Acta Psychologica 139 (2012) 404–416

Contents lists available at SciVerse ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

The end-state comfort effect facilitates joint action

Oliver Herbort [a,*,1], Arno Koning [b], Janet van Uem [b], Ruud G. J. Meulenbroek [b]

[a] University of Würzburg, Department of Psychology, Röntgenring 11, Würzburg, Germany
[b] Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, PO Box 9104, 6500HE Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Received 23 May 2011
Received in revised form 29 December 2011
Accepted 3 January 2012
Available online xxxx

PsycINFO classification:
2323
2330

Keywords:
Joint action
Prediction
End-state comfort effect
Goal inference
Grasping

ABSTRACT

Motor experts can accurately predict the future actions of others by observing their movements. This report describes three experiments that investigate such predictions in everyday object manipulations and test whether these predictions facilitate responses to the actions of others. Observing video excerpts showing an actor reaching for a vertically mounted dial, participants in Experiment 1 needed to predict how the actor would rotate it. Their predictions were specific to the direction and extent of the dial rotation and improved proportionate to the length of the video clip shown. Testing whether such predictions facilitate responses, in the subsequent experiments responders had to undo an actor's actions, back-rotating a dial (Exp 2) and a bar (Exp 3). The responders' actions were initiated faster when the actors' movements obeyed the so-called end-state comfort principle than when they did not. Our experiments show that humans exploit the end-state comfort effect to tweak their predictions of the future actions of others. The results moreover suggest that the precision of these predictions is mediated by perceptual learning rather than by motor simulation.

1. Introduction

Many tasks require the sequential execution of multiple actions to reach the task goal. When putting away a book, for instance, one must first grasp the book, then take it to a bookcase, and finally place it on the shelf with the other books. Each of the required individual actions can be executed in different ways. The degrees of freedom inherent to each element of the action sequence allow one to make anticipatory adjustments to each sub-action in order to reach the overarching goal as efficiently as possible (Cohen & Rosenbaum, 2004; Herbort & Butz, 2010, 2011a; Rosenbaum, van Heugten, & Caldwell, 1996, Rosenbaum et al., 1990, 1996). In our example it would be most efficient to pick the book up at its lower end if it is to be placed on a high shelf. Anticipatory adjustments not only afford successful behavior, but they also give away our action intentions in that a bystander may infer the final goal from the way an action is being carried out (Brass, Schmitt, Spengler, & Gergely, 2007; Cuijpers, van Schie, Koppen, Erlhagen, & Bekkering, 2006; Kilner, Friston, & Frith, 2007).

1.1. Predicting the actions of others

The ability to infer the intention of the actions of others may be essential to successfully complete everyday tasks. Collaborative motor tasks involving joint action strongly rely on people sharing motor intentions and action goals especially when these cannot be communicated verbally (Bratman, 1992; Newman-Norlund, Noordzij, Meulenbroek, & Bekkering, 2007; Newman-Norlund, van Schie, van Zuijlen & Bekkering, 2007 Povinelli, 2001). In competitive sports or when facing a threatening situation, it is often crucial or even lifesaving to be able to correctly predict the future actions of others (Parasuraman et al., 2009).

The capacity to predict the outcome of another's actions has been documented widely, particularly in relation to sports. Experts in ball sports are able to deduce detailed parameters of a ball's future trajectory from the movements their counterparts make while propelling the ball. Expert tennis players can even infer the direction of a shot from the kinematics of their opponent's body movements before ball-racket contact (c.f. Aglioti, Cesari, Romani, & Urgesi, 2008; Huys, Smeeton, Hodges, Beek, & Williams, 2008; Ward, Williams, & Bennett, 2002; Williams, Huys, Cañal-Bruland, & Hagemann, 2009). Here, the predictions were made by well-trained experts and the reported movement outcomes were typically of high relevance to the sport under study. The capacity to correctly infer the goals of another's actions is, however, not limited to trained experts. Most children and adults are able to infer whether and how

---

* Corresponding author. Tel.: +49 7071 75464.
E-mail addresses: oliver.herbort@uni-tuebingen.de (O. Herbort), a.koning@donders.ru.nl (A. Koning), r.meulenbroek@donders.ru.nl (R. G. J. Meulenbroek).

[1] Present address: University of Tübingen, Department of Computer Science, 72076 Tübingen, Germany.

0001-6918/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.actpsy.2012.01.001

someone is going to use or displace an object from the grip the person selects, for instance (Boria et al., 2009, Ortigue, Sinigaglia, Rizzolatti, & Grafton, 2010; Parasuraman et al., 2009). Moreover, the kinematics of another's action permits an observer to infer whether the actor intends to deceive him (Grèzes, Frith, & Passingham, 2004; Runeson & Frykholm, 1983; Sebanz & Shiffrar, 2009), or whether this person has competitive or cooperative intentions (Sartori, Becchio, & Castiello, 2011). Finally, the kinematics of human script even allows the course of a letter sequence to be predicted (Kandel, Orliaguet, & Viviani, 2000). In sum, when prompted to observe and categorize another's actions in everyday situations, most people are able to infer categorical goals and thus the future actions of the other.

1.2. Research question

We aim to extend the literature by examining whether adult subjects capitalize on the end-state comfort effect (see below) to predict the future actions of others in object-manipulation tasks. Our second aim was to establish whether these predictions affect behavior in joint-action tasks in which the participants are not explicitly instructed to observe and judge the behavior of their counterparts. These tasks did not necessarily require a prediction of the other's upcoming actions to be completed.

The end-state comfort effect refers to the finding that, when manipulating objects, humans tend to adjust the orientation of their grip, i.e., the orientation of the hand relative to the object to be displaced, to future manipulations. For example, Rosenbaum et al. (1990) observed that a horizontally placed bar was grasped with an overhand grip when the bar was to be rotated by 90° in clockwise direction, but was grasped with an underhand grip when it needed to be rotated counterclockwise (c.f. Johnson, 2000; Rosenbaum, Cohen, Jax, Weiss, & van der Wel, 2007, 1996, Rosenbaum et al., 1996). When an object allows for a continuous selection of grasp orientations (e.g. a control knob), participants parametrically adjust their grasp orientation to the direction and extent of the desired rotation (Herbort & Butz, 2010, 2011a): participants used a pronated (right) forearm when asked to turn a dial clockwise, but a supinated forearm when asked to turn it counterclockwise. Moreover, the degree of anticipatory pronation or supination depended on the extent of the dial rotation. Thus, the initial grasp orientation carries crucial information that allows an observer to predict the direction into and the extent by which someone is about to rotate the object.

In Experiment 1 we examine whether participants can use the end-state comfort effect to predict the upcoming actions of others. With Experiments 2 and 3 we test whether, in a joint-action situation, these predictions facilitate the initiation of complementary actions without explicitly prompting the responder to gauge the movements of a collaborator.

2. Experiment 1

Our first experiment was designed to verify whether, when viewing an actor reaching for a dial, participants pick up the anticipatory forearm rotations to help them to predict the way the actor is about to rotate a dial. Participants were only presented video excerpts showing the reach-and-grasp stage of the actor's forearm motion; they never saw the actual dial rotation. For each video clip, they were asked to indicate to which of five possible positions they expected the dial was about to be turned. As we were interested in action observation, the actor was filmed and responses were coded from a third-party perspective. If people are indeed sensitive to anticipatory forearm rotations of others, the participants' performance, in terms of correctly predicted end-point positions of the dial, was expected to exceed chance level.

2.1. Method

2.1.1. Participants

Participants were all undergraduate students, with 36 (23 males, mean age 25 years) completing the experiment at the Radboud University Nijmegen (The Netherlands) and 27 (10 males, mean age 27 years) doing so at the University of Würzburg (Germany). According to Coren's (1993) Lateral Preference Inventory (LPI), 54 participants were right-handed, 7 left-handed, and 2 ambidextrous (mean LPI score 2.9). All reported normal or corrected-to-normal vision. They were compensated for their time with course credits or money.

2.1.2. Stimuli

Using a SONY digital video camera (DCR-TRV330E) at 24 frames per second (fps) the actor (the first author) was filmed seated behind a vertically mounted red dial with a red pointer pointing straight up, allowing the participants full view of his active hand and arm. Fig. 1a shows several frames of one of the video clips. IREDs were placed on the distal end of the actor's forearm to record his movements (Optotrak® 3020, Northern Digital, Waterloo, Canada). In the starting position the actor always had both hands palm down on his upper legs. Fixed around the red dial facing the actor and hence not visible to the participant, was a black semi-circular disk that was mounted with five evenly distributed LEDs. One LED coincided with the pointer's vertical default position (designated 0°), with two LEDs arranged at either side of the default position at a ± 45° and a ± 90° angle.

The recording procedure aimed at eliciting natural behavior as much as possible. After a preparatory auditory signal alerting the actor to the upcoming trial, one of the five LEDs lit up, prompting the actor to reach for and grasp the dial and rotate it to the position indicated by the LED. If the 0° LED lit up, the actor was to grasp the dial without rotating it. After completion of each action, the relevant LED turned off signaling the actor to place his hand back on his leg. Next, a differently pitched beep prompted him to rotate the dial back to its upright 0° position. At all times the actor used his dominant right hand. Note that we only used rotations of up to 90° because we had earlier found that participants also pre-rotate their arms when preparing to turn a knob by only small extents (Herbort & Butz, 2010). Five videotapes were made for each of the five final pointer positions.

The forearm position and orientation data were analyzed to select five movements (one for each target angle) of similar duration in which the dial was reached for and grasped with a forearm orientation that was close to the average for the respective target position. The reach and grasp movements together took between 850 ms and 940 ms. At the moment the dial was grasped, the forearm orientations were 54.1°, 41.9°, 16.6°, −48.2°, and −64.7° for the five (1–5) respective target rotations (positive values denote supinations). Please note, forearm rotations also affect the spatial orientation of more distal parts of the body, such as the hand, thumb, or fingers. Thus, when referring to forearm orientations, through which rotations of the hand are chiefly generated (Herbort & Butz, 2011a), we also imply the rotations of the hand and fingers. The 24 frames showing the movements up until the hand came into contact with the dial were exported and later presented at 24 fps, resulting in clips with an overall length of 1 s.

2.1.3. Design and procedure

Participants were shown different excerpts of the five video clips. Besides the full-length clip (all 24 frames), they could also be shown the first quarter only (0–25%, frames 1 through 6), the second quarter (25–50%, frames 7 through 12), the third quarter (50–75%, frames 13 through 18), the fourth quarter (75–100%, frames 19 through 24), the first half (0–50%, frames 1 through 12), the second half (50–100%, frames 13 through 24), or the first three quarters of the clip (0–75%,

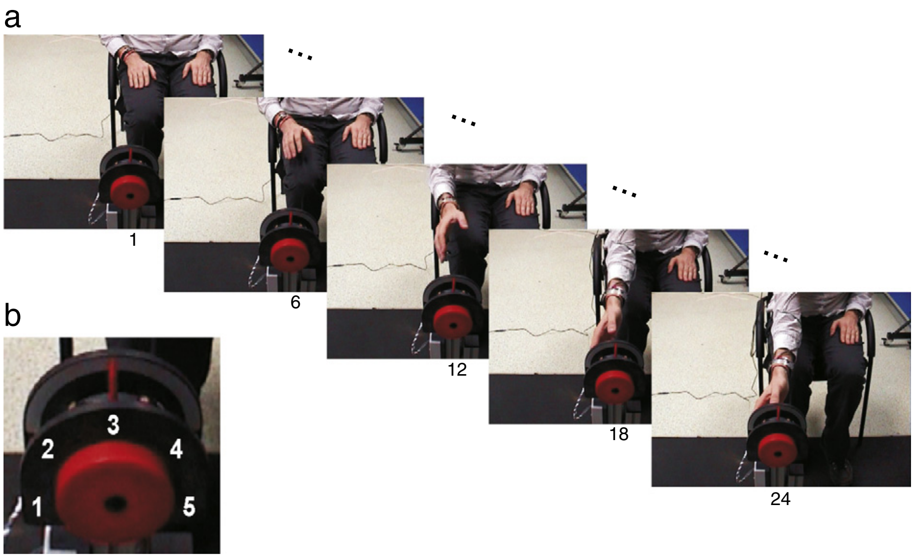

Fig. 1. Stimuli of Experiment 1. (a) Depicted are frames at 0%, 25%, …, 100% (frame numbers 1, 6, 12, 18, 24) of one of the video clips showing different stages of the reach-and-grasp movement toward the dial. (b) The response screen that was presented to cue the participants' selection of the targeted dial rotation, showing: an enlarged close-up of the dial with numbered end positions from the responder's perspective.

frames 1 through 18). In the case of partial clips, the participants saw a black screen during the non-pertinent frames. Including the blacked-out frames, all video clips hence always lasted 1 s. Fig. 1a shows the frames at 0%, 25%, …, 100% of an exemplar clip.

Given the two independent variables comprising the five target dial positions (1, 2, 3, 4, 5) and eight video segments (0–25%, 25–50%, 50–75%, 75–100%, 0–50%, 50–100%, 0–75%, 0–100%), in a single session each participant was presented a total of 40 (5 × 8) unique trials. The experiment started with a short explanation of the procedure. The participants were told they were going to see a series of short video clips of a man about to rotate a dial to an unknown position and asked to predict, based on the segment shown, how they thought the actor was going to rotate it. Each trial started with a blank screen (1000 ms), followed by a fixation cross (1000 ms). Next, a clip was shown (total duration 1000 ms), followed by another blank screen (1000 ms), after which a window appeared with the question “To which number will the dial be turned?”, below which the dial was depicted with the five numbers indicating the target positions superimposed on it (Fig. 1b). The picture of the dial was an enlarged outtake from a single frame, thus showing the dial from the customary third-party perspective. The participants were instructed to press the relevant key (1 through 5) of the top row of a standard keyboard. All 40 trials were randomly presented to each participant in a single session. Prior to the experimental trials, three practice trials were presented, which were not analyzed. The full experiment took between 5 and 10 min to complete.

2.2. Results

The data were pooled because preliminary analyses revealed no significant differences in overall accuracy between the two populations (Nijmegen vs. Würzburg) or between left-handed and right-handed participants.[2] For analysis, the percentage of correct responses for each participant and video segment was computed. We did not include target dial position as a factor, because some were subject to response biases. The biases resulted from the uneven distribution of the overall frequencies of the different responses. For example, regardless of the target position, the participants pressed 3 in 71% of all trials for the first-quarter (0–25%) segments, with 1 and 5 being selected in only 5 to 6% of the trials. Moreover, this preference for option 3 decreased with the length and stage of the segment shown. Online Resource 1 provides a table of the frequency of correct responses per video segment and target dial position. If repeated tests were computed, the alphas were adjusted to yield a global alpha or 0.05 according to the Bonferroni–Holm procedure.

2.2.1. Effects of video segment on prediction accuracy

Fig. 2 shows that the percentage of correct responses exceeded chance level (20%) for all video segments, two-sided one-sampled t-tests, all $T(62) > 3.2$, $ps < .002$, c.f. Online Resource 1. Next, we investigated whether there was a significant increase in prediction performance when increasingly longer segments (i.e., larger number of frames) were shown that all started with the onset of the reach-

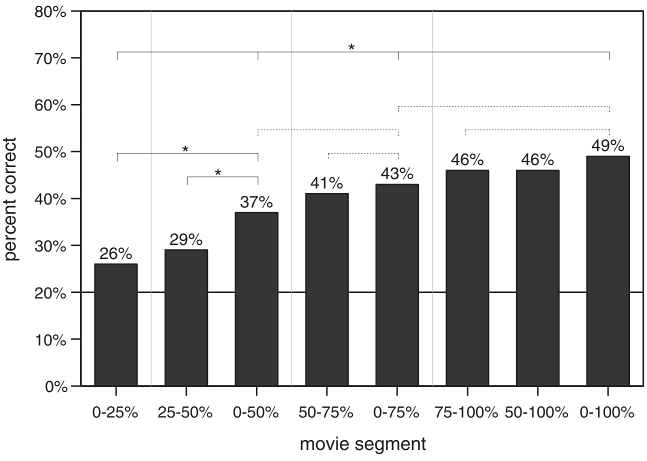

Fig. 2. Percentage of correct responses. The percentage of correct responses for the different video segments across trials. The bold black line at 20% indicates chance level for responses. The brackets with asterisks indicate significant differences and the dotted brackets without asterisk indicate tested but non-significant differences.

---

[2] The left-handed and ambidextrous participants performed numerically better than their right-handed peers (49%, 51%, and 38% correct responses, respectively). However, an ANOVA with video segment as the within-subject factor and handedness as the between-subject factor (left-handed, ambidextrous, right-handed) showed the effect of handedness and the interaction to be non-significant, $F(2, 60) = 1.71$, $p = .19$, $F(14,420) = 1.12$, $p = .35$. We accordingly analyzed the data independently of handedness.

and-grasp movement. To this end, we ran a repeated measures ANOVA with segment as the within-subject factor (the 0–25%, 0–50%, 0–75%, 0–100% clips) and percentage of correct responses as the dependent variable. This yielded a highly significant main effect for segment, $F(3,186)=18.9$, $p<.001$. Post-hoc t-tests revealed significant differences between the 0–25% and 0–50% clips, but not between the 0–50% and 0–75% segments or the 0–75% and 0–100% ones when corrected for a global alpha of 0.05, 0–25% vs. 0–50%: $T(62)=4.3$, $p<.001$, c.f. Online Resource 1. As expected, prediction performance increased with the number of frames presented, but only significantly so in the first half of the clip.

We also compared the percentage of correct responses of segments that started at different points but ended with the same frame: 25–50% vs. 0–50%, 50–75% vs. 0–75%, and 75–100% vs. 0–100%. Paired t-tests again revealed a significant difference for the 25–50% and 0–50% segment comparison only, $T(62)=3.2$, $p<.01$, c.f. Online Resource 1. Thus, seeing more frames indeed only significantly improved response accuracy if the segment showed the early stages of the movement.

2.2.2. Error types

To gain more insight into the response choices of individual participants and the type of errors they made, we analyzed the responses given when viewing the full clip (0–100%) in more detail (the data for all video segments can be found at Online Resource 1). Table 1 lists the responses (in percentages) for each of the five possible dial rotations for the full clip (0–100%), with the rows showing the targeted end-positions and the columns the participants' responses. The percentages (in bold font) on the diagonal axis indicate the percentages of correct responses.

The table reveals two main findings. First, the data show that the participants did not only correctly predict the direction but also the target position of the upcoming rotation in most cases (target positions 1, 2, 5), with the exception of the trials in which dial position 4 was targeted. Here the participants pressed option 5 more frequently than they did option 4. As to target position 3, the response frequencies were distributed around the correct response, but with a bias toward responses 1 and 2. Nevertheless, these biases do not invalidate our conclusion that the participants were not only sensitive to the targeted rotational direction but also able to specify in the majority of trials to which position the dial was about to be rotated.

Furthermore, the responses in trials with target dial positions 1, 2, 4, and 5 had a bimodal rather than a normal distribution, $\chi^2(2)\text{s}\geq 54.4$, $ps<.001$. In these trials the response options that were the 'opposite' of the most frequently chosen one (i.e., 5, 4, and 1, respectively) were also selected relatively frequently, with the 'in-between' responses being chosen relatively infrequently. Considering the trials with target dial position 5, for instance, the participants correctly responded in 57.1% of the trials and opted for response 1, its exact opposite, in 22.2% of the trials. Responses 2, 3, and 4, which were closer to the correct response than response 1, were given less frequently than response 1.

Table 1
Percentages of responses for all categories split by target dial position for the 0–100% movie segment.

		Responses				
		1	2	3	4	5
Target dial position (correct answer)	1	**52.4%**	22.2%	3.0%	4.8%	17.5%
	2	11.1%	**58.7%**	7.9%	17.5%	4.8%
	3	6.3%	30.2%	**44.4%**	17.5%	1.6%
	4	19.0%	7.9%	1.6%	**33.3%**	38.1%
	5	22.2%	3.2%	0%	17.5%	**57.1%**

2.3. Discussion

Experiment 1 showed that when people observe a reach-and-grasp movement toward an object, they can predict the goal of the intended object manipulation above chance level. Our participants based their predictions on the anticipatory rotations of the forearm and hand the actor adopted to complete a specific dial rotation comfortably. Moreover, most often they did not need to see the anticipatory movement in its entirety: in many trials the video clips lasting only 250 ms sufficed to achieve a prediction performance above chance level. Prediction accuracy further increased when longer or later segments of the video clip were shown. The overall accuracy level of up to 49% is further evaluated in the General discussion.

A surprising aspect of the analyses of the responses for the full video clip (0-100%) was that for most target dial positions the participants' predictions followed a bimodal distribution. We see two potential reasons why 17.5%–22.2% of all responses were the (exact) opposite of the correct or most frequently chosen target (Table 1).

First, the short duration of the video excerpts and the subsequent appearance and design of the response screen may have induced some participants to base their responses on other spatial features discernable immediately before the dial was grasped rather than solely on the arm-hand rotations. For example, when the end-state comfort effect was fully apparent, the actor's index finger was pointing toward the dial position opposite the targeted one (see Fig. 1a, where position 5 is targeted, while the index finger points to position 1). Such salient visual features may have prompted some participants to select the opposite position. Two different response patterns are then feasible. On the one hand, a participant may be guided by other or additional spatial cues in some trials and by the arm-hand rotations in others, resulting in a moderately inconsistent response pattern. On the other hand, some participants may respond mostly on the basis of alternative spatial features whereas others do so mostly on the basis of the movement goal. The responses of the first group should then consistently show an inverted and those of the other group a 'normal', non-inverted pattern. If participants indeed based their responses on cues other than the arm-hand rotations, one would expect 'inverters' and 'non-inverters' to also differ with respect to their ability to discriminate among segments with different target positions.

Second, even though we used a still from a video clip showing the dial from the observer perspective in the response window, some participants may have erroneously interpreted it as presenting the dial positions from the actor's viewpoint. Their inverted responses would then reflect a correct prediction. Here, one would expect their responses to consistently be inverted and their ability to distinguish target positions to match that of the non-inverters.

Looking for consistently inverted response patterns, we correlated the target positions with the actual responses individually for each participant (see Online Resource 1, for details on the rationale and additional analysis). Consistent, accurate non-inverters should show a high positive correlation, the less accurate and thus inconsistent participants low or no correlation, while consistent and accurate inverters should show a high negative correlation. For the majority of participants (44 of 63) the correlation was positive, $T(38)>3.0$, $ps<.01$. However, we found a significant negative correlation for 15 participants, indicating that they consistently gave inverted responses, $T(38)\text{s}<-2.5$, $ps<.05$. There were only four participants with non-significant correlation coefficients. Finally, the ability to discriminate among different target positions was similar for the inverters and non-inverters (median $\text{r}=-.73$ vs. $r=.75$, see Online Resource 1 for further details). This suggests that the larger proportion of the inverted responses was primarily due to a misinterpretation of the dial perspective we presented to help them code their response. This would mean that the original results underestimated the actual response accuracy. After correction for inverted responses based on the sign of each participant's correlation coefficient, the average percentage of correct responses across segments increased to 49% and to 60% for the full clip.

Another interesting finding is, that in trials with target dial position 4, the most participants expected the rotation to be aimed at position 5. One explanation would be, that the difference between these two target orientations was too small to discern with confidence. However, one would then also expect a similar confusion in trials targeting position 5. It should also be noted here that the participants exhibited an overall bias toward response 5, which was generally selected more frequently than response 4.

Response 3 was selected the least number of times, possibly accounting for the comparatively small percentage of correct responses for trials with target position 3. The response bias toward position 1 or 2 when position 3 was targeted most likely resulted from the arm-hand orientation for position 3 being closer to the orientations adopted for dial positions 1 and 2 than those adopted for the other target rotations.

In sum, response accuracy well exceeded chance level, allowing the conclusion that people are to a fair degree sensitive to the end-state comfort effect in others. Moreover, they seem to be able to predict future actions of others in some detail, surpassing simple binary judgments as "going to the left" or "going to the right".

3. Experiment 2

Konvalinka, Vuust, Roepstorff, and Frith (2010) proposed that "successful interpersonal coordination depends on the ability to (a) predict the other's subsequent action, and (b) to adapt promptly" (p. 9). With Experiment 1 we showed that participants were able to predict to which position an actor was going to rotate a dial by observing the way the actor reached for the dial. In this second experiment we extend on these findings in two ways.

Our first aim was to test whether participants can put such action predictions to use when asked to promptly respond to a collaborator's actions. Even though Experiment 1 provided evidence of people's ability to make action predictions, it is an open question whether and how this ability affects action coordination in a real-life joint-action task. Such effects may only be minor because in a tightly coupled joint-action task participants often have little time and few resources to predict the upcoming movements of others or are not prompted to explicitly monitor a collaborator's movements. In a more naturalistic setting a joint-action task might facilitate responses to cues in another's movements as anticipation skills tend to improve when people are allowed to respond in a more natural way. Tennis players, for example, predicted the direction of a videotaped service more accurately, when asked to respond with strike-like movements than when asked to respond verbally (Farrow & Abernethy, 2003). In Experiment 2 we hence tested whether observing a collaborator's reach-and-grasp movement toward the dial (in terms of hand and forearm rotations) would affect the efficiency of the counterpart's responses.

Our second aim was to verify whether the predictions made in Experiment 1, which were based on the anticipatory movements of a single actor, could also be made based on the movements of different and naïve fellow participants. To address these questions, we asked two participants to alternately operate the same dial we used in the first experiment. One participant was instructed to grasp and turn the dial from a starting position to one of four possible end positions that was exclusively known to him/her. The other participant was asked to undo this rotation by restoring the dial to its original position. As we focused on the extent to which the behavior of the first participant affected that of the second, we labeled the first participant 'actor' and the second 'responder'.

The actor's reach-and-grasp movement and the turning action provide the responder with two critical pieces of information. The orientation of the pointer after the rotation is of course the imperative stimulus specifically informing the responder what to do. However, because people adjust their grasp orientation to the upcoming rotational movement, the actor's arm and hand rotation while reaching for the dial already cues the end position of the dial pointer (e.g. Herbort & Butz, 2010). Thus, by observing the actor's reach-and-grasp movements, the responder may already predict the action goal.

We manipulated the informative content of the actor's anticipatory movement by introducing end-state comfort validity as an additional factor. We expected predictions to be impaired when the actor's 'normal' reach-and-grasp movement was disturbed. Accordingly, in so-called end-state comfort valid (ESC-valid) trials the actor was allowed to plan and execute the dial rotations in a normal fashion, which would consequently result in 'natural' anticipatory forearm and hand rotations in accordance with the end-state comfort effect. In so-called ESC-invalid trials the actor was misinformed about the true turn angle until the onset of the reach-and-grasp movement. We expected that this would interfere with movement planning and result in anticipatory movements that would differ from those in the ESC-valid trials. It is important to note that the actors were not explicitly instructed to use an ESC-valid or ESC-invalid mode: ESC-validity was manipulated by either giving the actor the opportunity to preplan the reach-and-grasp movement (ESC-valid) or not (ESC-invalid). The result section will reveal that this manipulation was effective.

If responders indeed use the forearm and hand rotation of the actor's reach-and-grasp movement to predict the actor's turning action, it would be harder for them to respond in ESC-invalid trials, resulting in longer reaction times relative to those for the ESC-valid trials. Additionally, the kinematics of the responder's movement might also be affected in ESC-invalid trials. To ensure that the actor programmed the reach-and-grasp movement consistent with the initially provided turn angle, 80% of the trials were ESC-valid and 20% ESC-invalid. Because the forearm and hand orientation is mostly determined by the direction of the upcoming dial rotation and in order to keep the number of different trial types low, we decided to only switch the dial direction and not the turn angle in the ESC-invalid trials.

3.1. Method

3.1.1. Participants

Altogether 13 dyads were recruited. However, due to technical problems the data of only 11 dyads (22 participants; mean age 22 years, range 19–28, 12 women) could be analyzed. All participants were students of the Radboud University Nijmegen and all were right-handed according to Coren's (1993) LPI (mean LPI-score = 3.9). They received either course credits or payment and all gave their informed consent. The dyads already knew each other when signing up and were of similar age (mean absolute age difference was 1.8 years). Four were mixed sex dyads, four all female, and three all male.

3.1.2. Procedure

The participants were randomly assigned a role (actor or responder). Wearing headphones, they were seated in chairs (height 50 cm) placed at a distance of 60 to 80 cm to the dial, which required them to lean slightly forward in order to reach it. All dyads operated the dial with their right hand. Before the start of the experiment, the dyads were familiarized with the task. Fig. 3 outlines the trial procedure (exemplar trials can be viewed in video 1).[3] Each trial began with both participants resting their arms on their thighs palm down and the dial

---

[3] The video clip can be downloaded from http://www.uni-wuerzburg.de/fileadmin/ext00209/user_upload/Videos/endstate_comfort_effect_enables_prediction_exp_2.mov. The tone signals you hear were edited in and correspond to what the actor (left channel) and observer (right channel) heard in their headphones.

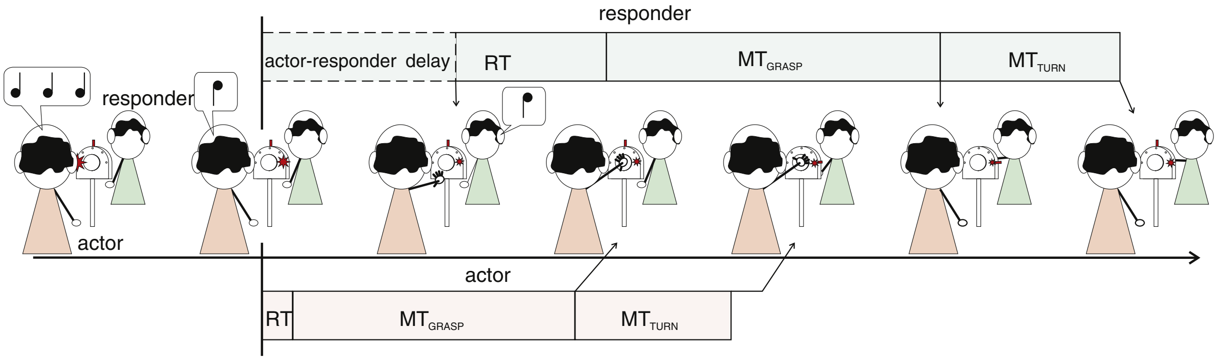

Fig. 3. Trial procedure of Experiment 2. The sequence of events in an ESC-invalid trial for actor (front) and responder (back). The cartoons show how the dyads were positioned, the dial and pointer, and the respective LEDs lighting up, reflecting the split-second change in the targeted end position of the dial. The boxes specify the timeline of the trial based on average reaction times, movement times, and a 400 ms actor-responder delay, allowing the actions of actor and responder to partially overlap.

pointer pointing upward (designated 0°). Next, the target LED switched on and a sequence of three low-pitched tones (440 Hz, 50 ms) followed by one high-pitched tone (880 Hz, 50 ms) was played to the actor only. The onsets of the tones were separated by 800 ms intervals. The actor was instructed to turn the dial to the location of the target LED and to start the movement at the precise moment the high-pitched tone sounded. To keep the actors' turning movements short, they were instructed to move the dial quickly and not necessarily to the precise target position. In the ESC-valid trials the original target LED remained lit, while in the ESC-invalid trials, as soon as the high-pitched tone was played, the target LED switched off and the LED cueing a rotation in the exact opposite direction lit up. Thus, in ESC-invalid trials the turn angle could jump from −90° to 90°, from −45° to 45°, from 45° to −45°, or from 90° to −90°. After a variable delay (300, 400, or 500 ms), a high-pitched tone sounded in the responder's headphones instructing him/her to quickly and accurately return the pointer to the upward position. We introduced the variable delay to avoid anticipatory response initiations by the responder. After 3200 ms the LED switched off again, with the next trial being initiated 1600 ms later.

Each trial type was a combination of three factors: turn angle (−90°, −45°, 45°, 90°), ESC-validity (valid, invalid), and actor-responder delay (300 ms, 400 ms, 500 ms). There were five experimental blocks separated by short breaks, each consisting of four presentations of all ESC-valid trials and one presentation of all ESC-invalid trials, resulting in 80% ESC-valid and 20% ESC-invalid trials, with each block comprising 60 and the full experiment 300 trials.

3.1.3. Data recording and analysis

A two-camera, optical motion-tracking system (Optotrak® 3020, Northern Digital, Waterloo, Canada) recorded the dial and forearm displacements at 100 Hz. IREDS attached to a semi-circular disk mounted close to the dial pointer served to record the dial displacements, while the dyad members both wore an IRED wrist-cuff close to the right wrist joint to help record their forearm rotations. Position and orientation data were smoothed with a 10 Hz, 2nd order low-pass Butterworth filter. The forearm orientation data were aligned so that 0° corresponded to the average orientation of the forearm resting on the thigh before movement initiation (with positive values denoting supinations for actor and responder). Based on the proximities of the participants' forearms to the dial and dial rotation rates, a semiautomatic process segmented the dial displacements into actor and responder movements. Next, hand movement onset (defined as the moment the forearm velocity first exceeded 50 mm/s), dial-turn onset (the first time the dial rotation exceeded 50 deg/s), and dial-turn completion (the last time the dial rotation exceeded 50 deg/s) were extracted for both participants using the segmented data. Reaction times (RT_{ACTOR}, $RT_{RESPONDER}$) were defined as the interval between the presentation of the high-pitched tone and movement onset for the respective dyad members. Actor and responder times were accordingly defined with respect to different events. Reach-and-grasp times ($MT_{GRASP,ACTOR}$, $MT_{GRASP,RESPONDER}$) were defined as the interval between movement onset and the onset of the dial turn of each respective dyad member, while their dial-turning times ($MT_{TURN,ACTOR}$, $MT_{TURN,RESPONDER}$) were defined as the interval between the onset of the dial turn and its completion. The respective forearm orientations at the time of grasping ($FO_{GRASP,ACTOR}$, $FO_{GRASP,RESPONDER}$) were defined as the forearm orientation at the onset of the dial turn. Additionally, the pointer orientation after the offset of the actor's dial turning movement was extracted (PO_{ACTOR}).

All trials were visually checked and discarded if there were substantial gaps[4] in the data recordings between the actor's movement onset and the end of the responder's turning movement or if the data could not be segmented (8.9%). From the analyzable data, trials were excluded in which the participants did not respond correctly[5] or in which the overall movement time ($RT+MT_{GRASP}+MT_{TURN}$) of either actor or responder deviated from the respective trial type's mean by more than two standard deviations (7.5%).

3.2. Results

Preliminary within-subject ANOVAS with factors turn angle, ESC-validity, and actor-responder delay were conducted independently for the dependent actor and responder variables for those dyads that delivered data for each cell of the ANOVA. The analysis revealed only one significant main effect and only one significant interaction that included the actor-responder delay.[6] $RT_{RESPONDER}$ was smaller for longer actor-responder delays, $F(2,16)=33.8$, $p<.001$. For $MT_{TURN,ACTOR}$ the three-way interaction reached significance, $F(6,48)=3.9$, $p=.032$. The interaction resulted because $MT_{TURN,ACTOR}$ was shorter in ESC-invalid trials than it was in ESC-valid trials, but only for the 90° angle and the 300 ms actor-responder delay. Thus, the actor-responder delay did not interact with the factors turn angle and ESC-validity for all dependent variables except for $MT_{TURN,ACTOR}$. $MT_{TURN,ACTOR}$ was only unsystematically affected by the actor-responder delay and the three-way interaction did not carry over to any of the responders' dependent variables. Because the actor-

---

[4] Gaps of several 100 ms were considered substantial, as were those during which key events of the movements were likely to have occurred. Gaps resulted from occlusions caused by the participant's body, the other participant, or the dial setup. Shorter gaps were closed by linear interpolation before filtering.

[5] For a trial to be considered correct, the pointer needed to have landed within 15° of the upward default position at both trial onset and end. As the distance between the LED panel and the pointer introduced systematic parallax errors, the pointer orientation resulting from the actor's dial rotation had to be within 15° of the actor's average pointer orientation for the respective turn angle.

[6] Due to the partial overlap in the actor's and responder's action executions, the actions of the responder could have affected those of the actor. Hence, the factor actor-responder delay was also included for the actor. We report Greenhouse–Geisser corrected p-values but uncorrected dfs.

responder delay is not the main interest of our study and in order to be able to include all dyads into the analysis, we decided to average each dependent variable for each combination of turn angle and ESC-validity. Separate within-subject ANOVAs with factors turn angle (90°, 45°, −45°, −90°) and ESC-validity (ESC-valid, ESC-invalid) were conducted for all variables and the two dyad members.

3.2.1. Actor: forearm orientations and movement times

Fig. 4a shows that the manipulation of turn angle and ESC-validity affected the actor's behavior in the expected way. $FO_{GRASP,ACTOR}$ strongly depended on the turn angle, $F(3,30)=17.7$, $p=.002$. More importantly, the dependency of $FO_{GRASP,ACTOR}$ on turn angle, and hence the magnitude of the end-state comfort effect, was weaker in ESC-invalid trials than it was in ESC-valid trials, $F(3,30)=5.1$, $p=.031$, implying that the actors were unable to fully compensate for the sudden change of turn angle. Fig. 5 provides the average time-normalized trajectories of the forearm orientations for all the different trial types. The error bars reveal that the standard deviation of the mean anticipatory forearm rotations between participants was rather high. The data show that the trajectory of the forearm orientations in ESC-valid trials started to divert from those in ESC-invalid trials at an early stage.

PO_{ACTOR} depended on turn angle, $F(3,30)=4856.3$, $p<.001$. Additionally, the interaction between turn angle and ESC-validity reached significance, $F(3,30)=3.7$, $p=0.03$. For the turn angles 90°, 45°, and −45°, actors brought the pointer to virtually identical orientations in ESC-valid (98°, 56°, −54°) and ESC-invalid trials (98°, 56°, −54°). In trials with turn angle −90°, there was a small difference between ESC-valid (−95°) and ESC-invalid (−97°) trials. ESC-validity thus had a minor but significant effect on PO_{ACTOR}.

$MT_{TURN,ACTOR}$ was higher for the longer turns, $F(3,30)=101.1$, $p<.001$. Otherwise, neither RT_{ACTOR}, $MT_{GRASP,ACTOR}$, nor $MT_{TURN,ACTOR}$ was affected significantly by ESC-validity, turn angle, or the two-way interaction, all $Fs<2.1$, all $ps>.16$ (Fig. 4c–e). Thus, the experimental setup created a situation in which three conditions were established. First, in ESC-valid trials the actor's reach-and-grasp movements allowed the upcoming turn to be predicted. Second, in the ESC-invalid trials the grasping movements deviated from those adopted in the ESC-valid trials. And third, ESC-validity did not significantly affect the other variables, with the exception of a small effect on PO_{ACTOR} in trials with a −90° turn angle.

3.2.2. Responder: forearm orientations and movement times

Fig. 4b shows that $FO_{GRASP,RESPONDER}$ strongly depended on turn angle, $F(3,30)=69.7$, $p<.001$, but that it did not interact with ESC-validity, $F(3,30)=1.2$, $p=0.31$. Thus, also the responders exhibited an end-state comfort effect, but the modulating effect of ESC-validity in the actors' reach-and-grasp movements did not carry over to those of the responders. Fig. 5 shows that the trajectories of the forearm orientations in ESC-valid and ESC-invalid trials largely overlapped.

The differences in $RT_{RESPONDER}$ between ESC-valid and ESC-invalid trials show that the actors' adjustments to his/her forearm and hand orientations in the ESC-invalid trials impaired the responders' ability to promptly select and initiate the correct response, $F(1,10)=5.6$, $p=0.04$ (Fig. 4c). Please note that the effect of ESC-validity on PO_{ACTOR} cannot account for this effect, as ESC-validity affected PO_{ACTOR} for the −90° trials

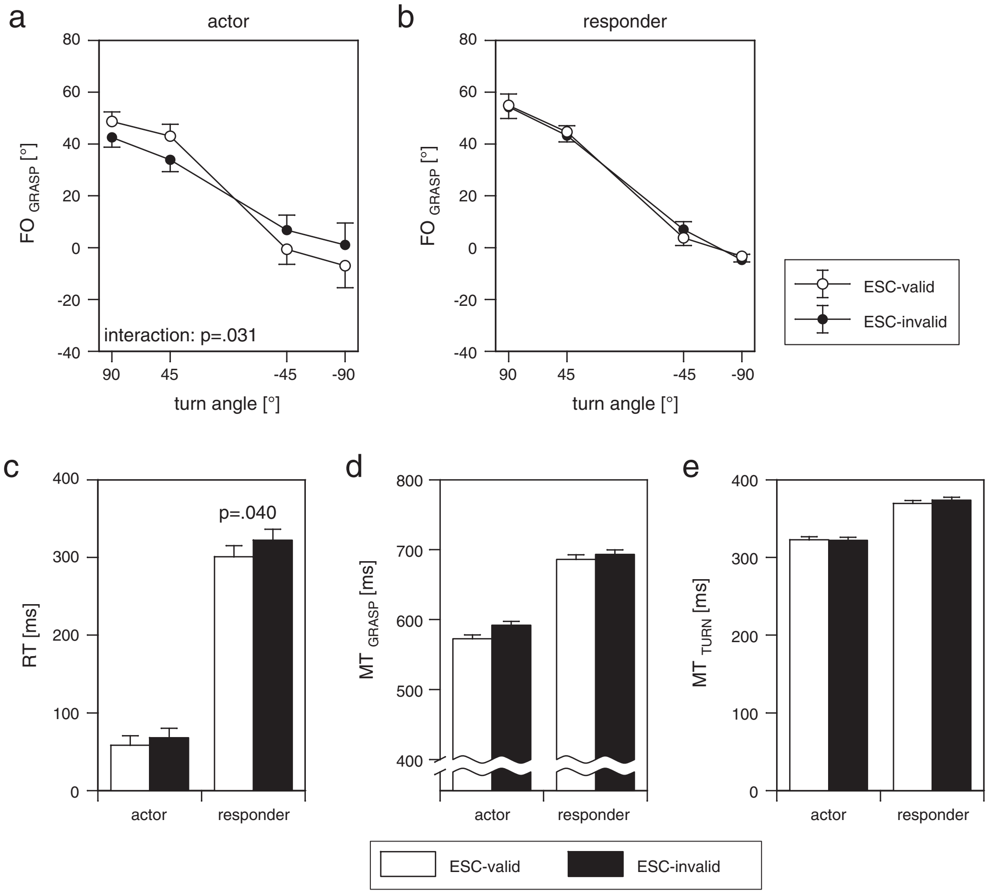

Fig. 4. Forearm orientations at grasping, reaction times, and movement times. Mean forearm orientations at grasping (positive values denote supinations) for actor (a) and responder (b),and the reaction times (c), movement times for the reach-and-grasp actions (d), and movement times for the dial rotation (e) for actors and responders in ESC-valid and ESC-invalid trials. Error bars show 95% within-subject confidence intervals for comparison between ESC-valid and ESC-invalid conditions (Loftus & Masson, 1994).

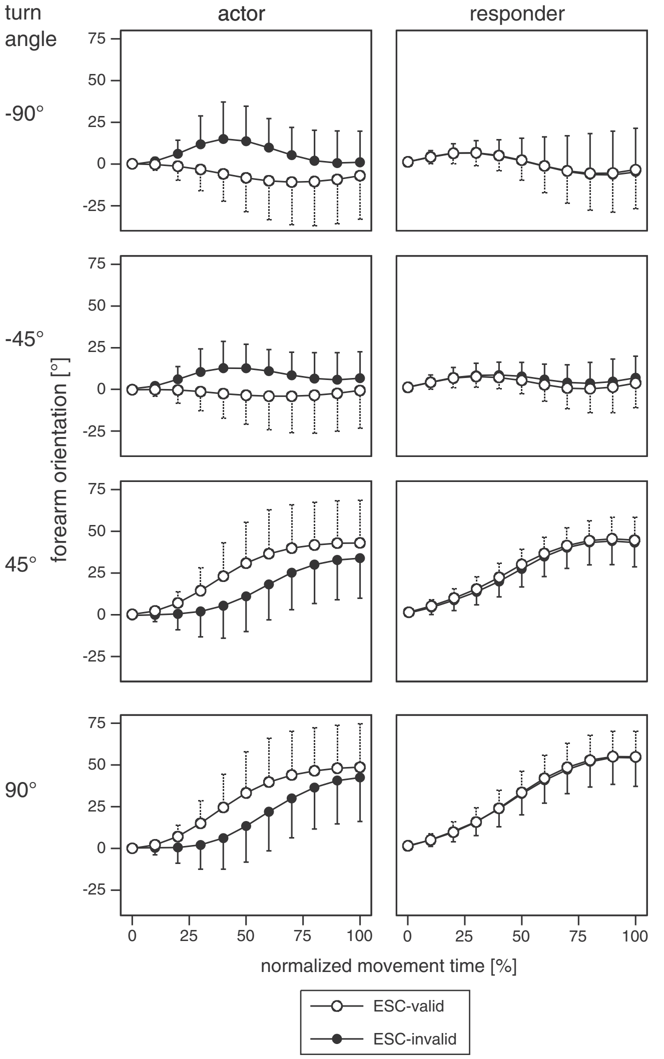

Fig. 5. Time series of reach-and-grasp movements. The charts show the average time-normalized trajectories of the actors' and responders' forearm orientations during the reach-and-grasp movements for the different trial types (with positive values denoting supinations). The error bars indicate between-subject standard deviations.

only, while it affected $RT_{RESPONDER}$ for all turn angles. $MT_{GRASP,RESPONDER}$ and $MT_{TURN,RESPONDER}$ were unaffected by turn angle and ESC-validity, with the exception that longer turns took more time to complete, $F(3,30) = 45.3$, $p < .001$; all other $Fs < 2.9$, all other $ps > .12$ (Fig. 4d–e).

3.3. Discussion

Experiment 2 showed that the responders were indeed able to use the actors' anticipatory hand orientations to speed up the initiation of their complementary actions. Moreover, it is surprising how sensitive they were to these cues. First, even though the actor's forearm orientations differed only by a few degrees between ESC-valid and ESC-invalid trials, this difference nevertheless affected the responder's performance. Second, due to the continuous nature of the task, there was some overlap in the grasp orientations the actor adopted in ESC-valid and ESC-invalid trials. Third, after the experiment, most responders reported that they had focused on the dial pointer and not on the actor's movements. Despite these facts, the differences in the actors' movements still affected the responders' reaction times. Together, these results show that even when people are not explicitly guarding another's movements, even subtle cues are processed, affecting the speed with which they respond to the other's actions.

Even though Experiment 2 showed that the participants were indeed sensitive to the movements of their counterparts, it remains to be discussed how the performance differences in ESC-valid and ESC-invalid trials can be explained. The data of Experiment 1 suggest that the ability to predict the outcome of an action sequence from visual cues facilitated the initiation of complementary actions in ESC-valid trials but impeded it in ESC-invalid trials. An alternative explanation would be that, rather than the prediction of the upcoming actions, the performance differences resulted from the fact that ESC-invalid trials captured more attention than ESC-valid trials, possibly because the movements in ESC-valid trials were more natural. Also, ESC-valid trials far outnumbered the ESC-invalid trials. To shed more light on this issue, we conducted a third experiment.

4. Experiment 3

Experiment 3 was devised to test whether the results of Experiment 2 could be exclusively explained by attentional factors or whether the predictability of the impending action also played a role. To this end, we simplified the paradigm in our third experiment. We this time opted for a task that was similar to the task introduced by Rosenbaum et al. (1990), in which a bar had to be grasped and rotated. Moreover, rather than a naïve participant, the actor now was a confederate of the experimenter. To rule out potential confounds of Experiment 2, this confederate was instructed to grasp a vertical bar using a thumb-up or thumb-down grip with equal probability. Both grips afforded natural motions and occurred equally frequently. Additionally, to facilitate the responders to glean information from the actor's actions, they were now instructed to react to the actor's actions as quickly as possible without having to wait for an auditory go-signal or a target cue.

More specifically, the current task features a vertically oriented bar that the actor (the confederate) picked up, rotated, and placed horizontally in front of a naïve participant. The responder was instructed to grasp the bar as quickly as possible and return it to its original vertical position. We predicted that the responder's ‘undo’ action would be initiated sooner when the actor showed the end-state comfort effect than when he did not.

4.1. Method

4.1.1. Participants

Fifteen students (14 women, 18–32 years, mean age 23 years) at the Radboud University Nijmegen participated in the experiment, for which they either received course credits or payment. None had taken part in either or both of the earlier experiments. All were right-handed according to the handedness scale of the Edinburgh Handedness Inventory (Oldfield, 1971). The data of one participant were discarded because of technical problems.

4.1.2. Procedure

Fig. 6 shows the experimental setup and trial procedure. One half of the 30-cm long, 3-cm diameter bar weighing 300 g was painted black and the other white. At the start of each trial the bar was placed vertically, black end pointing upward, in a fittingly large cylinder (height: 5 cm, inside diameter: 3.2 cm) fixed to the table. Also mounted on the table and 27 cm apart were two 10-cm high supports for the horizontal placement of the bar and allowing it to be grasped and manipulated with an overhand or underhand grip. We defined an overhand grip as featuring a slightly pronated forearm orientation with the fingers enclosing the horizontal bar from above with the thumb pointing down, and an underhand grip as showing a supinated forearm orientation with the fingers enclosing the bar from below with the thumb pointing upward.

The actor was instructed to pick up the vertically placed bar with the dominant right hand, rotate it, and position it horizontally upon

Fig. 6. Setup and trial procedure of Experiment 3. a) The black and white bar resting on the elevated support and the cylinder used to constrain the vertical starting and end position of the bar. b) The cartoon illustrated the trial sequence, with the actor in the foreground and the responder in the background.

the two supports. At the start of each trial he was told via a headphone whether to use a thumb-down (overhand) or a thumb-up (underhand) grip and whether to put the bar down with the black end pointing left or right. The responder was instructed to undo the actor's action as quickly as possible by grasping the horizontal bar with the dominant right hand and returning it to its original vertical position, black end pointing up.

We manipulated three independent variables, the first being rotational direction. After the actor had rotated the bar, the black side could either face left or right (actor's viewpoint). As the responder needed to back-rotate the bar clockwise (from his/her perspective) if the black end faced right (actor's perspective), we call this the clockwise condition and the opposite back-rotation (black end facing left) the counterclockwise condition. Second, ESC validity varied as the actor's grip could result in either a comfortable (ESC-valid) or uncomfortable (ESC-invalid) final posture (see Johnson, 2000, for subjective ratings of various grasp orientations). The thumb-up/counterclockwise rotations and the thumb-down/clockwise rotation were ESC-valid, while the thumb-up/clockwise and thumb-down/counterclockwise rotations were ESC-invalid. Finally, the repetition factor varied in that one of the mentioned four experimental conditions was presented five times in each of the experimental blocks. The position of a trial within such a 5-trial block is referred to as the repetition factor. Each combination of ESC validity (ESC-valid, ESC-invalid), rotation direction (counterclockwise, clockwise), and repetition (1st, 2nd, … 5th), was repeated 10 times throughout the experiment. The order of the blocks was randomized. Consisting of 10 blocks and a total of 200 trials, with 2-minute breaks after 50, 100 and 150 trials, completion of the experiment took each participant about 1 h.

4.1.3. Data analysis

The position data of the IREDs on the actor's and responders' wrists were filtered with a second-order, low-pass Butterworth filter having a cut-off frequency of 8 Hz. Actor and responder movement sequences were segmented into the reach-and-grasp movement from a fixed home position on the table top toward the bar, the bar-rotation, and the movement after releasing the bar back to the home position. To identify these submovements, a semi-automatic search procedure was used that identified local minima in the absolute wrist velocities of actor and responder. To measure how tightly actor and responder coordinated their actions, the time between the end of the actor's bar rotation and the onset of the responder's bar rotation was extracted (responder latency). Furthermore, the grasp orientation of actor (thumb-up, thumb-down) and responder (overhand vs. underhand grasp) at the start and end of the bar-rotation segment was extracted.

Trials were excluded if the data could not be segmented (121 trials, 4.3%), if the actor had not complied with the instructions (171 trials, 6.4%)[7], or if the responder rotated the bar by ±270° instead of ±90° to reach the trial goal (2 trials, 0.07%). The dataset of one participant was discarded as most (75%) of the trials were corrupted. Trials with a responder latency that deviated more than two standard deviations from the mean of the respective trial type were considered outliers and removed (71 or 2.9% of the remaining trials).

4.2. Results

4.2.1. Responder latency

Fig. 7a shows the effects of rotation direction and ESC-validity on responder latency. A within-subject ANOVA[8] of the responder latency and within-subject factors rotation direction (counterclockwise, clockwise), ESC-validity (ESC-valid, ESC-invalid), and repetition (1st, 2nd, …, 5th) revealed that responder latency was shorter if the actor grasped the bar with an ESC-valid grip than when he did so with an ESC-invalid grip, $F(1,12)=5.5$, $p=.037$. Furthermore, responder latency was longer if the bar had to be rotated counterclockwise than when it had to be rotated clockwise, $F(1,12)=110.8$, $p<.001$. The effect of repetition failed to reach significance, $F(4,48)=2.8$, $p=.067$. Descriptively, responder latency reduced by 13 ms on average from the first presentation of a trial to its second presentation, and did not change much after the second repetition. No interaction approached significance, all $ps>.256$.

4.2.2. Grip selections

Fig. 7b shows which type of grip the responders selected. A within-subject ANOVA of the relative frequency of overhand grips and the within-subject factors rotation direction (counterclockwise, clockwise), ESC-validity (ESC-valid, ESC-invalid), and repetition (1st, 2nd, …, 5th) revealed a main effect of rotation direction, $F(1,12)=27.0$, $p<.001$. Responders grasped the bar consistently with an overhand grip (100%) for clockwise rotations. In contrast, for counterclockwise rotations they only selected the overhand grip in 32% of trials. There was no significant effect of ESC-validity, repetition, or any interaction, all $ps>.248$. Thus, in most trials the responders selected a grip that enabled them to complete the back rotation with a comfortable thumb-up posture rather than an awkward thumb-down posture. The consistent selection of an overhand grip for clockwise rotations enabled them to begin and end the bar rotation

---

[7] The actor used an incorrect grip in 48 trials, an incorrect rotation in 82 trials, and an incorrect grip and rotation in 41 trials

[8] We report Greenhouse–Geisser corrected p-values and uncorrected degrees of freedom for F-values.

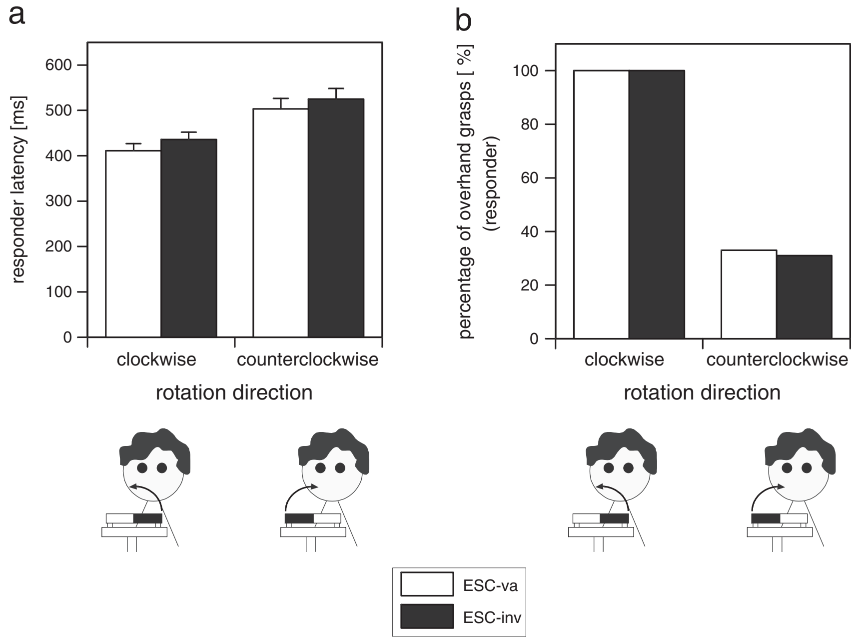

Fig. 7. Responder latency and grip selection in Experiment 3. Responder latencies (a) and percentages of overhand grips (b) selected by the responder as a function of rotation direction (from the responder's viewpoint) and ESC-validity of the actor's grip. The error bars show 95% within-subject confidence intervals for comparisons of ESC-valid and ESC-invalid trials (Loftus & Masson, 1994). As all participants selected grasps consistently, confidence intervals are invisible in the chart b.

comfortably, which is in line with previous reports (e.g. Rosenbaum et al., 1990). Besides preferring an underhand grip for the counterclockwise rotations, they also used overhand grips, possibly because goal-directed planning (favoring an underhand grip) and habitual action selection (favoring an overhand grip) compete in the case of counterclockwise rotations (Herbort & Butz, 2011b).

4.3. Discussion

Experiment 3 replicated the findings of Experiment 2. However, this time attentional factors were ruled out as a potential explanation for differences in the responders' reaction times for ESC-valid and ESC-invalid trials. In Experiment 3 the actors never had to correct sudden target switches during the execution of their reach-and-grasp movements, which could have captured the responders' attention and prolonged their responses in the previous experiment. Likewise, this time round the number of ESC-valid and ESC-invalid trials was the same, making it unlikely that ESC-invalid trials drew more attention because of their low frequency. Nevertheless, the coherence in the actors' reach-and-grasp movements in terms of the end-state comfort effect did affect the speed with which the responders initiated their reverse actions. On average, they reacted about 24 ms faster to ESC-valid grips than they did to ESC-invalid grips. This difference most likely results from the responder being able to correctly predict the actor's movements in ESC-valid trials, but being misdirected in ESC-invalid trials.

Another key difference between Experiments 2 and 3 is related to the dependent variable. Whereas the responders in Experiment 2 were instructed to start moving after an auditory signal, the responders in Experiment 3 could, in principle, initiate their movements at any time. Thus, instead of reaction times in Experiment 2, in Experiment 3 the time between the offset of the actor's and the onset of the responder's bar rotation was recorded. The time differences in this responder latency could hence have resulted from both the (re-)programming of an action prior to its initiation and the adjustment of an ongoing movement.

Finally, as ESC-valid trials were as frequent as ESC-invalid trials, the responder could not form associations between the actors' choice of grip and the subsequent bar rotation during the experiment. The responders' predictions must accordingly have been based on previous experience with the biomechanics of own or observed movements.

5. General discussion

In our three experiments we addressed the question whether people are sensitive to the end-state comfort effect in the movements of others and, if so, whether this sensitivity facilitates the prediction of specific parameters of and the responses to the other's subsequent action. The data showed the answer to all three questions to be “yes”. The participants in Experiment 1, who watched video excerpts showing the same actor reaching for and grasping a dial in preparation of different rotations, were able to predict the direction and extent of the impending rotation without ever seeing the actual dial turn. This implies that they took the other's end-state comfort effect into account and that they used this information to predict the upcoming movements and the ultimate action goal. Their predictions exceeded chance level even for those clips that only showed the early stages of the anticipatory movement, suggesting that they were highly sensitive to rather subtle cues in the actor's movements. In Experiments 2 and 3 we examined how this predictive ability affected responder performance in a collaborative setting. Here, participants watched another naïve participant or a confederate grasp and rotate a dial and bar, respectively. The responders initiated their reach-and-grasp movements and rotations faster if the movements of their counterpart obeyed the end-state comfort effect, which thus enabled them to predict the way in which the other was going to manipulate the object. In sum, also in joint tasks sensitivity to a collaborator's end-state comfort effect facilitates responses to his actions.

5.1. Perception of anticipatory movements of others

In Experiment 1 we found that the anticipatory forearm rotations of the actor enabled the observers to correctly predict the direction of the subsequent dial rotation above chance level, with the extent of the rotation also being estimated quite faithfully. Even though predictions were far from perfect, we were nevertheless impressed by the performance results. First, although predictions became more accurate when the participants were shown more of the clip, they were still able to extract salient information from the shorter segments showing the early stages of the reach-and-grasp movements. Second, predictions were based on anticipatory forearm rotations, which are not exhibited to the same extent by all humans and which may be reduced or absent in some participants or tasks (e.g. Herbort & Butz, 2010, 2011b, Janssen, Crajé, Weigelt, & Steenbergen, 2010). As the participant never saw the actor actually rotating the dial, they had no possibility to calibrate their predictions during the experiment. Thirdly, whereas the literature shows that experts (e.g., in sports) are able to make rather precise predictions of the outcome of highly relevant actions of others, in our first experiment the participants responded to actions the execution of which they would usually not devote much attention to and the outcome of which is usually in no way critical in everyday life (c.f. Sartori et al., 2011).

5.2. Responding to actions of others

The dyad members in Experiments 2 and 3 responded more quickly to the others' actions when the movements obeyed the end-state comfort principle. In Experiment 2 it facilitated movement initiation and in Experiment 3 reduced the time between the end of one member's action and the onset of the other member's reverse action. Various explanations have been put forward to account for the action facilitation in joint-action contexts. According to the direct matching hypothesis, identical circuits in the motor system are active when perceiving and executing actions, enabling imitation (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996). Indeed, when people react to the actions of others, imitative responses are faster than non-imitative ones (Brass, Bekkering, & Prinz, 2001; Tessari, Rumiati, & Haggard, 2002). The direct matching hypothesis could account for the results in Experiment 2 as both the actor and the responder had to execute essentially identical actions to achieve the goal. For example, when the actor had to turn the dial clockwise, the responder also needed to make a clockwise turn to move the pointer back to the default (0°) position. Then, the responses to ESC-valid actions were faster than those to ESC-invalid actions because only the first basically allowed an imitative response, while the ESC-invalid trials required non-imitative responses due to the actors' reduce end-state comfort effect.

While the participants in Experiment 2 may have successfully applied an imitation strategy, this was not possible for the participants of Experiment 3. Here, imitation was not feasible to undo the confederates' actions because they started with a vertically oriented grip and the responders with a horizontally oriented grip, requiring the responders to complement the actor's actions rather than imitate them (Newman-Norlund, Noordzij, Meulenbroek, & Bekkering, 2007; Newman-Norlund, van Schie, van Zuijlen, & Bekkering, 2007). Moreover, the responders' grip selections were not affected by those of the actors, also rendering an imitation strategy unlikely. Together, this suggests that the speed up in action coordination resulted from the responders predicting the upcoming actions of their counterparts.

Different models can account for such predictions. On the one hand, they may have been based on learned associations. In the perceptual learning of action sequences, the previous observations of rotation actions may have enabled the responders to form associations between the different reach-and-grasp movements and the subsequent rotations (e.g. Howard, Mutter, & Howard, 1992; c.f. Farrow & Abernethy, 2002). On the other hand, Simulation Theory suggests that people can infer the goals of others by simulating their behavior with their own action planning circuitry (Gallese & Goldman, 1998; Wolpert, Doya, & Kawato, 2003, but see Brass et al., 2007; Saxe, 2005). In this case, the observed actions would be expected to interfere with the planning of the own actions. Additionally, predictions of another's actions should depend on properties of the observers' motor repertoire. As such effects would not be predicted by pure perceptual sequence learning accounts, we will next discuss whether our data provide any evidence for the specific predictions of Simulation Theory.

A prerequisite of simulation theory is that observing the actions of others activates those neural states that would also be tapped if one would execute a similar action oneself. If this prerequisite was met, one would expect that observing the actions of the actor would have interfered with the action planning of the responder, because the responder had to plan his own actions and interpret the actor's movements more or less concurrently. Such alignment or interference effects of the movement kinematics have been reported for other joint-action tasks (e.g. Kilner, Paulignan, & Blakemore, 2003; Meulenbroek, Bosga, Hulstijn, & Miedl, 2007; Richardson, Marsh, Isenhower, Goodman, & Schmidt, 2007). Accordingly, one would expect the responders to have shown a reduced end-state comfort effect in ESC-invalid trials. However, the actor's grip selection had no significant influence on that of the responder in Experiments 2 and 3. Our findings hence do not support the notion that the actor's forearm rotations were mirrored by the responders' motor networks, as would have been predicted by Simulation Theory.

This notion is further supported by the data of Experiment 1. That is to say, it has been found that the variability of the end-state comfort effect can be considerable within a group of participants. For example, the difference between $\mathrm{FO_{GRASP}}$s before 90° and −90° rotations ranged from 5° to 159° ($SD=34°$) across participants in Experiment 2 (ESC-valid trials only). Nevertheless, all participants rotated the arm in the direction opposite to the dial turn. If they relied on their own motor systems to predict the actor's upcoming dial rotations, these interindividual differences should have impaired their predictions of the *extent* of the rotations but not their predictions of the *direction*. As an example, consider an observer who usually strongly rotates his arm before grasping an object for rotation. If this observer is to predict how an actor who usually rotates his arm only slightly in such cases is about to rotate the object, and if the observer uses his own motor system for his predictions, he is likely to underestimate the *extent* of the dial rotation, but would still be able to accurately predict the direction of the dial rotation. Thus, according to Simulation Theory many of the participants in Experiment 1 should have frequently over- or underestimated the extent of the dial rotation. However, at least for the full video clip, they were able to predict the parameter fairly accurately for most target positions. Thus, the response distribution also suggests that the participants did not rely on motor simulation. As suggested above, one possible reason for this is that a prerequisite for applying motor simulation was lacking in the task.

To conclude, our findings did not provide any evidence of the involvement of motor simulation processes in our object-manipulation tasks. We accordingly favor perceptual sequence learning from everyday observations of the actions of others as the basis for predictions in simple object-manipulation tasks (e.g. Howard et al., 1992). Nevertheless, although it seems likely that in our experiments responder predictions were not based on motor simulation processes, this does not imply that such processes are not involved in other tasks.

5.3. The end-state comfort effect

Besides providing information on how people capitalize on the end-state comfort effect to respond to the behavior of others, Experiment 2 also provides new insights into the end-state comfort effect itself. First, the data of the regular, undisturbed movements in ESC-valid trials show that forearm orientation was mostly determined by the direction of the upcoming dial turn and only little by its extent. Experiment 2 thus complements previous reports that showed that, rather than the precise end-state of the arm after rotation, the *rotation direction* is the major determinant of grasp orientation (Herbort & Butz, 2010, 2011a; Robert, Beurier, & Wang, 2009).

Also, the switches in the rotation direction had no significant effect on reaction times or the duration of the actor's reach-and-grasp movements. Thus, actors were able to quickly reprogram and adjust their forearm rotations. It is known that grip orientations can be corrected online to accommodate sudden changes in the orientation of the object to be grasped (Fan, He, & Helms Tillery, 2006). Given that the actors in Experiment 2 frequently initiated their reach-and-grasping movement before the target LED switched, our results extend previous findings by showing that the forearm orientation can also be adapted quickly to changes in the intended object *manipulation*. This supports the notion that the end-state comfort effect results from a simple selection process that can be easily adapted even in the course of an ongoing movement (Herbort & Butz, 2011a).

Finally, in the trials in which the target switched to its opposite, the actors quickly adjusted their grip orientation to the new target, although not completely so. For example, in the switch trials in which the dial had to be rotated by 90° or −90°, they used the forearm orientations that they also adopted for the 45° or −45° dial turns in the non-switch trials. This, and the fact that reaching did not slow down significantly, suggest that when reaching for the dial, the precise forearm orientation was not a strict constraint for action selection.

5.4. End-state comfort effect in joint action

Recently, there has been an increased interest in advance motor planning in joint action. Among other aspects, it was studied how participants hand over objects to a collaborator who wants to use the object for different tasks (Gonzalez, Studenka, Glazebrook, & Lyons, 2011; Ray & Welsh, 2011). The experiments showed that the participants were able to anticipate how their counterpart could best grasp the object to be able to use it readily. Thus, the information people glean from another's preferred, comfortable end postures enables them to align their own actions such that they facilitate the subsequent actions of the other. In our experiments, rather than assist their trial partners, the participants had to react to their actions. The results we obtained thus complement the previous findings by showing that knowing about another's preference for certain end-postures allows people to predict and complement the other's actions.

5.5. Conclusion

The reported experiments yielded two major findings. First, they demonstrated that humans take the end-state comfort effect into account to predict future actions of others, confirming that they are highly sensitive to movement cues in simple, everyday movements of other people. They further showed that the ability to make these predictions facilitates the way people respond to the actions of others, enabling them to coordinate joint actions better. Moreover, that these predictions are still efficient under the conditions of different real-life joint action tasks demonstrates that the movements of other persons can be interpreted very quickly and even if without explicitly directing attention toward them.

Acknowledgments

The authors acknowledge funding from the Emmy Noether program of the German Research Foundation (grant BU1335/3-1). They are indebted to Hubert Voogt for his technical assistance in conducting Experiments 2 and 3, thank Robrecht van der Wel for his valuable comments on earlier drafts of the paper, and thank Hanneke Meulenbroek–van der Meulen for improving the language.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.actpsy.2012.01.001.

References

Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience*, *11*(9), 1109–1116.

Boria, S., Fabbri-Destro, M., Cattaneo, L., Sparaci, L., Sinigaglia, C., Santelli, E., et al. (2009). Intention understanding in autism. *PLoS ONE*, *4*(5), e5596.

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, *106*(1–2), 3–22.

Brass, M., Schmitt, R. M., Spengler, S., & Gergely, G. (2007). Investigating action understanding: Inferential processes versus action simulation. *Current Biology*, *17*(24), 2117–2121.

Bratman, M. E. (1992). Shared cooperative activity. *The Philosophical Review*, *101*(2), 327–341.

Cohen, R. G., & Rosenbaum, D. A. (2004). Where grasps are made reveals how grasps are planned: generation and recall of motor plans. *Experimental Brain Research*, *157*(4), 486–495.

Coren, S. (1993). The lateral preference inventory for measurement of handedness, footedness, eyedness, and earedness: norms for young adults. *Bulletin of the Psychonomic Society*, *31*(1), 1–3.

Cuijpers, R., van Schie, H. T., Koppen, M., Erlhagen, W., & Bekkering, H. (2006). Goals and means in action observation: A computational approach. *Neural Networks*, *19*(3), 311–322.

Fan, J., He, J., & Helms Tillery, S. (2006). Control of hand orientation and arm movement during reach and grasp. *Experimental Brain Research*, *171*(3), 283–296.

Farrow, D., & Abernethy, B. (2002). Can anticipatory skills be learned through implicit video-based perceptual training? *Journal of Sports Sciences*, *20*, 471–485.

Farrow, D., & Abernethy, B. (2003). Do expertise and the degree of perception-action coupling affect natural anticipatory performance? *Perception*, *32*(9), 1127–1139.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *2*(12), 493–501.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*(2), 593–609.

Gonzalez, D. A., Studenka, B. E., Glazebrook, C. M., & Lyons, J. L. (2011). Extending end-state comfort effect: Do we consider the beginning state comfort of another? *Acta Psychologica*, *136*(3), 347–353.

Grèzes, J., Frith, C., & Passingham, R. E. (2004). Brain mechanisms for inferring deceit in the actions of others. *Journal of Neuroscience*, *24*(24), 5500–5505.

Herbort, O., & Butz, M. V. (2010). Planning and control of hand orientation in grasping movements. *Experimental Brain Research*, *202*(4), 867–878.

Herbort, O., & Butz, M. V. (2011). Habitual and goal-directed factors in (everyday) object handling. *Experimental Brain Research*, *213*(4), 371–382.

Herbort, O., & Butz, M. V. (2011). The continuous end-state comfort effect: Weighted integration of multiple biases. *Psychological Research*. doi:10.1007/s00426-011-0334-74 Advance online publication.

Howard, J. H., Mutter, S. A., & Howard, D. V. (1992). Serial pattern learning by event observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(3), 1029–1039.

Huys, R., Smeeton, N. J., Hodges, N. J., Beek, P. J., & Williams, A. M. (2008). On the dynamic information underlying visual anticipation skill. *Perception & Psychophysics*, *70*(7), 1217–1234.

Janssen, L., Crajé, C., Weigelt, M., & Steenbergen, B. (2010). Motor planning in bimanual object manipulation: Two plans for two hands? *Motor Control*, *14*(2), 240–254.

Johnson, S. H. (2000). Thinking ahead: The case for motor imagery in prospective judgements of prehension. *Cognition*, *74*(1), 33–70.

Kandel, S., Orliaguet, J. -P., & Viviani, P. (2000). Perceptual anticipation in handwriting: The role of implicit motor competence. *Perception and Psychophysics*, *62*(4), 706–716.

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

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: An account of the mirror neuron system. *Cognitive Processing*, *8*(3), 159–166.

Konvalinka, I., Vuust, P., Roepstorff, A., & Frith, C. D. (2010). Follow you, follow me: Continuous mutual prediction and adaptation in joint tapping. *The Quarterly Journal of Experimental Psychology*, *63*(11), 2220–2230.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*, 476–490.

Meulenbroek, R. G. J., Bosga, J., Hulstijn, M., & Miedl, S. (2007). Joint-action coordination in transferring objects. *Experimental Brain Research*, *180*(2), 333–343.

Newman-Norlund, R. D., Noordzij, M., Meulenbroek, R. G. J., & Bekkering, H. (2007). Exploring the brain basis of joint action: Co-ordination of actions, goals and intentions. *Social Neuroscience*, *2*, 48–65.

Newman-Norlund, R. D., van Schie, H. T., van Zuijlen, & Bekkering, H. (2007). The mirror neuron system is more active during complementary compared with imitative action. *Nature Neuroscience*, *10*, 817–818.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, *9*, 97–113.

Ortigue, S., Sinigaglia, C., Rizzolatti, G., & Grafton, S. T. (2010). Understanding actions of others: The electrodynamics of the left and right hemispheres. A high-density EEG neuroimaging study. *PLoS One*, *5*(8), e12160. doi:10.1371/journal.pone.0012160.

Parasuraman, R., de Visser, E., Clarke, E., McGarry, W. R., Hussey, E., Shaw, T., et al. (2009). Detecting threat-related intentional actions of others: Effects of image quality, response mode, and target cuing on vigilance. *Journal of Experimental Psychology: Applied*, *15*(4), 275–290.

Povinelli, D. J. (2001). On the possibilities of detecting intentions prior to understanding them. In B. F. Malle, L. J. Mosas, & D. A. Baldwin (Eds.), *Intentions and intentionality. Foundations of social cognition* (pp. 225–248). Cambridge, MA: MIT Press.

Ray, M., & Welsh, T. (2011). Response selection during a joint action task. *Journal of Motor Behavior*, *43*(4), 329–332.

Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R., & Schmidt, R. (2007). Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human Movement Science*, *26*(6), 867–891.

Robert, T., Beurier, G., & Wang, X. (2009). Arm postural anticipation for rotating a spherical object. *Computer Methods in Biomechanics and Biomedical Engineering*, *12*(1), 217–218.

Rosenbaum, D. A., Marchak, F., Barnes, H. J., Vaughan, J., Slotta, J. D., & Jorgensen, M. J. (1990). Constraints for action selection: Overhand versus underhand grips. In M. Jeannerod (Ed.), *Attention and performance*, *Vol. XIII*. (pp. 321–345). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rosenbaum, D. A., van Heugten, C. M., & Caldwell, G. E. (1996). From cognition to biomechanics and back: The end-state comfort effect and the middle-is-faster effect. *Acta Psychologica*, *94*, 59–85.

Rosenbaum, D. A., Cohen, R. G., Jax, S. A., Weiss, D. J., & van der Wel, R. (2007). The problem of serial order in behavior: Lashley's legacy. *Human Movement Science*, *26*, 525–554.

Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, *112*(4), 585–615.

Sartori, L., Becchio, C., & Castiello, U. (2011). Cues to intention: The role of movement information. *Cognition*, *119*(2), 242–252.

Saxe, R. (2005). Against simulation: The argument from error. *Trends in Cognitive Sciences*, *9*(4), 174–179.

Sebanz, N., & Shiffrar, M. (2009). Detecting deception in a bluffing body: The role of expertise. *Psychonomic Bulletin and Review*, *16*(1), 170–175.

Tessari, A., Rumiati, R., & Haggard, P. (2002). Imitation without awareness. *NeuroReport*, *13*(18), 2531–2535.

Ward, P., Williams, A. M., & Bennett, S. J. (2002). Visual search and biological motion perception in tennis. *Research Quarterly of Exercise and Sport*, *73*(1), 107–112.

Williams, A. M., Huys, R., Cañal-Bruland, R., & Hagemann, N. (2009). The dynamical information underpinning anticipation skill. *Human Movement Science*, *28*(3), 362–370.

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*, *358*, 593–602.