Great apes’ understanding of biomechanics: eye-tracking experiments using three-dimensional computer-generated animations

Sato, Yutaro; Kitazaki, Michiteru; Itakura, Shoji; Morita, Tomoyo; Sakuraba, Yoko; Tomonaga, Masaki; Hirata, Satoshi

Sato, Yutaro ...[et al]. Great apes’ understanding of biomechanics: eye-tracking experiments using three-dimensional computer-generated animations. Primates 2021, 62(5): 735-747 https://doi.org/10.1007/s10329-021-00932-8

Great apes' understanding of biomechanics: eye-tracking experiments using three-dimensional computer-generated animations

Yutaro Sato[1]
Michiteru Kitazaki[2]
Shoji Itakura[3]
Tomoyo Morita[4]
Yoko Sakuraba[1,5]
Masaki Tomonaga[6]
Satoshi Hirata[1]

1. Wildlife Research Center, Kyoto University, 2-24 Tanakasekiden, Sakyo, Kyoto 6068203, Japan
2. Department of Computer Science and Engineering, Toyohashi University of Technology, 1-1 Hibarigaoka, Tempakucho, Toyohashi, Aichi 441-8580, Japan
3. Center for Baby Science, Doshisha University, 4-1-1 Kizugawadai, Kizugawa, Kyoto 6190225, Japan
4. Institute for Open and Transdisciplinary Research Initiatives, Osaka University, 1-1 Yamadaoka, Suita, Osaka 5650871, Japan
5. Center for Research and Education of Wildlife, Kyoto City Zoo, Okazaki Koen, Okazakihoshojicho, Sakyo, Kyoto 6068333, Japan
6. Inuyama, Aichi 4840004, Japan

Abstract

Visual processing of the body movements of other animals is important for adaptive animal behaviors. It is widely known that animals can distinguish articulated animal movements even when they are just represented by points of light such that only information about biological motion is retained. However, the extent to which nonhuman great apes comprehend the underlying structural and physiological constraints affecting each moving body part, i.e., biomechanics, is still unclear. To address this, we examined the understanding of biomechanics in bonobos (*Pan paniscus*) and chimpanzees (*Pan troglodytes*), following a previous study on humans (*Homo sapiens*). Apes underwent eye tracking while viewing three-dimensional computer-generated (CG) animations of biomechanically possible or impossible elbow movements performed by a human, robot, or nonhuman ape. Overall, apes did not differentiate their gaze between possible and impossible movements of elbows. However, some apes looked at elbows for longer when viewing impossible vs. possible robot movements, which indicates that they may have had knowledge of biomechanics and that this knowledge could be extended to a novel agent. These mixed results make it difficult to draw a firm conclusion regarding the extent to which apes understand biomechanics. We discuss some methodological features that may be responsible for the results, as well as implications for future nonhuman animal studies involving the presentation of CG animations or measurement of gaze behaviors.

Keywords Biomechanics · Body movement · Bonobo · Chimpanzee · Computer-generated animation

Introduction

Visual processing of the movements of other animals is important for adaptive behaviors (e.g., Johansson 1973; Nakayasu and Watanabe 2014; Simion et al. 2008). To assess the perception of biological motion (BM), the kinematics of body movements are often presented as point-light displays (PLDs) (Johansson 1973). Animals can differentiate PLDs representing BM from other control PLDs to some extent (e.g., Blake 1993; De Agrò et al. 2021; Nakayasu and Watanabe 2014; Regolin et al. 2000). Such abilities are observed even in newborn animals such as newly hatched chicks (*Gallus gallus*) (e.g., Regolin et al. 2000; Vallortigara and Regolin 2006; Vallortigara et al. 2005) and human newborns (*Homo *sapiens*) (Bardi et al. 2011, 2014; Bidet-Ildei et al. 2014; Simion et al. 2008). Vallortigara et al. (2005) found that chicks spontaneously approached PLDs displaying the BM of conspecifics (hen), heterospecifics (cat), and vertebrate-like scramble motion, which suggests that chicks have some innate detection mechanism for vertebrate motion. Simion et al. (2008) also found that newborn human babies preferentially looked at heterospecific (hen) BM, which suggests that such innate reactivity may be shared over a wide range of animal taxa.

Furthermore, several laboratory experiments have suggested that nonhuman animals can detect abnormal biomechanical movement to some extent. Computer-generated (CG) animation is particularly useful for displaying impossible movements that cannot be performed by living animals (Campbell et al. 2009; Chouinard-Thuly et al. 2017; Watanabe and Troje 2006). Watanabe and Troje (2006) conducted a study using CG animations of pigeons, and found that pigeons appeared to learn to discriminate between normal conspecifics and conspecifics with abnormal movements (hopping without head-bobbing) by attending particularly to head movements. Similarly, Nakayasu et al. (2017), who used CG fish animations, observed reduced social approach behaviors in medaka (*Oryzias latipes*) when shown conspecifics with atypical locomotion compared with normal conspecifics. These studies used CG animations to successfully uncover new information regarding the understanding of body movements in animals.

However, this line of research has not yet included our closest evolutionary relatives, nonhuman great apes (hereafter referred to as “apes”). Regarding the understanding of body structure, a chimpanzee (*Pan troglodytes*) differentiated a PLD representing BM from a randomized version in a visual search task, although such differentiation was not limited to BM (Tomonaga 2001). Furthermore, Gao and Tomonaga (2020a) examined the effect of image inversion on recognition performance (i.e., the inversion effect). Their findings indicated that chimpanzees used configural processing for images of a normal body, but not for those of a body with scrambled body part locations, which indicates that they may have knowledge regarding the position of each body part [for a similar finding with capuchin monkeys, see Matsuno and Fujita (2018)]. However, the researchers did not directly examine the discrimination of normal from abnormal bodies (Gao and Tomonaga 2020a), and they used static images without motion cues. Thus, little is known regarding how apes understand the unique structural and physiological constraints underlying dynamic body movements, namely, biomechanical restraints. Such information is critical for distinguishing body movement processing between nonhuman animals and humans.

Humans appear to be capable of distinguishing biomechanically possible and impossible movements of particular body parts from early ontogeny (Geangu et al. 2015; Longhi et al. 2015; Morita et al. 2010, 2012; Reid et al. 2005, 2008; Senna et al. 2017). To investigate the underlying developmental process, Morita et al. (2012) carried out a crosssectional study in which human adults and infants were presented with CG animations of biomechanically possible or impossible elbow movements (i.e., bending upward or downward, respectively), performed by either a human or robot. They found that adults and 12-month-old infants, but not 9-month-old infants, looked at the impossible elbow movements for longer than the possible movements in both the human and robot animations. This indicates that both adults and infants have knowledge regarding the biomechanics of human movement, and that they extend this to the movement of nonhuman agents with a similar structure [for similar studies, see Komori et al. (2006) and Morita et al. (2010)]. Additionally, the adults exhibited pupil dilation, which is a physiological indicator of affective arousal, when viewing impossible vs. possible human movements. This could be associated with emotional responses such as surprise, fear, or disgust, as self-reported by another participant group (Morita et al. 2010). In view of this, and in conjunction with the previous finding that observing the movements of others activates sensorimotor representation for the observed movements [e.g., Iacoboni et al. (1999); cf. a recent review by Riečanský and Lamm (2019) on pain and empathy], Morita et al. (2012) concluded that adults identify biomechanical implausibility on the basis of such mirror-like mechanisms as well as knowledge regarding human movements acquired via experience, whereas 12-month-old infants do so primarily on the basis of the latter. This interpretation of the pupillometry data was in accord with other experiments suggesting that the mirroring of somatic components of observed actions might play a role in distinguishing biomechanically possible and impossible movements (Avenanti et al. 2007; Costantini et al. 2005; Romani et al. 2005).

Here, based on the study of Morita et al. (2012), we carried out a series of eye-tracking experiments to examine how apes understand biomechanics. Specifically, like in Morita et al. (2012), ape participants were presented with CG animations of a human and a robot displaying biomechanically possible or impossible arm movements, while participants underwent eye tracking (experiments 1 and 3). We also presented the participants with ape animations displaying similar movements (experiments 2 and 3). We assumed that if the apes looked at the elbow regions for longer when viewing impossible vs. possible movements, then this could indicate that they understood how the elbow should move. That is, their increased attention could reflect a novelty or a violation of predictions, as is usually assumed in similar developmental psychology studies (e.g., Longhi et al. 2015; Morita et al. 2012; Reid et al. 2005).

Pupillometry is known to be challenging in nonhuman animals mainly because of difficulties with head fixation [Massen et al. (2019), but see Hepach et al. (2021) for an attempt at using pupillometry to examine arousal in chimpanzees]. Our pupillometry results may have been affected by the visual properties of the stimuli (see General discussion). Thus, unlike Morita et al. (2012), we mainly report ape gaze behaviors. However, we include a brief summary of our pupillometry results to encourage future implementation of this technique, which may facilitate the examination of ape affective responses associated with autonomic nervous system activity [Kret et al. 2020; for a review on human and monkey studies see Peinkhofer et al. (2019)]. To this end, we provide detailed data visualization in Supplementary file 1 (Figs. S10–S18).

Experiment 1

First, we presented apes with human and robot animations that were slightly modified from those used in Morita et al. (2012). The participant apes lived with conspecific groupmates and had observed human movements via daily interactions with caretakers, researchers, or visitors. If apes have some understanding of human biomechanics, they should look at human elbows for longer when viewing impossible vs. possible movements. Furthermore, if apes have some understanding of biomechanics and, like humans, can infer the biomechanics-like properties of a novel robot agent, then they should also look at the elbows of the robot animation for longer when viewing impossible vs. possible movements. Note that it was unlikely that the apes would make inferences based on the direct experience of seeing similar robotic agents because the apes had rarely seen a robot or watched a three-dimensional CG animation of a robot.

Methods

Participants

Sixteen apes participated in experiment 1: six bonobos (*Pan paniscus*; four females, two males; range 15–46 years old) and six chimpanzees (five females, one male; range 10–23 years old) at Kumamoto Sanctuary, Wildlife Research Center, Kyoto University (KS); and four chimpanzees (one female, three males; range 6–42 years old, including estimated age) at Kyoto City Zoo (KCZ) (Supplementary file 1, Table S1). The experiments depended on voluntary participation, and the apes were not subjected to deprivation of food or water.

Apparatuses

The apes underwent eye tracking using an infrared eye tracker [300 Hz, Tobii TX300, Tobii Technology (KS); 60 Hz, Tobii X120, Tobii Technology (KCZ)]. Movie stimuli were presented on a 23-inch (approx. 58 cm) liquid crystal display monitor with a viewing distance of approximately 70 cm, and sounds were played from speakers. The eye tracker and speakers were mounted under the monitor. The apes viewed the monitor through a transparent panel [1-cm-thick polycarbonate panel (KS); 2-cm-thick acrylic panel (KCZ)]. On each day, anti-fog spray was applied to the panel so that neither the view nor the eye tracker recordings were obscured by condensation created by the breath of the participants. Bonobos and chimpanzees at KS participated in the experiments alone or in a dyad in an indoor enclosure (2.8 × 5.4 × 3.5 m; Fig. 1a), or in an experimental booth (3 × 3 × 2 m), respectively. Chimpanzees at KCZ participated in an indoor enclosure (~ 18.47 m_2 that included separate surrounding rooms for human experimenters; Fig. 1b) that was accessible to all their groupmates during testing. The apes viewed the monitor while sipping juice from a nozzle attached to the panel. Juice was delivered from a custommade juice dispenser hanging on the wall or from a bottle through a straw. In KS, a human experimenter softly held the head of some of the chimpanzees during the session, and the session was conducted without juice for one individual. The juice served to reduce the head motion of participants during the eye tracker recordings. We did not adjust the amount of juice delivered according to a participant’s gaze, so the juice did not serve as a reinforcer like in previous relevant eye-tracking experiments.

Stimuli and procedure

We adjusted the size and duration of the original stimulus materials used in Morita et al. (2012). Specifically, we made 6-s animations (1920 × 1080 pixels) in which either a human or a robot figure repeatedly (1) held both arms forward, (2) bent both elbows, and (3) stretched both arms out every 1.5 s (Fig. 2a, b). These animations are shown in Supplementary files 2–5.

Before the first session for each participant, we performed two-point automated calibration in which we presented participants with a small image at two corners of the screen one by one. We also performed the calibration whenever necessary. To assess the calibration accuracy, at the start of each session, nine images were presented on the screen and an experimenter manually compared the position of the recorded gaze with one of these images. The experimenter occasionally showed a real object by hand in front of the image on the screen to attract the participants' attention to the position of the image. We sought to maintain recording accuracy such that recording errors were typically within 1° (Kano and Tomonaga 2009; Kano et al. 2011; Hirata et al. 2010). Nevertheless, the errors may have been larger at KCZ because the eye tracker angle may have occasionally been different from that during calibration, and the calibration accuracy could not be checked clearly due to the participants' restlessness (in a few cases).

Each trial comprised a 6-s animation, and a session included four trials, such that the participants viewed all four animations in a session (i.e., human and robot animations displaying possible and impossible movements). Each participant underwent two sessions on separate days; data were missing for the second session for one participant at KCZ. The order of the four trials within a session was counterbalanced across participants using Tobii Studio software, which generated the same number of order patterns as there were trials using a Latin square, and assigned one pattern to each participant. The animations were presented to the chimpanzees with accompanying sounds that were timed to accompany the arm movements to attract attention [like stimuli for infants in Morita et al. (2012)]. A 2-s attention catcher (several moving illustrations with sounds) preceded each animation. Sounds were not played to the bonobos because they are known to become nervous upon hearing unfamiliar sounds. The calibration, gaze and pupil diameter recordings, and stimulus presentation were performed using Tobii Studio software v.3.4.8 (or v.3.2.3 for bonobos).

Analyses

Following Morita et al. (2012), we analyzed the data from the first trial of each animation to reduce the potential influence of stimuli habituation. Indeed, we observed reduced gaze time towards the screen from the first to the second presentation, irrespective of the models or movements (see Supplementary file 1, Figs. S1–S3 for the distributions of screen viewing time). The animations in this study were monotonous, as they depicted a repetition of the same arm movements. Thus, the apes may have rapidly habituated to the animations. Some apes also appeared more restless during the second vs. the first presentation. When participants did not view the stimuli for ≥ 1.5 s in the first trial, but did so in the second trial, we used the data from the second trial because it seemed inappropriate to include data from trials in which the apes were not attentive to the stimuli. We decided upon this criterion after considering that each action sequence took 1.5 s to complete (i.e., it took approximately 1.5 s for the position to change from holding both arms forward, bending both elbows, and stretching both arms out again), and that stricter criteria would reduce the sample size to a large extent (see Supplementary file 1, Figs. S1–S3). We also analyzed the first and second trials separately while excluding trials with a screen viewing time of < 1.5 s, and confirmed that the overall results, reported here, matched the analysis of the first trial only (summarized in Supplementary file 1, Table S2, Figs. S4–S6).

Figure 2 shows each area of interest. For each trial, we calculated the proportion of time spent looking at the elbows (i.e., elbow gaze time divided by the total time spent viewing the screen) and compared it between possible and impossible movements separately for the human and robot animations. We conducted statistical analyses using R software v.4.0.2 (R Core Team 2020). We used a Wilcoxon signed-rank test (two-tailed, $\alpha = 0.05$) with the exactRankTests v.0.8.31 R package (Hothorn and Hornik 2019). We used a non-parametric test for the proportion of gaze time because this variable had a heavily skewed distribution. We also calculated Cliff's d as the effect size statistic (Cliff 1993) using the orddom v.3.1 R package (Rogmann 2013). Cliff's d ranges from − 1 to 1, increases in value as more participants exhibit a difference in an expected direction (i.e., a higher proportion of elbow gaze time when viewing impossible vs. possible movements) and vice versa, and takes a value of 0 when there is no such difference. Asymmetric confidence intervals (CIs) were calculated based on the effect size and a variance estimate [Feng and Cliff (1993), cited in Long et al. 2003; Supplementary file 1].

We analyzed pupillometry data from the trials that were used in the gaze behavior analysis described above. We preprocessed the pupillometry data (see Supplementary file 1 for details) using the PupillometryR v.0.0.3 R package (Forbes 2020) while carefully checking that the processing did not yield a large artifact. Figure 3 shows an example of the time series data during each step of processing. We analyzed approximately the first half of each animation because the apes did not usually persist in viewing the stimuli for the entire duration. Moreover, Morita et al. (2012) observed pupil dilation when participants viewed impossible human movement occurring around 2–3 s from animation onset [Morita et al. (2012) also observed pupil dilation for possible robot movements in a later period, but this is of less relevance to the current study]. As it took approximately 1.5 s to bend both elbows (the point at which the biomechanical violation was most apparent during an impossible movement) and stretch both arms out again, approximately the first half of the animations (approx. 3 s) showed this action sequence twice. For each trial, we downsampled the time series data by averaging data points in ten(KS) or two-point (KCZ) time-bins. This enabled us to reduce the dataset size while accommodating differences in the number of data points (KS, 300 Hz; KCZ, 60 Hz). We removed trials with > 15% missing values. For baseline correction, we averaged time-bin values corresponding to approximately 300–500 ms from the animation onset and used this as the baseline. We then subtracted it from the following time-bin values. We finally averaged these baseline-corrected values in time-bins corresponding to approximately 2–3 s from an animation's onset. We refer to this output as the "pupil dilation score," with a larger score indicating increased pupil diameter from baseline, and vice versa. The pupil dilation scores were compared between possible and impossible movements for the human and robot animations separately using paired *t*-tests.

Results

For the human animations, two individuals were removed from the analysis because their data were missing for either or both conditions (possible/impossible movements). We used data from the second trial instead of the first trial for one case (impossible movement). The screen viewing time did not significantly differ between the possible [median (interquartile range) 3.65 (2.94–4.17)] and impossible movements [4.27 (3.08–5.49)] (Wilcoxon signed-rank test: $n = 14$, $V = 75$, $p = 0.173$). Figure 4a shows the proportion of time spent looking at elbows. The proportion of elbow gaze time was not significantly different across the possible [0.01 (0 − 0.08)] vs. impossible movements [0.12 (0.05–0.27)] [$n = 14$, $V = 73$, $p = 0.057$, Cliff's d = 0.50, 95% CI = (− 0.02, 0.81)]. For the robot animations, two individuals were removed from the analysis for the same reason described above. We used data from the second trial for one case (possible movement). The screen viewing time did not significantly differ between the possible [4.25 (3.32–5.91)] and impossible movements [4.69 (3.24–5.50)] (Wilcoxon signed-rank test: $n = 14$, $V = 59$, $p = 0.715$). Again, the proportion of elbow gaze time was not significantly different across the possible [0.07 (0.03 − 0.10)] vs. impossible movements [0.10 (0.01–0.13)] [$n = 14$, $V = 54$, $p = 0.588$, Cliff's d = 0.21, 95% CI = (− 0.29, 0.62)]. Regarding pupillometry, nine and ten participants were removed from the analyses for human and robot animations, respectively, because their data were insufficient. The pupil dilation score was not significantly different between the possible and impossible movements in both the human ($t_6 = 0.79$, $p = 0.460$) and robot animations ($t_5 = 0.31$, $p = 0.772$; for results, see Supplementary file 1, Fig. S7).

Discussion

These results did not clearly suggest that the apes understood the biomechanics underlying the arm movements. Although the apes seemed to look at the human elbows for a longer period of time when viewing the impossible movement compared with the possible movement, the difference was not statistically significant. In this experiment, the model was either a heterospecific (human) or unfamiliar agent (robot); we did not include a conspecific agent. Conspecific body images, rather than human body images, may elicit enhanced processing of body structure or action cues (e.g., Gao et al. 2020; Gao and Tomonaga 2020b; Hattori et al. 2010; Kano and Call 2014). Therefore, in experiment 2, we further examined ape understanding of biomechanics using animations of conspecific-like apes.

Experiment 2

Using the same apparatus as experiment 1, we presented elbow movement in nonhuman great ape animations. We predicted that apes may have a better understanding of conspecific compared with robot and human biomechanics, and hence that they would look at ape elbows for longer when viewing impossible vs. possible movements.

Methods

Fifteen apes (the same participants as in experiment 1, except one fewer ape from KCZ) participated in experiment 2. We used a procedure that was identical to that in experiment 1, except that we used a nonhuman ape animation, created based on a chimpanzee figure, which displayed similar possible and impossible elbow movements (Fig. 2c; Supplementary files 6, 7). A session comprised two trials (i.e., presentation of a 6-s animation) so that the participants viewed both animations (i.e., possible and impossible movements) in a session. They underwent two sessions on separate days; data for the second session were missing for two participants at KCZ.

Results

One individual was removed from the analysis because their data were missing. We analyzed the data only from the first trials. The screen viewing time was not significantly different between possible [5.36 (4.93–5.81)] and impossible movements [4.78 (3.61–5.40)] (Wilcoxon signed-rank test: $n = 14$, $V = 30$, $p = 0.173$). Figure 4b shows the proportion of time spent looking at elbows. The proportion of elbow gaze time was not significantly different between possible [0.07 (0.02–0.14)] and impossible movements [0.11 (0–0.19)] [$n = 14$, $V = 47$, $p = 0.946$, Cliff's d = 0.07, 95% CI = (− 0.41, 0.52)]. Regarding pupillometry, nine participants were also removed from the analysis because their data were insufficient. The pupil dilation score (i.e., pupil diameter change from baseline, averaged approx. 2–3 s from animation onset) was not significantly different between possible and impossible movements ($t_4 = 1.58$, $p = 0.190$; for results, see Supplementary file 1, Fig. S8).

Discussion

Again, the results did not suggest that the apes understood the biomechanics of the movements shown in the ape animations. However, the apes may have perceived the animations as strange or uncanny because of differences between their appearance (texture, color, etc.) and movements (trajectory, speed, acceleration, etc.) and those of real apes [cf. uncanny valley (Mori 1970; Mori et al. 2012)], and hence did not focus on the elbow movements. In a previous study (Steckenfinger and Ghazanfar 2009), for example, long-tailed macaques (*Macaca fascicularis*) differentiated their gaze between realistic CG monkey images compared with real monkey images and unrealistic CG monkey images. Thus, we attempted to reduce the unfamiliar nature of the nonhuman ape animations in experiment 3.

Experiment 3

We generated a modified ape animation in which arm movements were made while the animated ape stood quadrupedally. We hypothesized that this posture might be more familiar to apes, and thus, that the revised animations could enable their enhanced processing of the biomechanics compared with the animations of the bipedally standing apes used in experiment 2 [cf. for the effect of posture familiarity on the inversion effect, see Gao and Tomonaga (2020b)]. We presented the animations of quadrupedally standing apes as well as the human and robot animations (bipedally standing as in experiment 1) to a new group of chimpanzees. As with the participants in experiment 1, the chimpanzees in this experiment were familiar with humans, but not with robots. We predicted that the apes would show clearer signs that they detected biomechanical violations in the ape animations, as indicated by more time spent looking at the elbows when viewing impossible vs. possible movements.

Methods

Ten chimpanzees (seven female, three males; range 19–54 years old, including estimated age) at the Primate Research Institute, Kyoto University (PRI), participated in experiment 3 (Supplementary file 1, Table S1). We presented an ape animation in which the ape bent its right elbow either inward or outward while standing quadrupedally (Fig. 2d; Supplementary files 8, 9). We also presented the human and robot animations from experiment 1, in which the elbows were bent while the human or robot stood bipedally. A session comprised six trials, such that the participants viewed all six animations in a session (i.e., ape, human, and robot animations displaying possible and impossible movements). Two sessions were completed on separate days. We did not include sound in the animations, as it seemed to have a minimal impact in the previous experiments. A 4-s attention-getter (two small photographs of chimpanzees presented sequentially with brown noise) preceded each animation. We used a different attention-getter in experiment 3 with the goal of an enhanced control for the baseline pupil diameter. Chimpanzees at PRI participated individually in an indoor experimental booth (1.8 × 2.15 × 1.75 m) and viewed the monitor through a transparent 2-cm-thick acrylic panel. The other procedures were identical to those in experiment 1.

Results

For the ape animations, no individuals were removed from the analysis, and data from only the first trials were analyzed. The screen viewing time was not significantly different between the possible [5.58 (5.04–5.69)] and impossible movements [5.59 (4.28–5.97)] (Wilcoxon signed-rank test: $n = 10$, $V = 27$, $p = 1$). Figure 5 shows the proportion of time spent looking at elbows. The proportion of elbow gaze time was not significantly different between the possible [0 (0–0.06)] and impossible movements [0 (0–0.07)] [$n = 10$, $V = 9$, $p = 0.844$, Cliff's d = 0, 95% CI = (− 0.45, 0.45)]. For the human animations, one individual was removed from the analysis because their data for one condition were missing. We analyzed data from the second trials for three cases (two for possible and one for impossible movements). Screen viewing time was not significantly different between the possible [5.89 (4.24–5.94)] and impossible movements [5.65 (3.43–5.78)] (Wilcoxon signed-rank test: $n = 9$, $V = 15$, $p = 0.426$). The proportion of elbow gaze time was not significantly different between the possible [0.14 (0.03–0.17)] and impossible movements [0.12 (0.06–0.29)] [$n = 9$, $V = 30$, $p = 0.426$, Cliff's d = 0.33, 95% CI = (− 0.32, 0.77)]. For the robot animation, no individuals were removed from the analysis, and only data from the first trials were analyzed. The screen viewing time did not significantly differ between the possible [4.33 (3.72–5.86)] and impossible movements [5.62 (3.99–5.99)] (Wilcoxon signed-rank test: $n = 10$, $V = 35$, $p = 0.492$). Intriguingly, the proportion of elbow gaze time was significantly higher when the apes viewed impossible [0.17 (0.12–0.26)] vs. possible movements [0.08 (0.04–0.12)] [$n = 10$, $V = 52$, $p = 0.010$, Cliff's d = 0.80, 95% CI = (0.15, 0.97)]. Regarding pupillometry, three, two, and one participants were further removed from the analysis for ape, human, and robot animations, respectively, because their data were insufficient. The pupil dilation score (i.e., pupil diameter change from baseline, averaged approx. 2–3 s from animation onset) was not significantly different between the possible and impossible movements for all animations (ape, $t_6 = 0.38$, $p = 0.719$; human, $t_6 = -0.77$, $p = 0.470$; robot, $t_8 = 0.89$, $p = 0.397$).

Discussion

The chimpanzees in experiment 3 did not differentiate gaze behaviors between possible and impossible human or ape movements. Thus, these results did not suggest that the chimpanzees understood the biomechanics underlying human and ape movements. In contrast, the chimpanzees looked at elbows for longer when viewing impossible vs. possible robot movements. Because the chimpanzees were not expected to be familiar with robotic agents, it is unlikely that they differentiated gaze behaviors based on past experience. Instead, this result may indicate that, like humans (Morita et al. 2012), chimpanzees can extend some knowledge regarding biomechanics to a novel robot agent. This interpretation is apparently at odds with the finding that the chimpanzees did not distinguish between possible and impossible movements in the human and ape animations. This discrepancy may in part be due to the uncanny nature of the CG animations. Specifically, the chimpanzees may have perceived the ape and human animations as uncanny because, while those agents looked like familiar animals, either conspecific or heterospecific, they were dissimilar in some way, such as in their surface texture. Consequently, the apes may have just visually explored the human and ape agents per se, rather than fixating on the elbows. In contrast, the robot with its object-like appearance was not likely perceived as uncanny, and so it may have been easier for the chimpanzees to attend to its movements rather than the agent per se. This view is supported particularly for the ape animation by the result that four of ten participants did not fixate on the ape elbows when viewing both possible and impossible movements, while participants did look at the elbows when viewing at least one of the possible or impossible movements in the human and robot animations.

Meta-analysis

Overall, the apes did not look at the elbows for a longer period of time when viewing impossible vs. possible movements, except in the case of the chimpanzees in experiment 3 when viewing the robot animations. Despite the similar stimuli and procedures, experiments 1 and 3 yielded different results for the robot animation. Moreover, the apes in experiment 1 seemed to be able to distinguish the human possible and impossible movements (although this was not statistically significant; see Fig. 4a), while this did not seem to be the case for the apes participating in experiment 3 (Fig. 5). To synthesize the results of the human and robot animations in experiments 1 and 3, we conducted a meta-analysis. Note that this was not possible for the ape animations because different animations were used in experiments 2 and 3 (i.e., standing bipedally/ quadrupedally).

We used Cliff's d and a variance estimate for each animation (human and robot) in each experiment (experiments 1 and 3), and ran random effect models (restricted maximum likelihood estimation) using the metaphor v.2.4.0 R package (Viechtbauer 2010). We used an unweighted Cliff's d, which did not give a different weight to each experiment, because it can perform better than a weighted one (Hess et al. 2005; Kromrey et al. 2005). We calculated CIs for the pooled effect using the estimated effect and its variance (Supplementary file 1), and inferred its significance based on whether the CIs included 0. In performing the analysis, we referred to Kim et al. (2018), and adopted the associated R codes for calculation and generation of the forest plot from Errington et al. (2018) and Iorns et al. (2019).

Figure 6 shows Cliff's d with the corresponding CI for each of experiments 1 and 3, as well as for the random effect model. CIs of the pooled effect included 0 for both the human [estimated effect: 0.42, 95% CI = (− 0.02, 0.72)] and robot animations [estimated effect: 0.51, 95% CI = (− 0.15, 0.85)]. This indicates that the difference in gaze behavior between possible and impossible movements was not significant when the two experiments were combined.

General discussion

In the present study, we examined the way in which bonobos and chimpanzees understand biomechanics using a series of eye-tracking experiments based on a previous study in human adults and infants (Morita et al. 2012). Specifically, we presented CG animations adapted from the human study to apes, and measured their gaze while they viewed the animations via eye-tracking technology. Morita et al. (2012) found that 12-month-old infants and adults looked at elbows for longer when they were shown biomechanically impossible arm movements compared with possible movements. This pattern was observed both when the movements were displayed by a human and when displayed by a robot. In experiment 1, apes were shown human and robot animations adapted from Morita et al. (2012). Although we found no clear differences in gaze behaviors between possible and impossible movements in either the human or robot animations, the apes showed slight signs that they could distinguish these movements in the human animations. In experiment 2, we presented bipedally standing ape animations. Yet again, the apes did not differentiate their gaze behaviors between possible and impossible movements. Finally, in experiment 3, we continued the line of inquiry from the prior experiments by presenting another group of chimpanzees with modified ape animations in which the apes stood quadrupedally, and hence had a more typical posture, as well as the human and robot animations used in experiment 1. The chimpanzees did not differentiate their gaze behaviors between possible and impossible movements in the human and ape animations. However, they did look at elbows for longer when viewing impossible vs. possible robot movements, as if they had some understanding of robot biomechanics-like properties. The reason why the apes in experiment 3 differentiated their gaze behaviors, while those in experiment 1 did not, is unclear. In the present study, apes were not particularly required or motivated to look carefully at the animations. Thus, some apes might not have specifically fixated on the elbows during the trials because of inattention or restlessness, although we set the minimum inclusion criterion based on screen viewing time. Relatedly, the chimpanzees in experiment 3 seemed to view the stimuli for slightly longer than the apes in experiment 1 (see Supplementary file 1, Figs. S1, S3), which indicates that different levels of attention paid to stimuli may have affected the results. Alternatively, the chimpanzees in experiment 3 (at PRI) might have had more general experience of observing artificial objects compared with the apes in experiment 1 (at KS and KCZ), which could have indirectly helped the former to distinguish between possible and impossible movements in the robot animations.

Given these mixed results, it is difficult to infer the extent to which apes understand biomechanics. These results could indicate that, unlike humans (e.g., Morita et al. 2012), apes have little or no understanding of biomechanics. However, it is equally plausible that the null results are entirely related to the methodological features of the present study, in particular the CG animations of arm movements. First, the uncanny nature of the CG animations may have superseded the saliency of the impossible movements. Thus, future studies should carefully assess ape perception of CG animations, which could be perceived as strange looking by apes, prior to testing (Chouinard-Thuly et al. 2017). Presentation of a familiarization video prior to the presentation of test stimuli (Reid et al. 2005) or a habituation-dishabituation technique [for related human infant studies, see Christie and Slaughter (2010) and Southgate et al. (2008)] might also help familiarize participants with the experimental stimuli. Second, while we used animations depicting arm movements, animations of other body parts (e.g., faces, hands, or legs) might be more suitable for presentation to apes. For example, human infants spend a lot of time looking at their hands (White et al. 1964) and can regulate their hand movements voluntarily so that their hand enters their visual field (van der Meer 1997; van der Meer et al. 1995). When toddlers manipulate objects, they often watch their hands or those of a social partner (Yoshida and Smith 2008). Thus, hands appear to be especially salient for humans (Geangu et al. 2015; Longhi et al. 2015; Senna et al. 2017), and possible vs. impossible hand movements can be discerned at a young age [at 6 but not at 4 months old (Geangu et al. 2015); several days after birth (Longhi et al. 2015); at 9 months old, but only in those who could perform a pincer grip (Senna et al. 2017)]. This discrimination occurred slightly later in Morita et al. (2012), who examined reactions to elbow movements [i.e., at 12 but not 9 months old; but for a different result, see Morita et al. (2010)]. Compared with elbow movements, hand movements might be more familiar and salient for apes as well. Third, while our stimuli depicted a purposeless, intransitive arm-bending movement, other types of movements (e.g., walking or climbing) or goal-directed actions (e.g., grasping an object) coupled with more extreme examples of biomechanical violation might have elicited a different response from the apes. For example, Reid et al. (2005) reported that 8-month-old infants differentiated biologically possible and impossible arm movements displayed during a reaching action, although this was only the case for infants with certain fine motor skills. Goal-directed actions may be more salient [for a study with human neonates, see Craighero et al. (2011)], and thus serve to elicit differentiated gaze behaviors between conditions. Furthermore, an early group of experiments yielded roughly converging findings indicating that the mirror-like motor resonance system in humans may be more or less activated by the observation of both intransitive movements and goal-directed actions, whereas such a system in macaque monkeys may only be activated by goal-directed actions, or perhaps only very subtly by the observation of intransitive movements (reviewed by Rizzolatti et al. 2014; Rizzolatti and Craighero 2004; Rizzolatti and Sinigaglia 2010). More recently, this view has been challenged by data indicating that so-called mirror neurons in monkeys are also, to some extent, activated via the observation of intransitive movements (Kraskov et al. 2009; Papadourakis and Raos 2017, 2019). Nevertheless, it is possible to speculate that nonhuman and human animals might differ in terms of reactivity to intransitive movements relative to that to goal-directed actions. Although further research is needed, this possibility is well in line with the overall null results of the present study. Therefore, movies depicting movement of more salient body parts such as the hands (Geangu et al. 2015; Longhi et al. 2015; Senna et al. 2017), or goal-directed actions (Geangu et al. 2015; Reid et al. 2005), might facilitate research regarding how apes perceive violations of biomechanics.

We do not expect the recording accuracy of the ape eye-tracking technology to have a profound impact on the results. We did not quantitatively measure the error and could not compare it with that in the previous study on humans (Morita et al. 2012). However, we closely followed procedures that have been established through previous ape eye-tracking studies (reviewed in Hopper et al. 2020). Experiments have shown that these procedures can enable reliable recording of ape gaze with recording errors of less than 1 ° (Kano and Tomonaga 2009; Kano et al. 2011; Hirata et al. 2010). Although there are also some differences in procedures between those studies and the present one (e.g., the use of juice), Kano and Tomonaga (2009), for example, reported comparable levels of error between chimpanzee and human participants of less than 0.5 °. Kano et al. (2011) also noted that the error was about 0.5–0.7 ° for all participant groups, including nonhuman and human apes. Notably, some of the chimpanzees from those studies also participated in the present study, namely, chimpanzees at KS and PRI. The error might be larger for several participants, in particular chimpanzees at KCZ, as we noted above (Stimuli and procedure in Methods). Thus, our results should be interpreted with caution. However, those participants accounted for a small proportion of the participant group in each experiment, and we even omitted some of their data from statistical analyses because they spent an insufficient amount of time viewing the screen. Thus, it is unlikely that the data from those individuals had a pronounced impact on the overall results. The relatively small size of areas of interest for elbows might affect the nuanced results of experiment 1. Relatedly, a potential limitation of the present study is that we did not consider in depth the role played by peripheral vision (Orquin and Holmqvist 2018). Stimuli depicting a body at closer range might be useful for the clearer differentiation of gazes directed at different body parts. However, our findings suggest that the arm movements of our stimuli were sufficiently salient to cause the apes to orient their gaze accordingly. In experiment 3, the chimpanzees looked at the robot elbows for longer when viewing impossible vs. possible movements. Remarkably, the screen viewing time was not significantly different between those conditions. Thus, the eye-tracking experiments were able to capture finely differentiated gaze behaviors. This underlines the chief advantage of non-invasive ape eye tracking, namely, a prominent spatiotemporal resolution that is difficult to achieve by manual coding of gaze or head orientation (Hopper et al. 2020; Kano and Call 2017).

Overall, we did not observe pupil dilation when the apes viewed impossible movements. The present study used realistic, graphically rich animations; however, we did not control for the effects of specific visual properties within the animation such as brightness and contrast, which could modulate pupil diameter (Bradley et al. 2017). As these could hamper the detection of pupil dilation associated with affective responses, the pupillometry results should be interpreted with caution. Nonetheless, at the very least, our study suggests that the measurement of ape pupil diameter is feasible, and we believe that detailed data visualization (Supplementary file 1, Figs. S10–S18) might be instrumental for future studies.

In conclusion, the findings of the present study, in which apes underwent eye tracking whilst watching CG animations, add to those of previous work on primates' visual processing of movements made by other animals. Specifically, we collected data on ape behavior via a similar method to that used in human developmental psychology research (Morita et al. 2012). Some of our results indicate that some apes may have an understanding of biomechanics that they can draw on to infer the movements of a novel robotic agent. However, the null hypothesis regarding gaze behavior and pupil response was not rejected. Thus, conclusive evidence regarding the extent to which apes understand biomechanics was not obtained in the current study. Nevertheless, the results of this study have implications for the design of nonhuman animal studies involving gaze behavior, pupillometry, and/or CG animations.

Acknowledgments We would like to especially thank Mr. H. Ishimoto and Mr. K. Nagamachi at the Department of Computer Science and Engineering, Toyohashi University of Technology, for making the highquality 3D animations. Further, we are grateful to Dr. N. Morimura, Dr. F. Kano, and all the staff at KS; Ms. H. Bando, Mr. M. Sato, Dr. Y. Yamanashi, and Dr. M. Tanaka at KCZ; and Ms. E. Ichino, Dr. Y. Kawaguchi, and members of the Language & Intelligence section of the PRI, for their kind help in conducting the current study. We also thank Dr. T. Matsuzawa for financial support. We thank Sydney Koke, MFA, from Edanz (https://jp.edanz.com/ac), for editing a draft of this manuscript. Finally, we are grateful to the editor and anonymous reviewers for their constructive feedback during the review process.

Author contributions Conceptualization: YS, MK, SI, SH. Methodology: MK. Formal analysis, investigation, data curation, writing—original draft, and visualization: YS. Resources: MK, TM, YS, MT, SH. Writing—review and editing: MK, SI, TM, YS, MT, SH. Supervision: SI, SH. Project administration: SH. Funding acquisition: SH, YS.

Funding This study was supported financially by the Japan Society for the Promotion of Science, Ministry of Education, Culture, Sports, Science and Technology, to YS (grant no. 19J22889), MT (grant nos. 15H05709, 18H04194), SH (grant nos. 26245069, 18H05524), and TM (grant no. 16H06283, LDG-U04); the Cooperative Research Program of the PRI; and the Great Ape Information Network. These funding sources had no involvement in the study design, collection, analysis and interpretation of data, manuscript writing, or in the decision to submit the article.

Availability of data and material The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request. Note that relevant data are visualized in this published article (and in Supplementary file 1).

Code availability The R code used during the current study is available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethics approval The animal husbandry and research protocols complied with the Guide for Animal Research Ethics of the Wildlife Research Center, Kyoto University for the KS experiments (chimpanzees, no. WRC-2018-KS006A; bonobos, no. WRC-2018-KS007A), the Research Ethics Guideline of KCZ for the KCZ experiments (no. 2019-KCZ-001), and the Guidelines for the Care and Use of Laboratory Primates of the PRI for the PRI experiments (no. 2019-001).

Consent to participate Not applicable.

Consent for publication All authors approved the submission of the article. The funding sources had no involvement in the decision to submit this article for submission.

References

Avenanti A, Bolognini N, Maravita A, Aglioti SM (2007) Somatic and motor components of action simulation. Curr Biol 17:2129–2135. https://doi.org/10.1016/j.cub.2007.11.045

Bardi L, Regolin L, Simion F (2011) Biological motion preference in humans at birth: role of dynamic and configural properties. Dev Sci 14:353–359. https://doi.org/10.1111/j.1467-7687.2010. 00985.x

Bardi L, Regolin L, Simion F (2014) The first time ever I saw your feet: inversion effect in newborns' sensitivity to biological motion. Dev Psychol 50:986–993. https://doi.org/10.1037/a0034678

Bidet-Ildei C, Kitromilides E, Orliaguet J-P et al (2014) Preference for point-light human biological motion in newborns: contribution of translational displacement. Dev Psychol 50:113–120. https://doi. org/10.1037/a0032956

Blake R (1993) Cats perceive biological motion. Psychol Sci 4:54–57. https://doi.org/10.1111/j.1467-9280.1993.tb00557.x

Bradley MM, Sapigao RG, Lang PJ (2017) Sympathetic ANS modulation of pupil diameter in emotional scene perception: effects of hedonic content, brightness, and contrast. Psychophysiology 54:1419–1435. https://doi.org/10.1111/psyp.12890

Campbell MW, Carter JD, Proctor D et al (2009) Computer animations stimulate contagious yawning in chimpanzees. Proc R Soc B: Biol Sci 276:4255–4259. https://doi.org/10.1098/rspb.2009.1087

Chouinard-Thuly L, Gierszewski S, Rosenthal GG et al (2017) Technical and conceptual considerations for using animated stimuli in studies of animal behavior. Curr Zool 63:5–19. https://doi.org/ 10.1093/cz/zow104

Christie T, Slaughter V (2010) Movement contributes to infants' recognition of the human form. Cognition 114:329–337. https://doi. org/10.1016/j.cognition.2009.10.004

Cliff N (1993) Dominance statistics: ordinal analyses to answer ordinal questions. Psychol Bull 114:494–509. https://doi.org/10.1037/ 0033-2909.114.3.494

Costantini M, Galati G, Ferretti A et al (2005) Neural systems underlying observation of humanly impossible movements: an fMRI study. Cereb Cortex 15:1761–1767. https://doi.org/10.1093/cer- cor/bhi053

Craighero L, Leo I, Umiltà C, Simion F (2011) Newborns' preference for goal-directed actions. Cognition 120:26–32. https://doi.org/ 10.1016/j.cognition.2011.02.011

De Agrò M, Rößler DC, Kim K, Shamble PS (2021) Perception of biological motion in point-light displays by jumping spiders. bioRxiv. https://doi.org/10.1101/2021.02.24.432515

Errington TM, Tan FE, Lomax J et al (2018) Replication studies. Retrieved from osf.io/p7ayb

Feng D, Cliff N (1995) Comparisons of power and size of *d* and *t* statistics under normal and nonnormal assumptions. University of Southern California, Los Angeles (**Unpublished manuscript**)

Forbes SH (2020) PupillometryR: an R package for preparing and analysing pupillometry data. J Open Source Software 5:2285. https:// doi.org/10.21105/joss.02285

Gao J, Tomonaga M (2020a) Body perception in chimpanzees (*Pan troglodytes*): the effect of body structure changes. J Comp Psychol 134:222–231. https://doi.org/10.1037/com0000214

Gao J, Tomonaga M (2020b) How chimpanzees and children perceive other species' bodies: comparing the expert effect. Dev Sci. https://doi.org/10.1111/desc.12975

Gao J, Kawakami F, Tomonaga M (2020) Body perception in chimpanzees and humans: the expert effect. Sci Rep 10:7148. https:// doi.org/10.1038/s41598-020-63876-x

Geangu E, Senna I, Croci E, Turati C (2015) The effect of biomechanical properties of motion on infants' perception of goal-directed grasping actions. J Exp Child Psychol 129:55–67. https://doi.org/ 10.1016/j.jecp.2014.08.005

Hattori Y, Kano F, Tomonaga M (2010) Differential sensitivity to conspecific and allospecific cues in chimpanzees and humans: a comparative eye-tracking study. Biol Let 6:610–613. https://doi. org/10.1098/rsbl.2010.0120

Hepach R, Vaish A, Kano F, Albiach-Serrano A, Benziad L, Call J, Tomasello M (2021) Chimpanzees' (Pan troglodytes) internal arousal remains elevated if they cannot themselves help a conspecific. J Comp Psychol 135:196–207

Hess MR, Kromrey JD, Ferron JM et al (2005) Robust inference in meta-analysis: an empirical comparison of point and interval estimates using the standardized mean difference and cliff's delta. Annual Meeting of the American Educational Research Association. Montreal, p 11−15

Hirata S, Fuwa K, Sugama K et al (2010) Facial perception of conspecifics: chimpanzees (*Pan troglodytes*) preferentially attend to proper orientation and open eyes. Anim Cogn 13:679–688. https:// doi.org/10.1007/s10071-010-0316-y

Hopper LM, Gulli RA, Howard LH et al (2020) The application of noninvasive, restraint-free eye-tracking methods for use with nonhuman primates. Behav Res. https://doi.org/10.3758/ s13428-020-01465-6

Hothorn T, Hornik K (2019) exactRankTests: exact distribution for rank and permutation tests. R package version 0.8–31. https:// CRAN.R-project.org/package=exactRankTests

Iacoboni M, Woods RP, Brass M et al (1999) Cortical mechanisms of human imitation. Science 286:2526–2528. https://doi.org/10. 1126/science.286.5449.2526

Iorns E, Tsui R, Denis A et al (2019) Meta-analysis. Retrieved from osf.io/c69jx

Johansson G (1973) Visual perception of biological motion and a model for its analysis. Percept Psychophys 14:201–211. https:// doi.org/10.3758/BF03212378

Kano F, Call J (2014) Cross-species variation in gaze following and conspecific preference among great apes, human infants and adults. Anim Behav 91:137–150. https://doi.org/10.1016/j.anbeh av.2014.03.011

Kano F, Call J (2017) Great ape social attention. In: Watanabe S, Hofman M, Shimizu T (eds) Evolution of the brain, cognition, and emotion in vertebrates. Springer, Tokyo, pp 187–206. https://doi. org/10.1007/978-4-431-56559-8_9

Kano F, Tomonaga M (2009) How chimpanzees look at pictures: a comparative eye-tracking study. Proc R Soc B: Biol Sci 276:1949– 1955. https://doi.org/10.1098/rspb.2008.1811

Kano F, Hirata S, Call J, Tomonaga M (2011) The visual strategy specific to humans among hominids: a study using the gap–overlap paradigm. Vis Res 51:2348–2355. https://doi.org/10.1016/J. VISRES.2011.09.006

Kim J, Afshari A, Sengupta R et al (2018) Replication study: melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. Elife 7:e39944. https:// doi.org/10.7554/eLife.39944

Komori N, Katayama J, Kitazaki M, Itakura S (2006) Do infants prefer possible human movements? The 5th International Conference of the Cognitive Science. Vancouver, pp 143–144

Kraskov A, Dancause N, Quallo MM et al (2009) Corticospinal neurons in macaque ventral premotor cortex with mirror properties: a potential mechanism for action suppression? Neuron 64:922–930. https://doi.org/10.1016/j.neuron.2009.12.010

Kret ME, Prochazkova E, Sterck EHM, Clay Z (2020) Emotional expressions in human and non-human great apes. Neurosci Biobehav Rev. https://doi.org/10.1016/j.neubiorev.2020.01.027

Kromrey JD, Hogarty KY, Ferron JM et al (2005) Robustness in metaanalysis: an empirical comparison of point and interval estimates of standardized mean differences and Cliff's delta. Joint Statistical Meetings. Minneapolis, pp 7−11

Long JD, Feng D, Cliff N (2003) Ordinal analysis of behavioral data. In: Schinka JA, Velicer WF (eds) Handbook of psychology: research methods in psychology, vol 2. Wiley, Hoboken, NJ, pp 635–661. https://doi.org/10.1002/0471264385.wei0225

Longhi E, Senna I, Bolognini N et al (2015) Discrimination of biomechanically possible and impossible hand movements at birth. Child Dev 86:632–641. https://doi.org/10.1111/cdev.12329

Massen JJM, Behrens F, Martin JS et al (2019) A comparative approach to affect and cooperation. Neurosci Biobehav Rev 107:370–387. https://doi.org/10.1016/j.neubiorev.2019.09.027

Matsuno T, Fujita K (2018) Body inversion effect in monkeys. PLoS ONE 13:e0204353. https://doi.org/10.1371/journal.pone.0204353

Mori M (1970) Bukimi no tani (the uncanny valley). Energy 7:33–35

Mori M, MacDorman KF, Kageki N (2012) The uncanny valley (from the field). IEEE Robot Autom Magaz 19:98–100. https://doi.org/ 10.1109/MRA.2012.2192811

Morita T, Katayama N, Kitazaki M, Itakura S (2010) Development of perception of human and robot body movement. J Robot Soc Japan 28:463–469. https://doi.org/10.7210/jrsj.28.463

Morita T, Slaughter V, Katayama N et al (2012) Infant and adult perceptions of possible and impossible body movements: an eyetracking study. J Exp Child Psychol 113:401–414. https://doi.org/ 10.1016/j.jecp.2012.07.003

Nakayasu T, Watanabe E (2014) Biological motion stimuli are attractive to medaka fish. Anim Cogn 17:559–575. https://doi.org/10. 1007/s10071-013-0687-y

Nakayasu T, Yasugi M, Shiraishi S et al (2017) Three-dimensional computer graphic animations for studying social approach behaviour in medaka fish: effects of systematic manipulation of morphological and motion cues. PLoS ONE 12:e0175059. https://doi. org/10.1371/journal.pone.0175059

Orquin JL, Holmqvist K (2018) Threats to the validity of eye-movement research in psychology. Behav Res 50:1645–1656. https:// doi.org/10.3758/s13428-017-0998-z

Papadourakis V, Raos V (2017) Evidence for the representation of movement kinematics in the discharge of F5 mirror neurons during the observation of transitive and intransitive actions. J Neurophysiol 118:3215–3229. https://doi.org/10.1152/jn.00816.2016

Papadourakis V, Raos V (2019) Neurons in the macaque dorsal premotor cortex respond to execution and observation of actions. Cereb Cortex 29:4223–4237. https://doi.org/10.1093/cercor/bhy304

Peinkhofer C, Knudsen GM, Moretti R, Kondziella D (2019) Cortical modulation of pupillary function: systematic review. PeerJ 7:e6882. https://doi.org/10.7717/peerj.6882

R Core Team (2020) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Regolin L, Tommasi L, Vallortigara G (2000) Visual perception of biological motion in newly hatched chicks as revealed by an imprinting procedure. Anim Cogn 3:53–60. https://doi.org/10.1007/s1007 10050050

Reid V, Belsky J, Johnson M (2005) Infant perception of human action: toward a developmental cognitive neuroscience of individual differences. Cogn Brain Behav 9:193–210

Reid VM, Hoehl S, Landt J, Striano T (2008) Human infants dissociate structural and dynamic information in biological motion: evidence from neural systems. Soc Cogn Affect Neurosci 3:161–167. https://doi.org/10.1093/scan/nsn008

Riečanský I, Lamm C (2019) The role of sensorimotor processes in pain empathy. Brain Topogr 32:965–976. https://doi.org/10.1007/ s10548-019-00738-4

Rizzolatti G, Craighero L (2004) The mirror-neuron system. Annu Rev Neurosci 27:169–192. https://doi.org/10.1146/annurev.neuro.27. 070203.144230

Rizzolatti G, Sinigaglia C (2010) The functional role of the parietofrontal mirror circuit: interpretations and misinterpretations. Nat Rev Neurosci 11:264–274. https://doi.org/10.1038/nrn2805

Rizzolatti G, Cattaneo L, Fabbri-Destro M, Rozzi S (2014) Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. Physiol Rev 94:655–706. https://doi.org/10.1152/physrev.00009.2013

Rogmann JJ (2013) orddom: ordinal dominance statistics. R package version 3.1. https://CRAN.R-project.org/package=orddom

Romani M, Cesari P, Urgesi C et al (2005) Motor facilitation of the human cortico-spinal system during observation of bio-mechanically impossible movements. Neuroimage 26:755–763. https:// doi.org/10.1016/j.neuroimage.2005.02.027

Senna I, Addabbo M, Bolognini N et al (2017) Infants' visual recognition of pincer grip emerges between 9 and 12 months of age. Infancy 22:389–402. https://doi.org/10.1111/infa.12163

Simion F, Regolin L, Bulf H (2008) A predisposition for biological motion in the newborn baby. Proc Natl Acad Sci 105:809–813. https://doi.org/10.1073/pnas.0707021105

Southgate V, Johnson MH, Csibra G (2008) Infants attribute goals even to biomechanically impossible actions. Cognition 107:1059–1069 Steckenfinger SA, Ghazanfar AA (2009) Monkey visual behavior falls into the uncanny valley. Proc Natl Acad Sci 106:18362–18366. https://doi.org/10.1073/pnas.0910063106

Tomonaga M (2001) Visual search for biological motion patterns in chimpanzees (*Pan troglodytes*). Psychologia 644:46–59

Vallortigara G, Regolin L (2006) Gravity bias in the interpretation of biological motion by inexperienced chicks. Curr Biol 16:R279– R280. https://doi.org/10.1016/j.cub.2006.03.052

Vallortigara G, Regolin L, Marconato F (2005) Visually inexperienced chicks exhibit spontaneous preference for biological motion patterns. PLoS Biol 3:e208. https://doi.org/10.1371/journal.pbio. 0030208

van der Meer AL (1997) Keeping the arm in the limelight: advanced visual control of arm movements in neonates. Eur J Paediatr Neurol 1:103–108. https://doi.org/10.1016/S1090-3798(97)80040-2

van der Meer AL, van der Weel FR, Lee DN (1995) The functional significance of arm movements in neonates. Science 267:693–695. https://doi.org/10.1126/science.7839147

Viechtbauer W (2010) Conducting meta-analyses in R with the metafor package. J Stat Soft 36:1–48. https://doi.org/10.18637/jss.v036. i03

Watanabe S, Troje NF (2006) Towards a “virtual pigeon”: a new technique for investigating avian social perception. Anim Cogn 9:271–279. https://doi.org/10.1007/s10071-006-0048-1

White B, Castle P, Held R (1964) Observations on the development of visually-directed reaching. Child Dev 35:349–364. https://doi. org/10.2307/1126701

Yoshida H, Smith LB (2008) What’s in view for toddlers? Using a head camera to study visual experience. Infancy 13:229–248. https:// doi.org/10.1080/15250000802004437

Figures

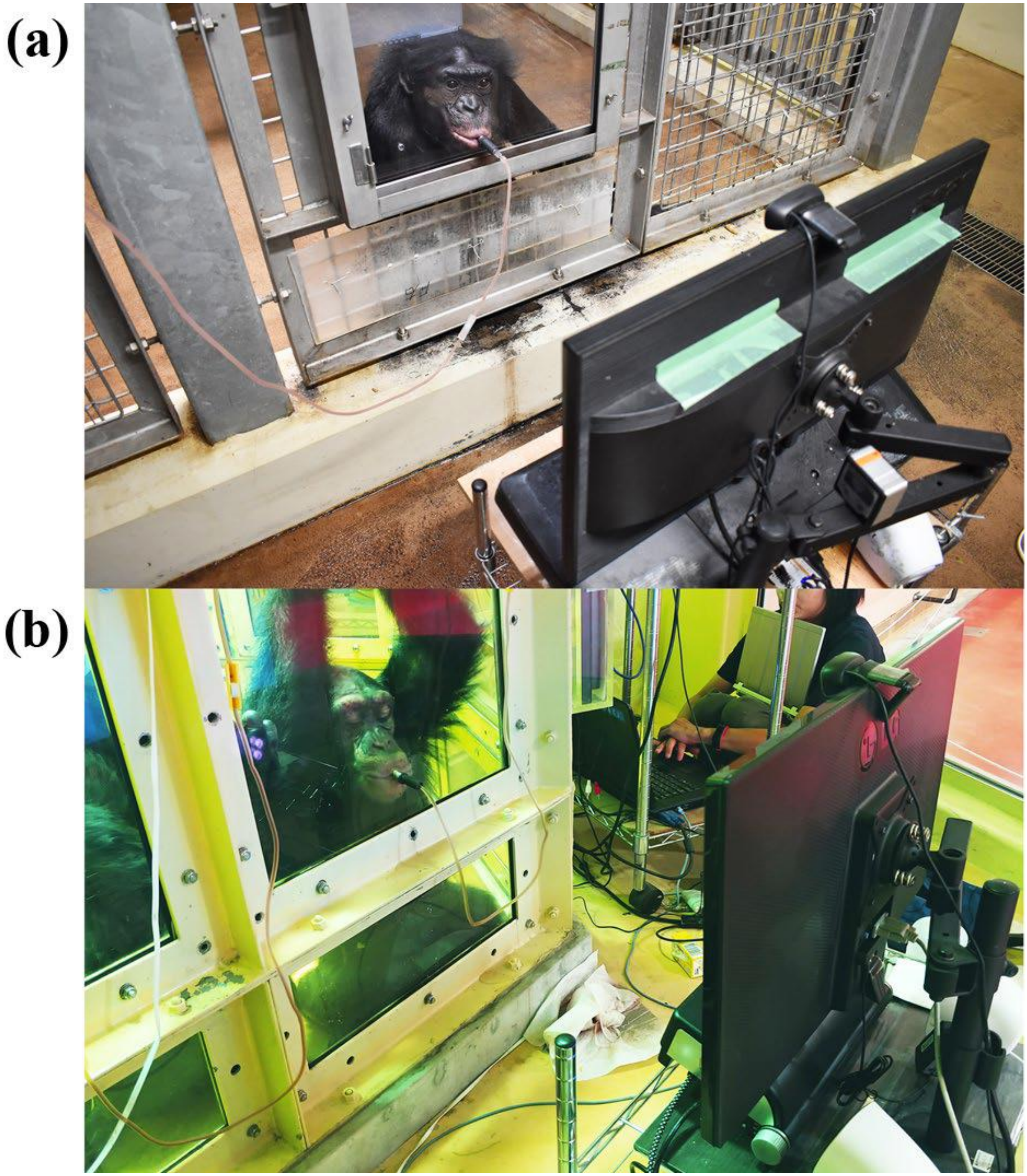

Fig. 1 Experimental settings at Kumamoto Sanctuary (a) and at Kyoto City Zoo (b)

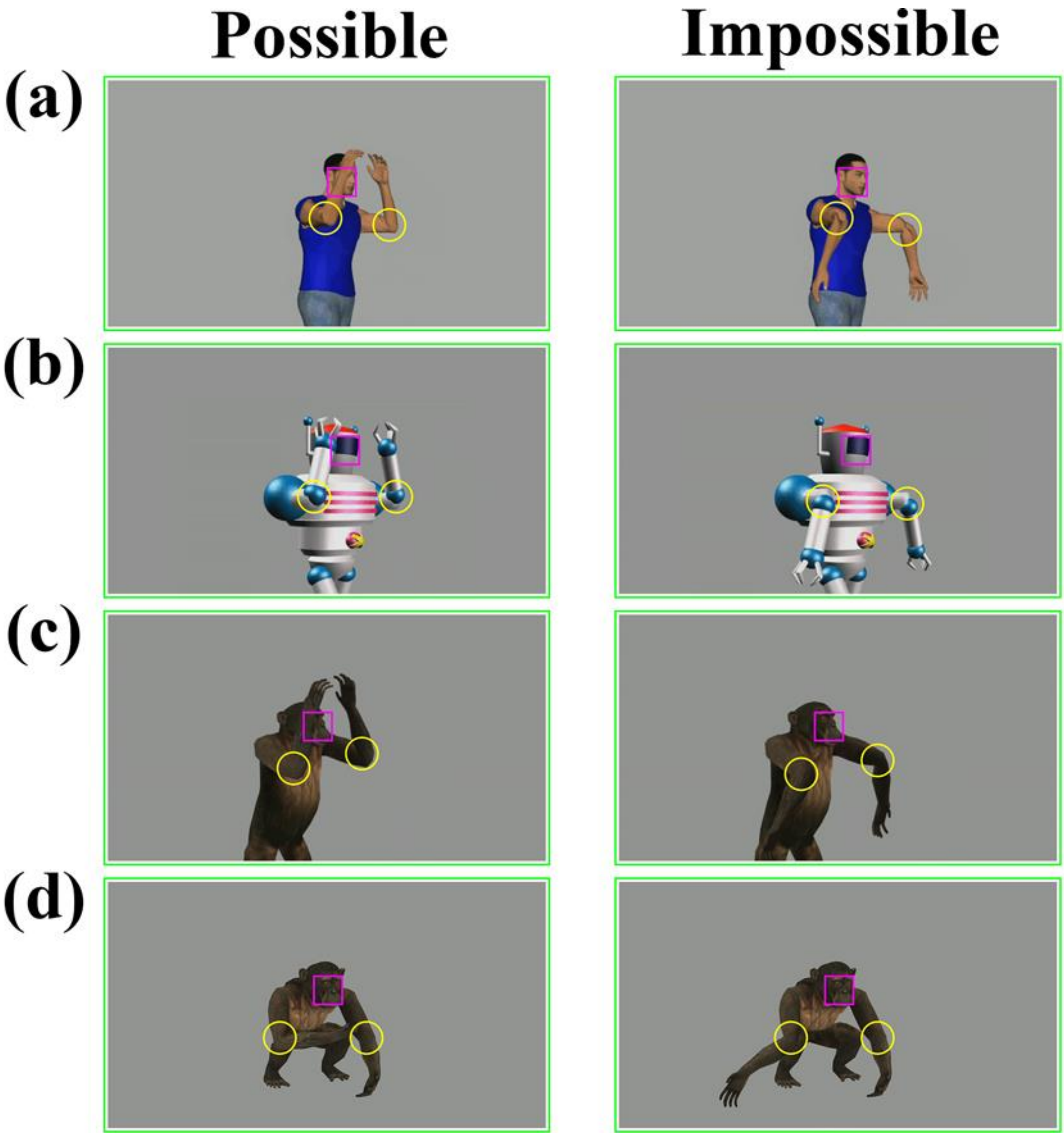

Fig. 2 Representative stills of animation stimuli: **a** human, **b** robot, **c** biped ape, and **d** quadruped ape. *Left* Possible elbow movements, *right* impossible elbow movements. *Circles* Areas of interest (AOI) for elbows, squares AOI for face (analysis of gaze time directed at face was not reported), *rectangles* AOIs for screen. These AOIs were not included in the stimulus animations. In the human, robot, and biped ape animations, the position of the whole body on the screen moved according to the elbow movements, unlike in Morita et al. (2012). Thus, the AOIs for elbows and face moved to track the corresponding body parts throughout the animation

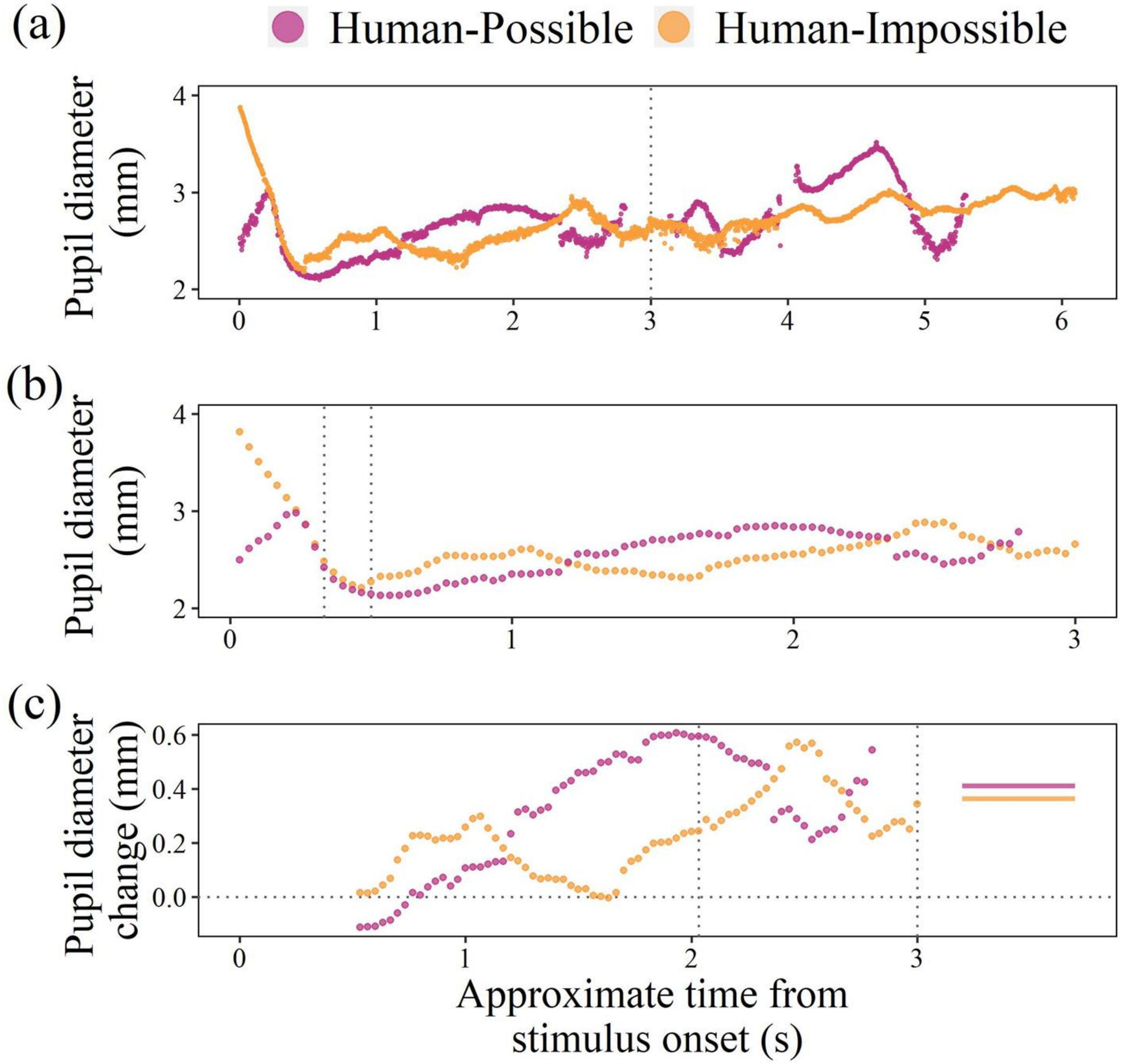

Fig. 3 **a** Example of pupillometry data from one individual when viewing the human animations. For visual comparison, two time series datasets were plotted simultaneously. Wefirst averaged pupil diameters across the right and left eyes. Approximately the first half of the data, corresponding to the *left* of the *vertical dotted line*, were analyzed in the following step. **b** We downsampled the time series data by averaging the data points in time-bins. The *area between the vertical dotted lines* (approx. 300–500 ms from animation onset) served as the baseline in the following step. **c** The values were then baseline corrected. The horizontal lines show the averaged values approximately 2–3 s from the stimulus onset (referred to as the “pupil dilation score” in the main text), which were subsequently analyzed using paired t-tests. Note that the axis scales differ between the plots. Similar plots for each individual are provided in Supplementary file 1

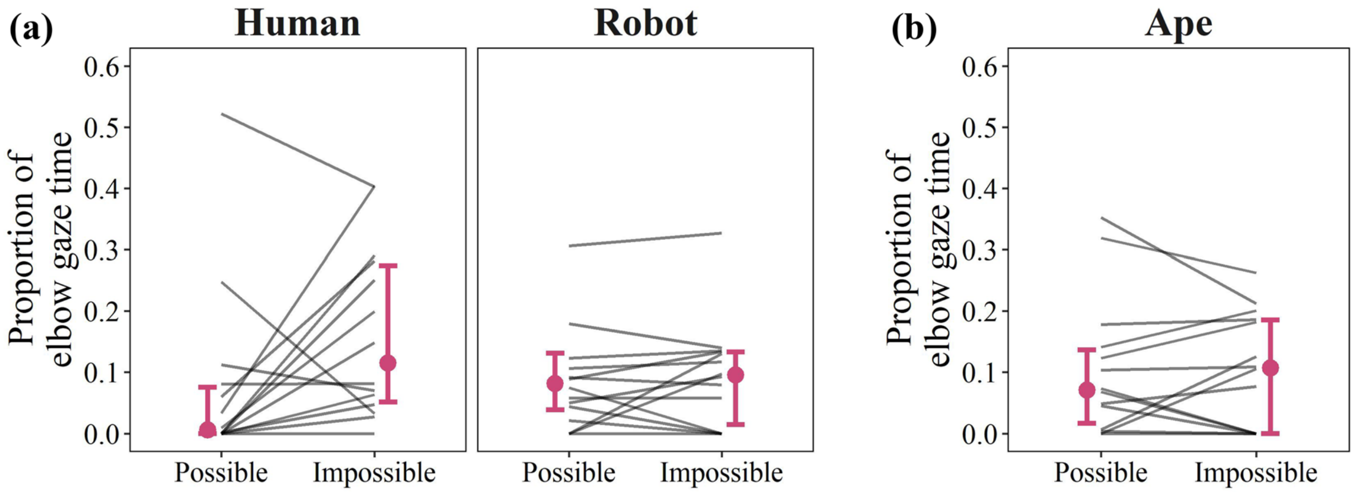

Fig. 4 **a** Proportion of time spent looking at elbows when viewing the human and robot animations (experiment 1). **b** Proportion of time spent looking at elbows when viewing the biped ape animations (experiment 2). *Dots* show the median, *error bars* show the first and third quartiles

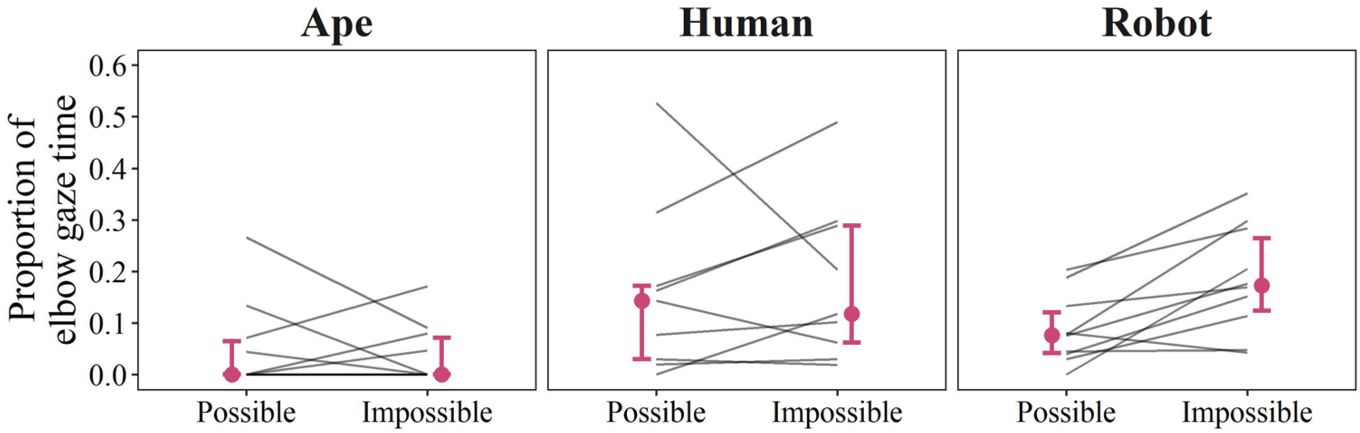

Fig. 5 Proportions of time spent looking at elbows when viewing human, robot, and quadruped ape animations (experiment 3). *Dots show* the median, *error bars* show the first and third quartiles

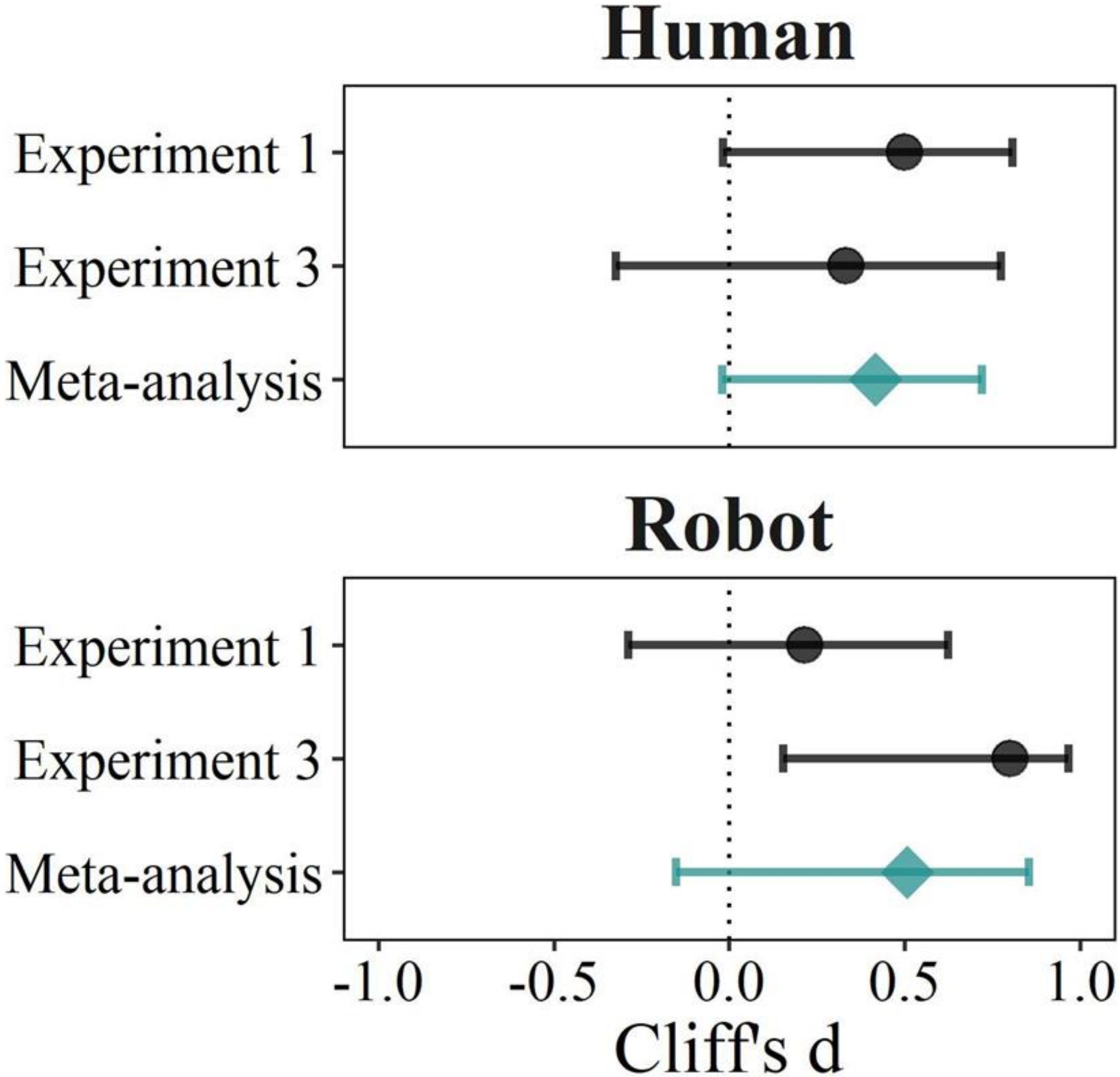

Fig. 6 Forest plot showing the results of the meta-analysis. *Upper panels* Results for the human animations, *lower panels* results for the robot animations. *Dots* show Cliff's d, *error bars* show the corresponding 95% confidence intervals