. It follows that the temporal end point accuracy does not allow firm conclusions about temporal speed-accuracy trade-offs.

Do the different trajectory patterns in the human and robot-looking avatar conditions merely reflect attentional rather than motor control effects? The human-looking avatar possesses several salient social cues (e.g., the face) that are not co-located with the task relevant part of the body (i.e. the hand) compared to the robot-looking avatar. Participants might have paid more attention to the task relevant body part in the robot condition compared to the human condition in the unperturbed trials. Consequently, they might have moved more directly towards the hand in the robot than human-looking avatar condition causing the retraction movement to be less pronounced in the robot than the avatar condition. If this were the case, then abolishing distracting social cues about the interaction partner should minimize differences between human and robot-looking avatar conditions. We tested this in a control condition, in which we rendered the whole body of the avatar invisible except for the hand. Hence, only the hands were visible while the rest of the avatar’s body was invisible. No other changes were introduced. To this end participants only interacted with the hand of the human and robot-looking avatar in these new conditions. Even when minimizing information about the interaction partner to only the task relevant information (i.e. the hand), participants showed a more pronounced retraction movement in the human compared to the robot-looking avatar condition as indicated by the non-overlapping 95% confidence intervals in the rightmost panel of Fig. 2B. We, therefore, conclude that differences in motion trajectories between the two avatar conditions are not owed to distracting visual cues or attentional effects.

3.1.3.2. Do participants perceive movement related avatar information differently between the human and the robot condition? If this were the case then some of the trajectory differences between the human- and robot-looking avatar condition might be owed to perceptual differences. We probed participants’ subjective impression regarding the avatar’s movement and the feeling of interactiveness (see Methods for more details). Participants filled out a questionnaire at the end of the experiment. They indicated what action they thought the avatar was executing and they rated the human- and robot-likeness of the movement on a 7 point Likert-scale. Moreover, participants also judged how strongly the participants felt that the avatar was interacting with them. Ten participants in the human-looking avatar condition and eleven participants in the robot-looking avatar condition said that the avatar was carrying out a high-five. The results demonstrate that a very
similar number of participants in both conditions perceived the avatar action as a high-five. Fig. 4 shows slight differences in the average ratings regarding the human and robot likeness of the movement. Yet, a mixed ANOVA with avatar condition as between subject factor and appearance as within subject factor indicates no significant main effects (main effect of avatar type: $F(1,26) = 3.64; p = 0.068$; main effect of appearance: $F(1,26) = 2.15; p = 0.154$) and no significant interaction between avatar type and appearance ($F(1,26) = 1.62; p = 0.214$). The interactivity ratings were also not significantly different between the robot and human-looking avatar group as indicated by an independent $t$-test ($t(26) = 0.49, p = 0.632$). Overall, these results suggest that participants in the human and robot-looking avatar condition exhibited different motor behaviors despite similar subjective perceptions of movement-related avatar information.

### 3.2. How well can motor control theories explain the movement trajectories?

Taken together the results show that human and robot interactions are associated with different movement trajectories. The question arises, which of the two movement types are more accurately predicted by motor control theories? To answer this question, we fitted two prominent motor control theories to the unperturbed movement data when the avatar was shown in full, namely the minimum jerk (Flash & Hogan, 1985) and the two-thirds power law model (Viviani & Terzuolo, 1982).

#### 3.2.1. Minimum jerk model

The minimum jerk model has been shown to be very powerful in predicting movement trajectories of many actions. We predicted the average trajectories along each of the three spatial dimensions for each participant and avatar condition with the minimum jerk model (Flash & Hogan, 1985). The overall fit of the minimum jerk model was very good (mean $R^2 = 91.29\%$). Nevertheless, the model provided a better fit for the forward-backward (z-axis) trajectories in the robot- than in the human-looking avatar condition $F(1,26) = 10.27, \eta^2_{\text{partial}} = 0.28, p = 0.004$ (Fig. 5A). No such differences were observed between upward and downward movement (y-axis) as well as left and right movements (x-axis). Therefore, the minimum jerk model predicts robot interaction better than human interaction along the spatial axis for which we observed the largest kinematic differences between robot and human interaction.

#### 3.2.2. Two-thirds power law

We found similar results for the two-thirds power law. The two-thirds power law has been very successful in describing the relationship between velocity and curvature for a wide variety of movements. We fitted the two-thirds power law to the 3D trajectory data and predicted the velocity from the curvature for each avatar condition separately (Fig. 5B). Despite an overall good fit ($R^2 = 61.02\%$), we found the two-thirds power law to better describe the relationship between velocity and curvature for robot than for human interaction, $F(1,26) = 6.56, \eta^2_{\text{partial}} = 0.20, p = 0.017$. Hence, local velocity-curvature relationships seem to be different for human than for robot interaction. Overall, this analysis demonstrates that two prominent motor control theories better account for human movement that is directed towards a robot- than a human-looking avatar. This is in line with the notion that motor control theories have been mainly developed from object directed movements. Our results demonstrate that motor control mechanisms differ significantly for object and human directed movements, hence, motor control theories need to be adjusted.

### 4. Discussion

We show that humans behave differently towards human and robot-looking avatars although both avatar confederates executed identical actions and participants' task instructions were identical. In contrast to predictions of prominent motor control theories, such as minimum jerk theory and the two-thirds power law, these results demonstrate that human motor control depends also on non-kinematic information, namely visual appearance of the interaction partner. Specifically, participants showed a larger retracting hand movement when interacting with a human- than a robot-looking avatar. Although these longer trajectories led to higher movement velocities, the accuracy of hitting the target was comparable (as indicated by radial error and variability) in both interaction conditions. Moreover, participants were earlier able...
to initiate corrective movements for robot than for human interactions. Participant’s perception of movement trajectories and interactivity of the robot- and human-looking avatars did not differ significantly rendering perceptual differences as a cause for the behavioral differences less likely. Taken together these results provide strong evidence that humans interact differently with robot than with human-looking avatars even if the kinematic patterns of the interaction partner give no reason to do so.

We analyzed the trajectories using two prominent motor control theories (Flash & Hogan, 1985; Hogan, 1984; Viviani & Terzuolo, 1982). The analyses showed that motor control theories cannot explain human-directed trajectories to the same degree as robot-directed trajectories. This is somewhat expected as motor control theories are mainly developed in the context of object-directed actions. At the same time, these results point to the necessity to also consider non-object-directed actions in the development and evaluation of these theories. Specifically, human-directed actions could be particularly interesting for motor theories as these actions are an essential and integral part of everyday human life. Consideration of human-directed movement can further increase the ecological relevance of motor control theories.

Overall, the findings demonstrate the existence of differences in motor control between human and robot interaction and that robot interaction is better explained by two well-known motor control theories.

Optimal control theories can account for non-kinematic information more easily in contrast to the above-mentioned motor control theories. The predictions of these theories depend on the parameters that are optimized. In principle, it is possible for optimal control theories to account for at least some of the results by optimizing for parameters that relate movement performance to visual appearance. Our results strongly highlight the necessity to include such parameters in existing optimal control theories if they intend to accurately predict motor control in human interaction.

The reported findings fit well to the observation that eye movement control depends on the presence of human related information. Specifically, humans exhibit reactive eye movement behavior for the observation of objects (e.g. building blocks) in the sense that eye movements closely follow the physical position of the object (Flanagan & Johansson, 2003). Yet, when humans observed social objects (e.g. hand) they showed predictive eye movements where their gaze moved to the anticipated goal position of the hand. Our results provide further

![Fig. 5. Average residual error between predicted and actual trajectory locations of the A. minimum jerk model (shown for each spatial dimension separately) and B. the two-thirds power law. Bars indicate one SE from the mean.](image-url)
support for the idea of different motor control strategies during human interaction. For example, differences between movement trajectories of human- and robot-looking avatar conditions (see Figs. 2 and 3) occur directly after the onset of the movement. These differences early on during the trajectory suggest that participants chose different motor control plans a-priori for robot and human interactions. Yet, our online motor control manipulation showed that these differences in motor control are not restricted to the choice of motor plans but also affect online motor control. It remains an open question to what degree these different motor control strategies are based on different motor control processes.

We can only speculate about the underlying cognitive mechanisms leading to different motor control strategies in human versus robot interaction. One key difference between human and robot interactions might be with regards to the type of agency participants attributed to the human- and robot-looking like avatar. Specifically, participants might have attributed agency to the human looking avatar but not to the robot looking avatar. This could explain why participants retracted the hand more in the human looking avatar condition but not in the robot looking avatar condition. In the robot looking avatar condition participants might have assumed that the movements are pre-programmed, and, hence, not be influenced by the participant’s own movements. However, in the human looking avatar condition, participants might have expected the avatar to react to their own movements and therefore participants might have tried to provide more cues about their own action to the avatar by increasing their own movement trajectory (retracting the hand more). However, if participants attributed different degrees of agency to the interaction partners we would expect differences in the ratings of interactivity in the different conditions. Since this was not the case, we render this possibility as unlikely.

The effect of anthropomorphism on perception has been investigated in several studies to probe sensorimotor contributions to action recognition (for a review see (Press, 2011)). The findings of these studies may partly explain the effect of visual appearance on movement trajectories. The sensorimotor account posits that the observation of an action activates the motor plans corresponding to the observed action. Yet, this activation of the motor system depends on the human-likeness of the observed agent (Press, 2011). According to this view, one would expect that observation of the high-five action of the avatar should activate the motor system in our experiment. Specifically, the robot avatar should activate the motor system involved in producing a high-five less than the human avatar because the robot avatar’s appearance is less human-like. Due to this lower activation of the motor system the motor response in the robot avatar condition might be ‘weaker’ than in the human avatar condition. While this account might be able to predict a difference in motor trajectories between the robot and human avatar condition, this view does not outline the exact movement parameters where these differences should occur and how the mapping from motor activation during action observation is related to motor activation of motor production. One cannot assume that the processes underlying action observation and social interaction have the same properties. We have previously shown that cross-modal transfer between the visual and motor modalities is small to non-existent in conditions which share properties of social interactions while it is not during mere action observation (de la Rosa et al., 2016).

Differences in the human-looking avatar and robot interaction are likely to reflect differences in a-priori cognitive states, e.g. expectations or biases. Our task required a fast and accurate response from participants, which left them little time to use higher cognitive processes during the execution of an action. Consequently, the observed trajectories are likely to partly mirror cognitive states set prior to action execution. Previous research has shown that both cognitive states affect motor behavior and that motor behavior can be used to deduce the cognitive state of the actor (Becchio et al., 2008; Cavallo et al., 2016; Georgiou et al., 2007). Moreover, cognitive states, such as expectations, are considered to be important as they might facilitate an intuitive and easy use of technology. For example, virtual reality allows humans to interact more naturally with a computer generated environment e.g. (de Kok et al., 2017). In robotics the experience of human robot interaction is improved if human expectations can be met during the interaction (see also (Hassenzahl & Tractinsky, 2006)). We suggest that visual appearance is likely to affect human machine interaction through yet unknown cognitive state variables.

One such cognitive state variable might be the negative feeling induced by visual appearance. A long-standing hypothesis in human-machine and human-robot interaction research that visual appearance might have a profound impact on the emotions of the human interaction partner. According to this hypothesis, known as uncanny valley, eerie emotions are elicited in human interaction partners as a robot or an avatar increasingly look like a real human. Although anecdotal evidence seems to match up with this hypothesis, it has been proven difficult to objectively measure and explain this phenomenon (see Wang et al. (2015) for a review). There is some evidence for an uncanny valley effect on human impressions when looking at images of objects that are morphed to look like humans (MacDorman & Ishiguro, 2006; Seyama & Nagayama, 2007). Much less is known about whether the uncanny effect also affects motor control. McMahen et al. (2016), for example, argue in their review that mid fidelity interaction devices often, but not always, lead to worse performance than low fidelity devices. Yet, these results should be interpreted with caution as different interaction devices require different input-output mappings. Performance differences, therefore, might simply reflect the higher cognitive load due to the higher motor complexity required by mid fidelity devices. A fair comparison would be if the same input-output mapping (e.g. natural interaction like a high-five) is used under different appearance conditions. In our study, we kept the interaction method the same while changing the interaction partners’ appearance. Our results show differences in motor control although the human and robot likeness of the avatar’s movement were overall judged quite similar by participants in both conditions. We, therefore, think that uncanny valley effects on motor control in our study is limited and cannot fully explain the observed effects.

Our study is also interesting for the discussion whether human-machine interaction relies on similar cognitive mechanisms as human-human interaction. Though it is clear to participants that robots have a different cognitive system compared to humans, they often apply similar human cognitive processes to computers (for a review see (Nass & Moon, 2000)). It has been suggested that applying social-cognitive mechanisms to human-machine interaction is beneficial for human-robot interaction, too (Krämer et al., 2012). For example, humans rely on gaze cues from humanoid robots to improve task performance in human-robot interactions (Staudte & Crocker, 2011). Our results suggest that this might not hold for all types of human behavior. Specifically, we show that differences occur in implicit motor behavior of the hand. It is likely that these differences are important for social interactions as movement trajectories can be used for reading others’ intentions (Georgiou et al., 2007). Moreover, it has been suggested that human-robot interaction setups therefore might be a useful tool to explore important motion features that social cognitive mechanisms rely on (Sciutti et al., 2015).

The life-size mixed reality setup with unconstrained movements used in this study closely resembles real-life situations. We therefore think that our results are likely to also apply to real-life where humans increasingly interact with robots e.g. in industrial environments. The different movement patterns suggest that well known human machine interaction (HMI) principles might not apply to real-life human and robot interaction situations alike. For example, Fitts’ law predicts that task difficulty (e.g. the size of target – here the avatar’s hand) and performance (e.g. accuracy, hand velocity) are inversely related. The results of our study are not fully in line with this suggestion as we find performance increases for human-compared to robot-looking avatars (e.g. higher velocity) despite constant task difficulty. Hence, our results
stress the importance to examine existing HMI principles under more naturalistic conditions.

In sum, our results show that human motor control is influenced by task irrelevant non-kinemetic information, namely visual appearance. We show that task irrelevant information is able to alter motor control. In addition, motor control in human interaction conditions cannot be readily explained by two prominent motor control theories. These findings suggest that different motor control principles might be at play during social and object interactions.

CRediT authorship contribution statement

**Stephan de la Rosa:** Conceptualization, Methodology, Investigation, Formal analysis, Writing - original draft, Writing - review & editing. **Tobias Meilinger:** Writing - original draft, Writing - review & editing. **Stephan Streuber:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing. **Laura Fademrecht:** Writing - original draft, Writing - review & editing. **Aurélie Saulton:** Writing - original draft, Writing - review & editing. **Isabelle Bülthoff:** Writing - original draft, Writing - review & editing. **Stephan de la Rosa:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing. **M. Alejandra Quiros-Ramírez:** Writing - original draft, Writing - review & editing. **Laura Fademrecht:** Writing - original draft, Writing - review & editing. **Heinrich Bülthoff:** Writing - original draft, Writing - review & editing. **Isabelle Bülthoff:** Writing - original draft, Writing - review & editing. **Rouwen Cañal-Bruland:** Writing - original draft, Writing - review & editing.

Acknowledgments

S.S. was funded by the DFG Centre of Excellence 2117 “Centre for the Advanced Study of Collective Behavior” (ID:422037984).

Appendix A. Supplementary data

Supplementary data to this article can be found online at [https://doi.org/10.1016/j.actpsy.2020.103168](https://doi.org/10.1016/j.actpsy.2020.103168).

References

Becchio, C., Sartori, L., Bulgheroni, M., & Castiello, U. (2008). Both your intention and mine are reflected in the kinematics of my reach-to-grasp movement. *Cognition, 106*, 894–912.

Becchio, C., Sartori, L., & Castiello, U. (2010). Toward you: The social side of actions. *Current Directions in Psychological Science, 19*, 183–188.

Cavallo, A., Koul, A., Ansuini, C., Capozzi, F., & Becchio, C. (2016). Decoding intentions from movement kinematics. *Scientific Reports, 6*(37036).

de Kok, I., et al. (2017). Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics) 10498 LNCS. 105–108.

de la Rosa, S., Ferslt, Y., & Bülthoff, H. H. (2016). Visual adaptation dominates bimodal visual-motor action adaptation. *Scientific Reports, 1–8*. [https://doi.org/10.1038/srep23829](https://doi.org/10.1038/srep23829).

de la Rosa, S., Streuber, S., Giese, M. A., Bülthoff, H. H., & Curió, C. (2014). Putting actions in context: Visual action adaptation aftereffects are modulated by social contexts. *PLoS One, 9*, 1–10.

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation.

Flanagan, J. R., & Hogan, N. (1985). The coordination of arm movements: An experimentally confirmed mathematical model. *The Journal of Neuroscience, 5*, 1688–1703.

Fl用量, C., & Frith, U. (2012). Mechanisms of social cognition. *Annual Review of Psychology, 63*, 267–313.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mindreading. *Trends in Cognitive Sciences, 2*, 493–501.

Georgiou, I., Becchio, C., Glover, S., & Castiello, U. (2007). Different action patterns for cooperative and competitive behaviour. *Cognition, 102*, 415–433.

Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature, 394*, 780–784.

Hassenzahl, M., & Tractinsky, N. (2006). User experience - A research agenda. *Behaviour & Information Technology, 25*, 91–97.

Hogan, N. (1984). An organizing principle for a class of voluntary movements. *The Journal of Neuroscience, 4*, 2745–2754.

Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *Neuronage, 14*, S103–S109.

Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Current Biology, 13*, 522–525.

Körting, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature, 427*, 244–247.

Krämer, N. C., Von Der Pütten, A., & Eimer, M. (2012). Human-agent and human-robot interaction theory: Similarities to and differences from human-human interaction. *Studies in Computational Intelligence, 396*, 215–240.

Lieberman, M. D. (2007). Social cognitive neuroscience: A review of core processes. *Neuroscience*. [https://doi.org/10.1146/annurev.psych.58.110405.085654](https://doi.org/10.1146/annurev.psych.58.110405.085654).

De Dinechin, K. F., & Ishiguro, H. (2006). The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies, 7*, 297–337.

McMahan, R. P., Lai, C. & Pal, S. K. in *59–70* (2016). [https://doi.org/10.1007/978-3-319-39907-2_6](https://doi.org/10.1007/978-3-319-39907-2_6).

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*, 81–103.

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *The American Psychologist, 17*, 776–783.

Perrett, D. I., & Wilmer, H. (1985). “John thinks that Mary thinks that ...” attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology, 39*, 437–471.

Press, C. (2011). Action observation and robotic agents: Learning and anthropomorphism. *Neuroscience and Biobehavioral Reviews, 35*, 1410–1418.

Scialti, A., Ansuini, C., Becchio, C., & Sandini, G. (2015). Investigating the ability to read others’ intentions using humanoid robots. *Frontiers in Psychology, 6*.

Scott, S. H. (2004). Optimal feedback control and the neural basis of volitional motor control. *Nature Reviews. Neuroscience, 5*, 532–546.

Sebanz, N. (2006). The social nature of perception and action. *Psychological Science, 15*, 99–104.

Seymura, J., & Nagayama, R. S. (2007). The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence Teleoperators and Virtual Environments, 16*, 337–351.

Staudte, M., & Crocker, M. W. (2011). Investigating joint attention mechanisms through spoken human-robot interaction. *Cognition, 120*, 268–291.

Srivastava, P., & Tervuulo, C. (1982). Trajectory determines movement dynamics. *Neuroscience, 7*, 431–437.

Wang, S., Lilienfeld, S. O., & Rochat, P. (2015). The uncanny valley: Existence and explanations. *Review of General Psychology, 19*, 403–407.

Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews. Neuroscience, 12*.

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 358*, 593–602.

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology, 11*, 1410–1418.

Wolpert, D. M., & Kawato, M. (1998). Primal representation of movement kinematics. *Science, 280*, 1120–1123.