Perceiving conspecifics as integrated body-gestalts is an embodied process

Klaus Kessler[1] and Sebastien Miellet[2]

[1]Institute of Neuroscience and Psychology, University of Glasgow, United Kingdom

[2] Department of Psychology, University of Fribourg, Switzerland

Running Head: Embodied Body-Gestalt Completion

Word count Abstract: 155

Word count Main Text (incl. Appendix, excl. Captions and References): 9086

Corresponding author

Dr. Klaus Kessler

Department of Psychology

Glasgow University

58 Hillhead Street, Glasgow G12 8QB, UK

e-mail: k.kessler@psy.gla.ac.uk

phone: +44 (0)141 330 4774

Abstract

We investigated the effect of posture congruence on social perception. Specifically, we tested the hypothesis that completing "body-gestalts", rather than being a purely visual process, is mediated by congruence in the postures of observer and stimulus. We developed novel stimuli showing a face and two hands that could be combined in various ways to form "body-gestalts" implying different postures. In three experiments we found that imitative finger movements were consistently faster when the observer's posture matched the posture implied by the configuration of face and hands shown on screen, suggesting that participants intuitively used their own body schema to 'fill in the gaps' in the stimuli. Besides shaping how humans perceive others' bodies, embodied body-gestalt (eBG) completion may be an essential social and survival mechanism, e.g. allowing for quick recovery from deceptive actions. It may also partly explain why humans subconsciously align themselves in everyday interactions: This might facilitate optimal co-representation at higher, conscious levels.

Keywords: Social perception, alignment, self-other mapping, embodiment, posture/motor resonance, body schema

Introduction

When we perceive others in everyday encounters, their bodies are often partially occluded. For instance, while sitting in a café reading a newspaper, a person's lower body would be occluded by the table while large parts of their upper body would be hidden behind the newspaper. Nonetheless, human observers perceive the visible body parts (e.g. the hands holding the newspaper) and the face as belonging to the same person without being aware of any discontinuities in the visual input. It seems that the missing parts are automatically filled in to form a coherent percept of the other person. Of course this does not apply to human bodies only: The human brain is generally capable of completing partially occluded objects and some of the most prominent examples are known as "gestalt" phenomena (e.g. Kanizsa, 1974). Hence, gestalt completion based on visual principles and/or prior visual experience might be a general mechanism that also applies to human bodies.

However, observing partially occluded human bodies could be a special case of gestalt completion. In contrast to inanimate objects, humans posses a body of their own that shares essential visual and kinematic features with the bodies of conspecifics. Moreover, for humans as social animals congruence between their own and others' bodies matters for a wide variety of social behaviours ranging from complex joint actions, where corresponding body parts have to be tracked and moved in highly coordinated fashions (e.g. dancing), to competitive actions such as person-to-person combat, where tracking the adversary's body parts in relation to one's own matters for very different reasons. Therefore, perceiving others' bodies, rather than involving a general visual mechanism for gestalt completion, might involve special-purpose mechanisms for gestalt completion distinct from those used in perceiving inanimate objects. Specifically, we hypothesised that human observers could rely on the representations of their own bodies for filling in the gaps in their percepts of largely occluded conspecifics (e.g. a person sitting at a table behind a newspaper).

Gestalt completion and different types of body representation

Based on lesion studies, Coslett and colleagues (e.g. Coslett, Buxbaum, & Schwoebel, 2008; Medina, Jax, & Coslett, 2009) defined the body schema as a continuously updated, dynamic representation of body part locations based on proprioceptive and efference copy information. Accordingly, the body schema is a dynamically updated representation involved in monitoring the body's state (i.e. posture) and in controlling action execution, suggesting a representational continuum that ranges from static postures to dynamic actions. This makes the body schema a prime candidate for subserving the dynamic representational mapping between self and others according to the requirements of the situation (Buxbaum, Giovannetti, & Libon, 2000; Buxbaum, Kyle, & Menon, 2005; Fogassi & Luppino, 2005; Meltzoff & Moore, 1994; Reed & Farah, 1995).

The concept of a body schema is best understood in contrast to a "structural body representation" and a "body image" that have also been defined based on lesion studies (but see Corradi-Dell'Acqua, Tomasino, & Fink, 2009; for a review, see Schwoebel & Coslett, 2005). In autotopagnosia patients seem to lack knowledge of the structural spatial relations between body parts. Their abilities to point to a specific part of the body when prompted by its name or by a picture are equally impaired (e.g. Buxbaum & Coslett, 2001). Structural representations of the human body seem to be derived primarily from visual input and prior visual experience (Buxbaum & Coslett, 2001; Corradi-Dell'Acqua et al., 2009; Sirigu, Grafman, Bressler, & Sunderland, 1991) and frequent visual exposure to a particular posture (e.g. "typing on a keyboard") may enhance particular structural relations as well as generate conceptual or semantic representations, i.e. a "body image" (Schwoebel & Coslett, 2005).

In some cases body-gestalt completion might be primarily supported by these visual-structural or semantic types of body representations, which would be in agreement with a domain-general view of gestalt completion based on prior visual experience. Although visual representations of bodies could be a specialised subclass (e.g. Downing, Jiang, Shuman, & Kanwisher, 2001 - congruent to the debate regarding specialised visual representations for faces in contrast to other objects; Gauthier, Skudlarski, Gore, & Anderson, 2000; Kanwisher, 2000; Tarr & Gauthier, 2000), perceiving and completing them into gestalts would still involve purely visual mechanisms much like those involved in perceiving inanimate objects. In fact, such mechanisms might not be able, by themselves, to distinguish a realistic sculpture from a real body. Importantly, we challenged this domain-general view and set out to test the hypothesis that there is a specialised body-gestalt completion mechanism based on the internal body schema. By definition, any such mechanisms could be recruited only for the processing of bodies and not for any other visual object.

This hypothesis extends a growing body of findings showing that humans and other primates "resonate" with their conspecifics at an automatic, implicit level by mapping their own repertoire of movements and postures onto their perception of others. Resonance in this context is typically defined as involving a 1:1 mapping between the overt or covert (internal) imitation behaviour of the observer and the model (Rizzolatti, Fadiga, Fogassi, & Gallese, 1999). Movement- or motor-resonance was investigated predominantly in the context of the so-called mirror system, typically using imitation or observation tasks that showed limbs in isolation (Bertenthal & Pinto, 1994; Biermann-Ruben et al., 2008; Brass, Bekkering, Wohlschlager, & Prinz, 2000; Catmur, Walsh, & Heyes, 2007; Iacoboni et al., 1999; Jonas et al., 2007; Kessler et al., 2006; Rizzolatti & Fabbri-Destro, 2010). In contrast, other research has emphasised the importance of the overall posture or gestalt perceived in conspecifics (Amorim, Isableu, & Jarraya, 2006; Johansson, 1973; Kessler, Gordon, Cessford, & Lages, 2010; Kessler & Thomson, 2010; Lestou, Pollick, & Kourtzi, 2008; McKay et al., 2011; Meltzoff & Moore, 1994; Reed & Farah, 1995; Saygin, 2007). Returning to our initial example of people being occluded by a table and a newspaper in a café, we set out to investigate whether the body schema - as the likely substrate of motor- and posture-resonance (e.g. Buxbaum et al., 2005; Fogassi & Luppino, 2005) – would contribute to the early stages of visual perception of conspecifics, and, in particular, if it would do so when conspecifics' bodies were largely occluded.

The current study

To test our hypothesis that motor- and posture-resonance contribute towards body gestalt completion (and not only conceptual visual knowledge of how bodies look like), we developed novel experimental stimuli shown in Figures 1, 3, 5, (and A in the Appendix). The principle idea was to present a face and two hands in varying configurations that could be perceived as largely occluded bodies in varying postures (cf. Ratings, see Appendix). Considering only the visual input (cf. Figs. 1, 3, 5 and A), the process could resemble figure completion in classic gestalt stimuli such as the Kanizsa figures (Kanizsa, 1974). However, our hypothesis was that, unlike classic gestalt phenomena or gestalt completion of inanimate objects, body-gestalt completion would be an embodied process and not a purely visual process. Importantly, the stimulus configuration of the hands in relation to the head was varied in our experiments such that in one case the head-hands relationship was congruent with the participant's posture, and in the other case it was not. Our prediction was that posture congruence would facilitate self-other resonance based on the body schema and would, consequently, facilitate body-gestalt completion.

In Experiments 1 and 2 posture congruence, or posture match (we use "posture congruence" and "posture match" interchangeably in this manuscript) was manipulated by keeping hand orientation the same across stimulus conditions while changing the location of the hands in relation to the head, such that in one case the head-hands relationship matched the participant's posture, and in the other case it did not. In Experiment 3 the location of the hands on the screen was kept constant while the orientation of the hands was changed in such a way that two different postures in relation to the head emerged that were either congruent or incongruent with the participant's posture.

We measured the effect of posture congruence (between stimulus and participant) on speed and accuracy of imitative responses to intransitive finger movements shown onscreen. Thus, motor-resonance could also contribute towards gestalt completion. In addition to these finger movements performed by one of the stimulus hands, we employed gaze cues to optimally engage observers with the face stimulus (Frischen, Bayliss, & Tipper, 2007 for a review) while directing their attention to one of the stimulus hands (Bayliss, Bartlett, Naughtin, & Kritikos, 2011). Thus, we exploited the fact that hand motion and eye gaze are normally coordinated (Amano, Kezuka, & Yamamoto, 2004; Flanagan & Johansson, 2003; Gowen & Miall, 2006) in order to induce a connection between face and hand stimuli. We expected that the strength of this connection would depend on the efficiency of body-gestalt integration between face and hands; so for a stimulus configuration that could be easily completed into a body-gestalt, we expected gaze cuing to facilitate the imitative finger responses. To re-iterate, our main prediction was that "easy" body gestalt completion would not only depend on the visual configuration of body parts onscreen in relation to prior visual experience, but would be significantly modulated by the relation of that configuration to the instantaneous posture of the participant.

Temporal dynamics of posture/motor resonance

In order to disentangle possible contributions of posture- and motor-resonance we allowed for a 500 ms static stimulus display prior to the onset of the gaze cue. Additionally, in Experiment 1 we varied the SOA between gaze cue and finger movement (17 ms vs. 500 ms) allowing us to understand the temporal dynamics of gestalt completion. That is, gestalt completion might already be facilitated by posture congruence in relation to the static stimulus prior to gaze and finger movements. If that were the case, we would expect two things. First, we should observe a facilitatory effect of a congruent posture on imitative responses at a very short SOA of 17 ms. Second, the facilitatory effect of gestalt completion in relation to the static stimulus should take precedence over subsequent manipulations of gaze, such as validity (Expts. 1-3), and of finger movements, such as movement-direction and -type (Expt. 3).

In all three Experiments we varied the validity of gaze cues. Two alternative predictions were possible regarding their effect. First, validly cued movements might be particularly facilitated by a congruent posture, while invalidly cued movements might be particularly surprising and suffer an extra cost, hence, resulting in opposite effects of posture congruence depending on gaze cue validity. Alternatively, body-gestalt completion, aided by posture congruence, might provide more flexibility with respect to predicting another's actions, allowing fast responses to validly cued movements, but also enabling quick shifts of attention when movements are invalidly cued. The second prediction is compatible with the assumption that posture resonance in relation to the static image could already facilitate gestalt completion, thus taking precedence over subsequent gaze cue validity.

As already mentioned, in Experiment 3 the orientation of the hands was changed while the location of the hands on the screen was kept constant. The change in orientation necessitated additional controls in terms of movement-direction and -type which will be described in detail in the context of Experiment 3. Here it is important to note that on half of the trials the finger response of the participant was congruent in direction or type with the finger movement on screen, while in the other half it was incongruent. Thus, posture congruence could be modulated by the congruence of movement-direction or -type, resulting in a posture congruence effect only when type or direction would be congruent too. However, if posture congruence was indeed as important as we claimed, and if it was indeed established already in relation to the static stimulus, then it would precede manipulations of the finger movement and would be expected to dominate the effects of movement-type/-direction in the statistical analysis. Experiment 3 was a strong test for our predictions, since Brass et al. (2001) had previously reported significant congruence effects of movement-type and -direction for imitative responses to hands presented in isolation. Nonetheless, we expected that the implicit body context in our study along with the proposed posture resonance mechanism for body-gestalt completion would prioritise posture congruence effects.

Our line of reasoning regarding the precedence of posture congruence over gaze cue validity and movement-type/-direction congruence appears plausible and socially relevant when taking into account competitive social interactions. Deception is quite common in these situations and the survival benefit of feigning the onset of a particular movement while actually performing another is obvious when thinking of person-to-person combat. An important cue that can be used to misdirect the adversary's attention is gaze (Frischen et al., 2007). At the same time, humans, who were able to recover quickly from such deceptive cues and did not fall easily for deceptive movements in general, would have had an additional survival advantage. Thus, representing another as an integrated body-gestalt could provide more flexibility in anticipating the other's actions by tracking alternative options simultaneously, hence, allowing quicker recovery from deceptive cues. In the present study we expected this mechanism to enable attentional switching in relation to invalid gaze cues (Expts. 1-3) and to aide correct responses to finger movements onscreen even if these were partially incongruent (whether in direction or type; Expt. 3).

Hypotheses

To summarise, we hypothesised that perceiving the largely occluded body of a conspecific would be a special case of gestalt completion. This is at odds with a domain-general view, where all partially occluded visual objects would be completed into gestalts based on prior visual experience. While we conceded that knowing 'how bodies look' could be of great importance for body-gestalt completion, we proposed that bodies afford an additional mechanism that is not shared with inanimate objects. Specifically, we suggested that our implicit knowledge of 'how our own body feels and moves' could substantially aide our visual perception of largely occluded bodies of others. In other words we proposed an embodied body gestalt (eBG) completion mechanism. This body-specific knowledge was identified as the body schema, which we expected to be involved in the perception of largely occluded bodies, particularly where the participant's instantaneous posture was congruent with the posture implied by the stimulus.

Regarding more specific aspects of the predicted posture congruence effect, we expected posture resonance to take effect in the static stimulus period and so prior to gaze and finger movement onsets. Posture congruence was therefore expected to take priority over manipulations of gaze (validity) and of finger movement (-direction or -type). This would result in response facilitation due to posture congruence for valid as well as for invalid gaze cues. It would also result in posture congruence being statistically relevant, even disregarding congruence in finger movement-direction or -type. Such a prioritisation of the overall gestalt over body-part particulars would suggest a mechanism for representing a conspecific as an integrated body-gestalt based on the observer's own body schema which would allow anticipating several of the conspecific's options to act simultaneously. A confirmation of our predictions would therefore highlight the importance of the proposed eBG mechanism for social interaction and survival in general.

Experiment 1

Methods

Participants

20 right-handed volunteers (15 female, average age 23.9) were recruited via a departmental subject pool and were either paid for their participation or received course credits.

Apparatus

The experiment was run in a dark room using EPrime® 2.0 on a Windows® PC with a 19'' 60 Hz CRT monitor set to a resolution of 1024x768 pixels. A chinrest ensured a constant viewing distance of one meter throughout the experiment.

Stimuli

The stimuli are shown in Figure 1 and contained one face and two hands that covered an area of 470 (width) by 568 (height) pixels centred on the 1024x768 screen. All trials employed the same face and hands shown in Figure 1, with the two hands being either above or below the face with their palms-up. The overall configuration could either be perceived as a person carrying a log (hands below the face) or as a person about to perform an underhand chin-up (hands above the face). Participants were always sitting in an upright position with their hands palms up resting on a table, using an upside-down keyboard for their imitative responses. Thus, their posture was congruent with the posture implied by the first stimulus configuration ('as if carrying a log') and we expected faster responses for this stimulus condition.

Although the stimulus configurations employed across our three experiments could be associated with particular intentions or situations (e.g. palms turned upwards as if carrying a log, hands raised above the head as if preparing for an underhand chin-up or an overhand pull-up, hands crossed in front of the body as if playing the piano with crossing of hands, etc.), these associations where rather loose and our stimuli did not involve any objects that would solidify a semantic interpretation of the postures implied by the stimuli. We expected this to enhance the contributions of posture and motor resonance over semantic posture understanding (cf. Goldenberg & Karnath, 2006; Schwoebel & Coslett, 2005).

To understand possible differences in familiarity with particular postures we asked 14 students at the University of Glasgow to rate the two stimulus configurations in this Experiment 1 and their implied postures together with the stimuli employed in the other two experiments and alongside anatomically impossible configurations that showed either two left or two right hands (either above or below the face) on a scale from 0 (very implausible) to 100 (very plausible) with respect to the plausibility of the implied posture (see Appendix for details). The two stimulus postures employed here were rated as significantly more plausible than their anatomically impossible counterparts (Wilcoxon tests, both $p < .001$), but did not differ significantly from each other ($p = .35$).

Procedure

Participants could determine when each trial began. Each trial started with a fixation cross (grey on a black screen) for 500 ms that indicated the position of the face in the upcoming stimulus (i.e. top or bottom). Participants were instructed to keep their eyes on the cross and, then, on the face. The face and the two hands were presented all at once and remained static for another 500 ms. Then a gaze cue was presented (see Figure 1): the face looked at one hand. While the gaze cue remained, there was a movement of either the cued or the uncued hand: the index or the little finger of this hand performed a tapping movement that lasted ~267 ms (8 frames at 2 screen refreshes each). Figure 1 shows examples of the maximum finger deflections. Participants were instructed to immediately imitate each finger movement as if seen through a mirror (specular imitation with 1:1 spatial mapping), thereby pressing the key underneath their finger (on the upside-down keyboard). After this participants were asked to indicate by means of another key press in which direction the face had looked (left or right). This ensured that participants engaged with the (unpredictive) gaze cue. Acoustic feedback was given for accuracy on both responses.

Prior to the test trials participants received two practice sessions. First, they practiced specular imitation of index and little finger movements with only the two hands present (6 trials). Second, they practiced the whole procedure as described above (a further 6 trials).

Insert Figure 1 about here

Dependent Measures and Design

Reaction times (RT, in msec) and accuracy (ACC, in % correct) of the imitative responses were recorded as the dependent variables. We anticipated that RT would be the more important measure of performance as the tasks were very simple and we expected ACC to be close to ceiling level. We nevertheless chose to report ACC in order to exclude the possibility of speed-accuracy-trade-offs. For RTs we only analysed correct responses and we employed medians for representing the individual performance in each cell of the design to reduce distortions by outliers. The three independent variables that we included as factors in the ANOVA design were the factors "stimulus posture" (hands below vs. hands above the face), "validity" (valid vs. invalid gaze cues), and "SOA" (17 ms vs. 500 ms between gaze-cue onset and finger-movement onset). We collapsed across movement side (left/right) and across fingers (index/little finger). There was a total of 64 experimental trials (8 in each cell of the ANOVA design).

Hypotheses

As described in the Introduction, we expected imitative responses to be facilitated (faster) when the hands were displayed below the face stimulus, because the participants' body posture with their hands resting palms-up on the table was congruent to the posture implied by that stimulus configuration. We therefore expected a "stimulus posture" effect, where stimuli with the hands below the face (implied posture: 'as if carrying a log') were processed faster (and more accurately) than stimuli with the hands above the face (implied posture: 'as if preparing for an underhand chin-up'), indicating that posture congruence facilitated body-gestalt completion.

Furthermore, the possibility of facilitated body-gestalt completion raised a question regarding the effect of gaze-cue validity. To re-iterate, two alternatives were possible regarding this effect. Firstly, a consistent body-gestalt perceived in the model might bias predictions towards expecting movements that are validly cued by gaze, facilitating these, while invalidly cued movements would be surprising and suffer a cost. If that was the case we would expect a strongly attenuated or even reversed posture congruence effect for invalid trials: slower/less accurate responses with a congruent compared to an incongruent posture (i.e. between stimulus and participant). Alternatively, a consistent body-gestalt might provide more flexibility with respect to predicting another's actions, allowing fast responses to validly cued movements, but also enabling quick shifts of attention when movements are invalidly cued. If that was the case we would expect that the advantage for a congruent over an incongruent posture would be observed for valid as well as for invalid trials.

Finally, an effect that would emerge already at the short SOA of 17ms (between the onset of the gaze cue and onset of the finger movement) could be regarded as support that gestalt completion was accomplished very quickly, most likely in relation to the static image at the onset of the trial.

Insert Table 1 about here

Results

Descriptive statistics for RT and ACC data are provided in Table 1. Separate ANOVAs were conducted for RTs and ACCs, respectively, and included the factors "stimulus posture" (hands below vs. hands above the face), "validity" (valid vs. invalid gaze cues), and "SOA" (17 ms vs. 500 ms between gaze-cue onset and finger-movement onset). As can be observed in Figure 2, Panel A, the main effect of SOA ($F(1,19)= 51.7$, $p<.0001$, $\eta^2_p=.731$) and the main effect of validity ($F(1,19)= 8.7$, $p=.008$, $\eta^2_p=.314$) reached statistical significance in the analysis conducted with RTs. Responses were faster at 500ms SOA and for validly cued finger movements. Most importantly the main effect of stimulus posture was significant too ($F(1,19)= 14.8$, $p=.001$, $\eta^2_p=.438$) (all other p>.1). Response times were faster when the participant's posture and the posture implied by the stimulus were congruent (i.e. hands below the face).

Insert Figure 2 about here

ACC data (Figure 2, right) revealed a significant ordinal interaction between "stimulus posture" and "validity" ($F(1,19)= 5.03$, $p=.037$, $\eta^2_p=.209$; see Figure 4, Panel B) besides a significant main effect of "stimulus posture" ($F(1,19)= 11.26$, $p=.003$, $\eta^2_p=.372$). Simple effects revealed (see Figure 2, Panel B) that the interaction was due to significantly lower performance on invalid trials compared to valid trials only when the participant's posture and the posture implied by the stimulus were incongruent (i.e. hands above the face: $p = .027$). In contrast, no significant difference ($p = .45$) between valid and invalid trials was observed when the postures were congruent (possibly due to performance reaching ceiling level in valid trials). The particularly attenuated performance in incongruent invalid trials was further confirmed by a significant difference to congruent invalid trials ($p = .002$), while congruent and incongruent trials did not significantly differ in the valid condition ($p = .2$).

Discussion

The significant stimulus-posture effect in Experiment 1 provides a first hint that "body-gestalt" completion is not merely visual but can be mediated by representations of the body, i.e. by posture/motor resonance. RTs were faster and ACC higher for the hands-head stimulus configuration that was congruent with the participant's posture. It seems likely, further, that the generation of a body-gestalt was rapid and possibly automatic given the significant effects at the shortest SOA (17ms). Finally, gestalt completion seemed to allow for more efficient attentional cueing (valid trials), yet, also for more efficient attentional switching (invalid trials) compared to trials where gestalt completion was hampered due to postural incongruence (cf. Hypotheses).

A further experiment was however needed to address a potential objection. Hands that formed the more effective body-gestalt stimulus together with the face were always displayed in the lower visual field while the hands in the less effective body-gestalt were always displayed in the upper visual field. Thus, however unlikely, the reported results could have been an effect of upper vs. lower visual field differences instead of differences in body-gestalt completion.

Furthermore, despite similar plausibility ratings obtained for the two stimulus postures (cf. Appendix), the incongruent stimulus posture with the hands above the head is probably observed less frequently in everyday social situations than the congruent stimulus posture (hands below the head). Hence, the observed body-gestalt completion effect may have been boosted by visual familiarity in addition to or even instead of posture congruence.

The strongest supporting evidence for the involvement of dynamic body schema representations in body-gestalt completion would be obtained by changing the participants' posture online during the experiment, revealing a direct effect of 'instantaneous' posture changes on body-gestalt completion.

In Experiment 2 we changed the participant's posture randomly across trials. The two postures could match either of those implied by the stimuli in Experiment 1 (see Figure 3): 1) hands raised above their head ('as if preparing for an underhand chin-up') and 2) hands resting on the table ('as if carrying a log'). By alternating posture congruence between both stimulus configurations, the potential alternative explanation in terms of upper and lower visual field could be addressed as well.

Experiment 2

Methods

Participants

24 right-handed volunteers (17 female, average age 22.17) participated in this study.

Materials and Procedure

In this Experiment we varied the participants' posture to induce congruence with either of the two postures implied by the stimuli in Experiment 1 (see Figures 1 and 3). Before each trial, the participant was either instructed to sit with their hands palms up, resting on the table or else instructed to raise their hands (palms up) above their head by resting their upper arms on a specifically build arm rest that was attached to the chin rest (see Fig. 3, middle). Participants responded by means of response pads attached to each hand (shown in Fig. 3, right). This allowed for free movements of the hands and arms according to the posture instructions, while the response fingers (index and little finger) always remained on top of the correct response keys. Participants were instructed about their posture randomly before each trial, creating either a posture match with the stimulus configuration where the hands were below the face ('carrying a log') or with the configuration where the hands were above the face ('underhand chin-up'). In order to maintain a comparable number of trials despite the newly introduced factor "participant's posture", we only employed an SOA of 17 ms in this experiment.

Insert Figure 3 about here

Hypotheses

We expected posture congruence between participant and visual stimulus to facilitate imitative responses regardless of the visual field in which the hands were displayed. Hence, in agreement with Experiment 1, when participants rested their hands palms up on the table (cf. Fig. 1) we expected to observe faster response times for stimuli with hands below the face ('carrying a log') compared to stimuli with hands above the face ('preparing for a chin-up'). However, extending Experiment 1, we expected these effects to be reversed when participants raised their hands above their heads as if preparing for an underhand chin-up (see Fig. 5, middle). We therefore predicted an interaction between "stimulus posture" (hands below vs. above the stimulus face) and "participant's posture" (hands below vs. above the participant's head). For simplicity we propose labelling this interaction as the "embodied Body-Gestalt" effect (eBG): informally, it is the facilitative effect of congruence between participant and the stimulus posture on gestalt completion. As in Experiment 1 we also expected to observe the eBG effect for both valid and invalid trials, supporting the notion that embodied gestalt completion was accomplished before the onset of the gaze cue and could facilitate expected processing (valid trials) or require attentional switching (invalid trials).

Insert Table 2 about here

Results

Descriptive statistics for RT and ACC data are provided in Table 2. Separate ANOVAs were conducted for RTs and ACCs, respectively, and included the additional factor "participant's posture" (hands below vs. above the head) along with the previously employed factors "stimulus posture" (hands below vs. above the face) and "validity" (valid vs. invalid gaze cues). For RTs, the main effects of "participant's posture" ($F(1,23)= 6.03$, $p=.022$, $\eta^2_p=.208$) and "validity" ($F(1,23)= 5.75$, $p=.025$, $\eta^2_p=.199$) reached significance together with the interactions between "participant's posture" and "validity" ($F(1,23)= 7.29$, $p=.013$, $\eta^2_p=.24$).

Most importantly, the interaction between "participant's posture" and "stimulus posture" reached significance too ($F(1,23)= 37.6$, $p<.0001$, $\eta^2_p=.62$), revealing the strongest effect size within the general linear model (see Figure 4, Panel A). All simple effects of this interaction reached significance ($p < .05$; cf. Fig. 4), which directly supported our hypothesis of an eBG effect. We had predicted that a congruent posture between participant and stimulus configuration onscreen would always generate the fastest RTs, implying that the RT disadvantage for a particular stimulus configuration (e.g. hands above the face) would reverse when the participant's posture was changed to match its implied posture (i.e. participants raised their hands above the head as well).

Insert Figure 4 about here

This result was mirrored in the analysis of ACC data (see Fig. 4, Panel B), where a significant interaction between "participant's posture" and "stimulus posture" was also observed ($F(1,23)= 11.6$, $p=.0024$, $\eta^2_p=.335$). A congruent posture between stimulus and participant produced more accurate data than when the postures were incongruent (for significance of simple effects see Fig. 4, Panel B). In addition, main effects of "stimulus posture" ($F(1,23)= 5.38$, $p=.029$, $\eta^2_p=.189$) and "validity" ($F(1,23)= 14.18$, $p=.001$, $\eta^2_p=.381$) were also observed (all other $p > .1$).

Discussion

The most important finding in Experiment 2 was that in addition to replicating the "stimulus posture" effect observed in Experiment 1 (left half of the top graph in Fig. 4) we were able to fully reverse the direction of the effect when the participant's posture was changed to match the stimulus configuration with the hands above the face. Based solely on Experiment 1 we could not decisively conclude whether the body-gestalt completion effect was truly embodied or partly or fully mediated by higher visual familiarity with the congruent posture. The observed interactions (RT and ACC) between the participant's and the model's posture in Experiment 2 suggested that the posture congruence advantage could indeed originate from embodied processes, since this advantage was also observed for the less visually familiar posture. This embodied body-gestalt (eBG) effect ruled out any possible differences in visual posture familiarity and between visual fields, supporting our hypothesis that online posture resonance between stimulus and participant significantly contributed to perceiving conspecifics

However, there is a potential alternative explanation, which we aimed to address in the final Experiment 3. Given our findings so far it is possible that the congruence between the stimulus hands and the hands of the participant was established not via a postural match, but via congruence in locations coded in relation to landmarks of the environment. Specifically, the position of the participant's and model's hands may have been coded in relation to the top and bottom of the screen, so that the match in relative positions might in principle account for any congruency benefit. That is, participants may have been faster to respond to hands displayed at the top of the screen with their own hands raised, simply because they perceived their own hands as being spatially closer to the top of the screen as well. In Experiment 3 we addressed this alternative explanation by keeping the location of the hands constant (always below the face), while changing their orientation in order to induce two different postures ('as if carrying a log' vs. 'as if playing a piano with hands crossed'). This manipulation cannot be resolved via environment-related spatial coding, since the locations are the same. Thus, if the eBG effect would persist, posture/motor resonance would be the simplest and most consistent explanation across all three experiments.

Experiment 3

Methods

Participants

24 volunteers (23 right-handed, 16 female, average age 22.1) participated in this study.

Materials and Procedure

In this final Experiment 4 we set out to replicate the eBG effect – formally, the interaction between “stimulus posture” and “participant’s posture” observed in Experiment 2 - with a different posture alteration. We again employed the ‘palms-up’ posture for stimuli and participants (as if ‘carrying a log’; cf. Expts. 1 and 2) but this time as a second posture we used crossed hands (with palms down) rather than raised hands (cf. Figure 5). By keeping the location of the hands constant (always below the face) while changing their orientation for inducing two different postures (‘as if carrying a log’ vs. ‘as if playing a piano with hands crossed’), a difference in environment-related spatial coding was no longer possible.

The posture with crossed hands was chosen because it maintains finger mapping while changing hand orientation. The locations of the fingers within the hand map exactly onto the finger locations within the hand posture with palms up (cf Fig, 5). This was essential for avoiding a spatial mapping mismatch between the moving finger on screen and the participant’s response finger (i.e. little finger mapping onto index finger and vice versa), particularly when the postures of the stimulus and the participant were incongruent. This would have created a confound between mapping mismatch and postural incongruence. (Similar considerations led to the exclusion of a large number of hand/body postures and posture combinations.)

Insert Figure 5 about here

As shown in Figure 5, in the palms-up posture (of stimulus and participant) a finger tap always moved upwards, while in the crossed-hands posture it always moved downwards. Thus, when postures differed so did the tapping directions. Therefore, congruence between the participant's response direction and the movement direction of finger movements onscreen had to be controlled in an additional step. This was achieved by including finger movements in both directions on the screen (i.e., upwards as well as downwards in both stimulus configurations, cf. Fig. 5). By doing so, half of all trials displayed a finger movement in the same direction as the response direction of the participant - who always pressed a key - while the other half of the trials showed a movement in the opposite direction (cf. Table 3).

Insert Table 3 about here

Finally, including finger movements in both directions as a control also introduced an additional type of finger movements: finger lifts in addition to taps. Importantly, only finger taps matched the participant's keypress responses in type, resulting in a complex relationship between movement-direction, movement-type, and postures as shown in Figure 5 and in Table 3 (details of the employed design are shown in Table 3). Thus, congruence of movement-direction and congruence of movement-type were confounded in our design and we addressed this issue by conducting two separate analyses. The first analysis included "movement-direction congruence" as a design factor, while the second included "movement-type congruence" (cf. Brass et al., 2001).

Hypotheses

First, given our proposal that perceiving conspecifics involves embodied body-gestalt completion (eBG completion), we expected to replicate the eBG effect in form of a two-way interaction between "stimulus posture" and "participant's posture" independently of "movement-direction congruence" or of "movement-type congruence". Alternatively, if eBG completion was at work but of no more importance than congruence in movement direction or type, then the eBG effect would be modulated by "movement-direction congruence" or "movement-type congruence" (or both), resulting in three-way interactions ("stimulus posture" x "participant's posture" x "movement-direction/-type congruence"). Finally, if the "stimulus posture" x "participant's posture" interaction could not be replicated at all, then our previous effects may have been the result of congruent spatial coding in relation to the screen and may have not been a reflection of embodied processing.

As in the two earlier experiments we also expected to observe an eBG effect for both valid and invalid trials, further supporting the notion that embodied gestalt completion was accomplished before the onset of the gaze cue and could facilitate expected processing (valid trials) or require attentional switching (invalid trials).

Finally, a response posture with crossed hands is perceived as particularly difficult and not only affects the integration of somatosensory and visuo-spatial frames of reference (Riggio, Gawryszewski, & Umilta, 1986) but also significantly modulates motor responses that are spatially mapped to visual stimuli (Hommel, 1993; Simon, Hinrichs, & Craft, 1970). Concordantly, a statistical trend (p = .055) was observed for the rating difference (see Appendix) between the posture with crossed hands (average = 63) and the posture with palms up (average = 89). We therefore expected a main effect of participant posture; in particular, we expected participants to respond more slowly with crossed hands. We were also apprehensive that the difficulty involved in responding with crossed hands might mask the eBG effect. Given this possibility, a successful replication of the eBG effect would underpin the involvement of the body schema in body-gestalt completion.

Insert Table 4 about here

Results

Descriptive statistics for RT and ACC data in each design cell are provided in Table 4. Two separate ANOVAs were conducted for RTs. These included either the control factor “movement-direction congruence” or “movement-type congruence” (congruent vs incongruent between stimulus and participant) in addition to the previously employed factors: “participant’s posture” (here: hands palms up vs. hands crossed), “stimulus posture” (here: hands palms up vs. hands crossed), and “validity” (valid vs. invalid gaze cues).

Importantly, neither “movement-direction congruence” nor “movement-type congruence” reached significance for RTs - neither as main effects (p = .25; p = .74) nor as three-way interactions together with “participant’s posture” and “stimulus posture” (pc = .74; p = .25). In contrast, the main effects of “participant’s posture” ($F(1,23)= 112.7, p<.0001, \eta^2_p=.831$) and “validity” ($F(1,23)= 13.4, p=.001, \eta^2_p=.369$) reached significance. As was to be expected, participants responded significantly faster with their palms up than with crossed hands and valid trials elicited faster responses than invalid trials. Most importantly, however, the interaction between "participant's posture" and "stimulus posture" was significant again ($F(1,23)= 10.5$, $p=.003$, $\eta^2_p=.314$). All simple effects of this interaction reached significance ($p < .05$; cf. Fig. 6).

Insert Figure 6 about here

The pattern of results (Figure 6, Panel A) therefore supported our main hypothesis that a congruent posture between participant and stimulus configuration onscreen would always generate the fastest RTs, implying that a RT disadvantage for a particular stimulus configuration (e.g. hands crossed) reversed when the participant's posture was changed to match it (i.e. participant responding with crossed hands).

The ACC data (Fig. 6, bottom) revealed a main effect of "participant's posture" ($F(1,23)= 5.38$, $p=.029$, $\eta^2_p=.189$), confirming that it was harder for participants to respond with crossed hands. The interaction between stimulus posture and validity also reached significance ($F(1,23)= 7.6$, $p=.011$, $\eta^2_p=.249$; all other $p > .1$). Finally, we show the interaction between participant and stimulus posture in Figure 6, Panel B, since the pattern was congruent with the RT effect (Panel A) although it did not reach significance ($p > .1$).

Discussion

Overall Experiment 3 corroborated Experiments 1 and 2 allowing for the following conclusions. First, we can rule out the possibility that congruent spatial encoding of the participant's and the model's hands in relation to the environment (e.g. screen) might have accounted for the effects in Experiments 1 and 2. Second, in the light of the quite dramatic change in proprioception introduced by means of the crossed-hands posture as well as the required additional control factors “movement-direction congruence” and “movement-type congruence”, we can conclude that there is compelling evidence for embodied contributions to body-gestalt completion. Finally, Experiment 3 also sheds light on possible hierarchies among such embodied contributions. Brass et al. (2001) used finger movements of hands shown in isolation (no body context) to investigate the impact of movement-direction and -type. They reported that both factors affected the speed of imitative responses, while in our Experiment neither direction nor type mattered statistically. Thus, we suggest that posture resonance took priority in this particular experimental context where hands and a head could be completed into body gestalts; of course, other differences in setup like the employed SOAs and the number of fingers that could move, may have also played a role. This line of reasoning is discussed in more detail below.

General Discussion

Conform to our hypotheses (see Introduction) we have shown that human observers rely on representations of their own bodies to rapidly fill in gaps in their percepts of largely occluded conspecifics.

With the findings in Experiment 1 we provided initial evidence that participants were indeed able to integrate our drastically reduced stimuli into body-gestalts. We obtained evidence that the facilitative effect of posture congruence on gestalt integration peaked very quickly (at 17 ms SOA) and allowed for fast attentional cueing (valid trials) as well as fast attentional switching (invalid trials). The latter was evidenced by two main effects of “validity” and “stimulus posture” and by the absence of a disordinal interaction (Fig. 2). This pattern of results was replicated and strengthened in Experiment 2, where we changed the participants' posture randomly across trials, inducing posture congruence with either of two stimulus configurations and showing that speed and accuracy of responses was increased when the posture of the participant matched the posture implied by the stimulus. This ruled out visual posture familiarity as an alternative explanation.

In Experiments 1 and 2 posture congruence was manipulated by keeping the hand orientation the same across stimulus conditions while changing the location of the hands in relation to the head. In Experiment 3 the location of the hands onscreen was kept constant while the orientation of the hands was changed in such a way that two different postures in relation to the head emerged This manipulation addressed the possibility that congruence between the stimulus hands and the hands of the participant was not established via a postural match but via congruence in locations coded in relation to the screen or any other environmental landmarks. By keeping the location of the hands constant (always below the face) a difference in environment-related spatial coding (e.g. in relation to the top vs. bottom of the screen) was no longer discriminative. While alternative explanations for each single experiment could still be conceived of, we suggest that embodied resonance is the simplest and most consistent explanation for the overall pattern of results. We propose labelling the observed effect as the "embodied body-gestalt" (eBG) effect.

Altogether our results also suggest that "body-gestalts" were completed very early, most likely in relation to the static stimulus, i.e. before the onset of the gaze cue and the subsequent finger movement. Based on the overall pattern revealed in our series of experiments - the lack of movement-direction/-type effects (Expt. 3), the lack of any disordinal interaction between eBG effect and validity (Expts. 1, 2, and 3), and the strongest eBG effect at 17 ms SOA (Expt. 1) - we conclude that completion of body gestalts involves early posture matching. We propose that eBG completion based on posture matching is initiated upon the presentation of the static stimulus configuration in anticipation of the subsequent gaze and finger movements, with the consequence that posture takes priority over the particulars of these subsequent movements such as “validity” and “movement-type” or “movement-direction”. This underpins the role of body-related representations in body-gestalt completion, but the exact nature of these bodily representations needs to be determined.

Possible body representations underlying body-gestalt completion

As described in the Introduction, at least three different types of bodily representations have been proposed in the literature. Functional dissociations in lesion studies have motivated distinguishing “structural body representation”, “body image”, and “body schema” (but see Corradi-Dell'Acqua et al., 2009; for a review, see Schwoebel & Coslett, 2005).

Deficient structural or topological body representations are commonly observed in autotopagnosia, where patients seem to have lost their knowledge of the spatial relations between their body parts and are equally impaired at pointing to a specific body part when prompted by its name or by a picture (e.g. Buxbaum & Coslett, 2001). Structural representations of the human body seem to be primarily derived from visual input (Buxbaum & Coslett, 2001; Corradi-Dell'Acqua et al., 2009; Sirigu et al., 1991), hence, such representations could have played a role in our experiments. In fact, our conclusion that posture resonance might already occur in relation to the initial, static period of the stimulus is compatible with the involvement of such a structural, topological body representation.

Repeated visual exposure to a particular posture could enhance structural representations compared to infrequent postures, possibly even resulting in a conceptual or semantic representation. An example would be a frequently observed posture such as ‘typing on a keyboard’, for which we indeed found the highest ratings in terms of posture plausibility (see Appendix, Fig. A). Thus, in some cases eBG completion could be further supported by this second, semantic type of body representation, commonly referred to as a “body image” (Schwoebel & Coslett, 2005).

However, the observed eBG effects cannot be explained by structural and/or semantic body representations: In Experiments 2 and 3 we changed posture alignment on a trial by trial basis and observed robust eBG effects. Although the effects occurred at a very short SOA (17ms), suggesting gestalt completion in relation to the static image, the effects depended on the congruence in posture between the ‘self’ and the ‘other’. This can only be explained by means of a dynamic body representation that quickly adjusts to changing relationships between an observer’s and another’s body.

The notion of a body schema refers to such a dynamic body representation and is thought to rely on proprioception and efference copies during action execution in addition to online visual input about the (moving) body (Schwoebel & Coslett, 2005; Wolpert, Ghahramani, & Jordan, 1995). A multisensory representation of this kind would allow for accurate posture estimates as well as for online posture- and motor-corrections. Maravita and Iriki (2004), reported representational plasticity in monkeys after extensive and active use of a tool (e.g. a handheld rake) and proposed enlarged body schema representations that encompassed the dimensions of the tool as explanation. This seems to confirm a tight relationship between the online representations of the body, proprioception, and the visually guided execution of actions. The body schema as a manifestation of this relationship has been therefore proposed to extend to self-other mappings (Buxbaum et al., 2005; Fogassi & Luppino, 2005), possibly allowing for the decoding of action intentions and for the prediction of future actions (Fogassi et al., 2005; Gallese & Goldman, 1998; Kessler et al., 2010; Kilner, Vargas, Duval, Blakemore, & Sirigu, 2004; Kourtis, Sebanz, & Knoblich, 2010; Ramnani & Miall, 2004; Wilson & Knoblich, 2005).

We conclude that although in some cases body-gestalt completion could be aided by visual-structural or semantic types of body representation - which would be in agreement with a domain-general view of gestalt completion – overall our results strongly suggest that bodies are a special case of gestalt completion, i.e. that body-gestalt completion uniquely involves a mapping of the observer's body schema onto perceived body parts.

Flexible hierarchies of embodied self-other mappings

The findings of Experiment 3 suggest that overall posture may have played a more essential role in our paradigm than finger movement kinematics. This suggests a hierarchy in body schema representations with overall posture taking precedence over effector kinematics. Specifically, posture congruence between participant and stimulus significantly explained the data pattern in Experiment 3, while an exact match in finger movement -direction or -type did not.

In contrast, Brass et al. (2001) reported that, for hands presented in isolation, imitation of finger movements depended significantly on congruence of movement type (lift vs. tap) and movement direction. While differences in setup, such as the employed SOAs and the number of fingers that could move, might play a role in explaining the discrepancy, the most striking difference between our study and Brass et al's is the lack of a body context (incl. gaze cues) in the latter. This allows us to reconcile the discrepant findings by proposing a flexible, context-dependent, hierarchical organisation of the body schema template humans employ for monitoring their own body and for mapping it onto others (cf. Meltzoff & Moore, 1994; Reed & Farah, 1995).

Specifically we propose that when a hand is presented in isolation and merely a finger movement is expected, the body template might 'zoom in' on hand kinematics and the specific mappings of these onto perceived hands (cf. Brass et al., 2001). But when the perceived body context is more complex and actions are more coordinated (e.g. gaze-hand coordination), then a match of the overall posture would be of advantage. Posture would provide an 'anchor' for the self-other mapping that would facilitate coordinated predictions of various possible actions (as well as deceptions) and would therefore carry more weight than the specific kinematic mapping of the fingers in relation to the hand. We propose that in the complex case the body schema would 'zoom out' of the hand-specific mapping and prioritise the overall gestalt. Extending previous notions suggested by others (Gallese, 2007; Goldenberg & Hagmann, 1997; Meltzoff & Moore, 1997; Reed & Farah, 1995; Reed & McGoldrick, 2007) we therefore propose that humans build a mental model of the people around them based on a multi-level body schema template of themselves. The multiple levels within the model could be focused upon dynamically as required: e.g. zooming in on effectors or zooming out on the overall gestalt depending on the context (for similar notions see Goldenberg & Hagmann, 1997; Goldenberg & Karnath, 2006; Goldenberg, Laimgruber, & Hermsdorfer, 2001).

Such a dynamic hierarchical model would predict another's actions with the necessary flexibility, particularly when unexpected actions occur. In social interactions deception is quite common. The survival benefit of feigning the onset of a particular movement while actually performing another is obvious when thinking of person-to-person combat. This still seems to be common practice in competitive sports like tennis, football, fencing, and all close combat sports. An important cue that can be used to misdirect the adversary's attention is gaze (Frischen et al., 2007) and we have used this feature here in the invalid trials. However, when participants were able to form an integrated "body-gestalt" or model of the person they had to imitate (congruent postures), switching attention away from the misleading gaze cue was more efficient compared to when gestalt integration was hampered (incongruent postures). Representing another as an integrated body-gestalt could therefore provide more flexibility in anticipating the other's actions by tracking alternative options simultaneously, hence, allowing quicker recovery from deceptive cues. Ability to recover quickly from being misled by a feigned movement or intentionally misdirected gaze could be part of the successful representation of an adversary and was probably a necessary survival skill of our ancestors. Such emphasis on cognitive flexibility rather than hard-wired resonance mechanisms coheres with recent theoretical proposals (Heyes, 2010; Jacob & Jeannerod, 2005; Mahon & Caramazza, 2008).

More generally, the eBG effect also suggests that self-other mappings could be hampered if the postures of two interacting people are strongly incongruent, possibly affecting higher levels of mental alignment as well (see also Garrod & Pickering, 2009; Pickering & Garrod, 2004 for discussions of multi-level alignment). This might explain why we tend to align our posture with other people during communication (Garrod & Pickering, 2009; Shockley, Santana, & Fowler, 2003) and social interaction (Chartrand & Bargh, 1999; Lafrance, 1985): Alignment at a basic, automatic level might facilitate higher-level aligned representations (Frith & Frith, 2007; Garrod & Pickering, 2009; Kessler & Rutherford, 2010; Kessler & Thomson, 2010; Kessler & Wang, 2012).

In conclusion, our findings not only reveal an immediate effect of a dynamic body representation (i.e. body schema) on the integrated perception of the bodies of conspecifics, but also highlight the relevance of this mechanism for establishing higher-level co-representations during coordinated or competitive social interactions.

Conclusions

Across three experiments and a rating study (see Appendix) we have provided evidence that human observers implicitly and rapidly integrate isolated body features into body-gestalts and that they do so by employing an online representation of their own body (eBG completion). We identified the body representation recruited during eBG completion as the body schema, while also allowing that frequently encountered postures might benefit from learned structural and conceptual body knowledge. The involvement of the body schema is clear evidence for a body-specific gestalt completion mechanism that is not shared with other objects.

Interestingly, switching attention away from a misleading gaze cue is also facilitated by eBG completion, suggesting that it may play a role in counteracting deception. This and the observation that overall posture dominated the particulars of the imitated effector, such as finger movement type and finger movement direction, allows us to propose flexible hierarchies for embodied self-other mappings. The multiple levels within the hierarchy could be focused upon dynamically as required, zooming in on effectors or zooming out on the overall gestalt depending on the context. Such automatic yet flexible hierarchical self-other mappings might manifest themselves in everyday social interactions: Humans might subconsciously mimic others in various aspects of their behaviour because it facilitates high-level aligned representations.

Ethical Statement

Informed consent was obtained from all participants in writing. All experimental procedures were approved by the local ethics committee and were in concordance with the Declaration of Helsinki and with the guidelines of the British Psychological Society and the American Psychological Association.

Acknowledgements

This research was supported by ESRC/MRC funding (RES-060-25-0010) awarded to KK. We would like to thank: Guillaume A. Rousselet for allowing us to use a picture of his face as stimulus; John Shaw for manufacturing the response pads; Rebecca Wilson, Ben Harkin, and Aleksandra Pastuszak for help with data collection; Stephen Butterfill for his astute comments and help with streamlining the manuscript; Chad Marsolek and four anonymous Reviewers for their very helpful and constructive criticism of earlier versions of the manuscript.

References

Amano, S., Kezuka, E., & Yamamoto, A. (2004). Infant shifting attention from an adult's face to an adult's hand: a precursor of joint attention. *Infant Behaviour and Development, 27*, 64-80.

Amorim, M. A., Isableu, B., & Jarraya, M. (2006). Embodied spatial transformations: "body analogy" for the mental rotation of objects. *J Exp Psychol Gen, 135*(3), 327-347.

Bayliss, A. P., Bartlett, J., Naughtin, C. K., & Kritikos, A. (2011). A Direct Link Between Gaze Perception and Social Attention. *Journal of Experimental Psychology-Human Perception and Performance, 37*(3), 634-644.

Bertenthal, B. I., & Pinto, J. (1994). Global Processing of Biological Motions. *Psychological Science, 5*(4), 221-225.

Biermann-Ruben, K., Kessler, K., Jonas, M., Siebner, H. R., Baumer, T., Munchau, A., et al. (2008). Right hemisphere contributions to imitation tasks. *Eur J Neurosci, 27*(7), 1843-1855.

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

Brass, M., Bekkering, H., Wohlschlager, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: comparing symbolic, spatial, and imitative cues. *Brain Cogn, 44*(2), 124-143.

Buxbaum, L. J., & Coslett, H. B. (2001). Specialised structural descriptions for human body parts: Evidence from autotopagnosia. *Cognitive Neuropsychology, 18*(4), 289-306.

Buxbaum, L. J., Giovannetti, T., & Libon, D. (2000). The role of the dynamic body schema in praxis: Evidence from primary progressive apraxia. *Brain and Cognition, 44*(2), 166-191.

Buxbaum, L. J., Kyle, K. M., & Menon, R. (2005). On beyond mirror neurons: internal representations subserving imitation and recognition of skilled object-related actions in humans. *Brain Res Cogn Brain Res, 25*(1), 226-239.

Catmur, C., Walsh, V., & Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Current Biology, 17*(17), 1527-1531.

Chartrand, T. L., & Bargh, J. A. (1999). The Chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology, 76*(6), 893-910.

Corradi-Dell'Acqua, C., Tomasino, B., & Fink, G. R. (2009). What Is the Position of an Arm Relative to the Body? Neural Correlates of Body Schema and Body Structural Description. *Journal of Neuroscience, 29*(13), 4162-4171.

Coslett, H. B., Buxbaum, L. J., & Schwoebel, J. (2008). Accurate reaching after active but not passive movements of the hand: Evidence for forward modeling. *Behavioural Neurology, 19*(3), 117-125.

Downing, P. E., Jiang, Y. H., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science, 293*(5539), 2470-2473.

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature, 424*(6950), 769-771.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science, 308*(5722), 662-667.

Fogassi, L., & Luppino, G. (2005). Motor functions of the parietal lobe. *Current Opinion in Neurobiology, 15*(6), 626-631.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychol Bull, 133*(4), 694-724.

Frith, C. D., & Frith, U. (2007). Social cognition in humans. *Curr Biol, 17*(16), R724-732.

Gallese, V. (2007). Before and below 'theory of mind': embodied simulation and the neural correlates of social cognition. *Philos Trans R Soc Lond B Biol Sci, 362*(1480), 659-669.

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

Garrod, S., & Pickering, M. J. (2009). Joint Action, Interactive Alignment, and Dialog. *Topics in Cognitive Science, 1*(2), 292-304.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3*(2), 191-197.

Goldenberg, G., & Hagmann, S. (1997). The meaning of meaningless gestures: A study of visuo-imitative apraxia. *Neuropsychologia, 35*(3), 333-341.

Goldenberg, G., & Karnath, H. O. (2006). The neural basis of imitation is body part specific. *Journal of Neuroscience, 26*(23), 6282-6287.

Goldenberg, G., Laimgruber, K., & Hermsdorfer, J. (2001). Imitation of gestures by disconnected hemispheres. *Neuropsychologia, 39*(13), 1432-1443.

Gowen, E., & Miall, R. C. (2006). Eye-hand interactions in tracing and drawing tasks. *Hum Mov Sci, 25*(4-5), 568-585.

Heyes, C. (2010). Where do mirror neurons come from? *Neuroscience and Biobehavioral Reviews, 34*(4), 575-583.

Hommel, B. (1993). Inverting the Simon Effect by Intention - Determinants of Direction and Extent of Effects of Irrelevant Spatial Information. *Psychological Research-Psychologische Forschung, 55*(4), 270-279.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science, 286*(5449), 2526-2528.

Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends Cogn Sci, 9*(1), 21-25.

Johansson, G. (1973). Visual-Perception of Biological Motion and a Model for Its Analysis. *Perception & Psychophysics, 14*(2), 201-211.

Jonas, M., Biermann-Ruben, K., Kessler, K., Lange, R., Baumer, T., Siebner, H. R., et al. (2007). Observation of a finger or an object movement primes imitative responses differentially. *Exp Brain Res, 177*(2), 255-265.

Kanizsa, G. (1974). Contours without gradients or cognitive contours. *Italian Journal of Psychology, 1*, 93-113.

Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience, 3*(8), 759-763.

Kessler, K., Biermann-Ruben, K., Jonas, M., Siebner, H. R., Baumer, T., Munchau, A., et al. (2006). Investigating the human mirror neuron system by means of cortical synchronization during the imitation of biological movements. *Neuroimage, 33*(1), 227-238.

Kessler, K., Gordon, L., Cessford, K., & Lages, M. (2010). Characteristics of Motor Resonance Predict the Pattern of Flash-Lag Effects for Biological Motion. *Plos One, 5*(1), e8258. doi:8210.1371/journal.pone.0008258.

Kessler, K., & Rutherford, H. (2010). The two forms of Visuo-Spatial Perspective Taking are differently embodied and subserve different spatial prepositions. *Frontiers in Psychology, 1*, doi: 10.3389/fpsyg.2010.00213.

Kessler, K., & Thomson, L. A. (2010). The embodied nature of spatial perspective taking: embodied transformation versus sensorimotor interference. *Cognition, 114*(1), 72-88.

Kessler, K., & Wang, H. F. (2012). Spatial Perspective Taking is an Embodied Process, but Not for Everyone in the Same Way: Differences Predicted by Sex and Social Skills Score. *Spatial Cognition and Computation, 12*(2-3), 133-158.

Kilner, J. M., Vargas, C., Duval, S., Blakemore, S. J., & Sirigu, A. (2004). Motor activation prior to observation of a predicted movement. *Nat Neurosci, 7*(12), 1299-1301.

Kourtis, D., Sebanz, N., & Knoblich, G. (2010). Favouritism in the motor system: social interaction modulates action simulation. *Biology Letters, 6*(6), 758-761.

Lafrance, M. (1985). Postural Mirroring and Intergroup Relations. *Personality and Social Psychology Bulletin, 11*(2), 207-217.

Lestou, V., Pollick, F. E., & Kourtzi, Z. (2008). Neural substrates for action understanding at different description levels in the human brain. *Journal of Cognitive Neuroscience, 20*(2), 324-341.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J Physiol Paris, 102*(1-3), 59-70.

Maravita, A., & Iriki, A. (2004). Tools for the body (schema). *Trends in Cognitive Sciences, 8*(2), 79-86.

McKay, L. S., Simmons, D. R., McAleer, P., Marjoram, D., Piggot, J., & Pollick, F. E. (2011). Do distinct atypical cortical networks process biological motion information in adults with Autism Spectrum Disorders? *Neuroimage*.

Medina, J., Jax, S. A., & Coslett, H. B. (2009). Two-component models of reaching: Evidence from deafferentation in a Fitts' law task. *Neuroscience Letters, 451*(3), 222-226.

Meltzoff, A. N., & Moore, M. K. (1994). Imitation, Memory, and the Representation of Persons. *Infant Behavior & Development, 17*(1), 83-99.

Meltzoff, A. N., & Moore, M. K. (1997). Explaining facial imitation: A theoretical model. *Early Development & Parenting, 6*(3-4), 179-192.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*(2), 169-+.

Ramnani, N., & Miall, R. C. (2004). A system in the human brain for predicting the actions of others. *Nat Neurosci, 7*(1), 85-90.

Reed, C. L., & Farah, M. J. (1995). The Psychological Reality of the Body Schema - a Test with Normal Participants. *Journal of Experimental Psychology-Human Perception and Performance, 21*(2), 334-343.

Reed, C. L., & McGoldrick, J. E. (2007). Action during body perception: Processing time affects self-other correspondences. *Social Neuroscience, 2*(2), 134-149.

Riggio, L., Gawryszewski, L. D., & Umilta, C. (1986). What Is Crossed in Crossed-Hand Effects. *Acta Psychologica, 62*(1), 89-100.

Rizzolatti, G., & Fabbri-Destro, M. (2010). Mirror neuron mechanism. *Encyclopedia of Behavioural Neuroscience, 200*(3-4), 240-249.

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1999). Resonance behaviors and mirror neurons. *Archives Italiennes De Biologie, 137*(2-3), 85-100.

Saygin, A. P. (2007). Superior temporal and premotor brain areas necessary for biological motion perception. *Brain, 130*(Pt 9), 2452-2461.

Schwoebel, J., & Coslett, H. B. (2005). Evidence for multiple, distinct representations of the human body. *Journal of Cognitive Neuroscience, 17*(4), 543-553.

Shockley, K., Santana, M. V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology-Human Perception and Performance, 29*(2), 326-332.

Simon, J. R., Hinrichs, J. V., & Craft, J. L. (1970). Auditory S-R Compatibility - Reaction Time as a Function of Ear-Hand Correspondence and Ear-Response-Location Correspondence. *Journal of Experimental Psychology, 86*(1), 97-&.

Sirigu, A., Grafman, J., Bressler, K., & Sunderland, T. (1991). Multiple Representations Contribute to Body Knowledge Processing - Evidence from a Case of Autotopagnosia. *Brain, 114*, 629-642.

Tarr, M. J., & Gauthier, I. (2000). FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience, 3*(8), 764-769.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychol Bull, 131*(3), 460-473.

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An Internal Model for Sensorimotor Integration. *Science, 269*(5232), 1880-1882.

Table 1: Cell means and standard deviations in Experiment 1.

Stimulus posture:	hands down	hands down	hands down	hands down	hands raised	hands raised	hands raised	hands raised
SOA:	17	17	500	500	17	17	500	500
Validity:	invalid	valid	invalid	valid	invalid	valid	invalid	valid
Mean RT:	741.23	687.75	639.35	605.04	778.75	748.19	655.17	618.92
StdDev RT:	142.28	112.77	119.86	131.45	151.75	145.01	111.06	104.86
Mean ACC:	97.92	98.44	98.44	98.44	94.79	94.79	99.48	97.92
StdDev ACC:	4.76	4.22	4.22	4.22	8.97	8.97	2.55	4.76

Table 2: Cell means and standard deviations in Experiment 2.

Participant posture:	hands down	hands down	hands down	hands down	hands raised	hands raised	hands raised	hands raised
Stimulus posture:	hands down	hands down	hands raised	hands raised	hands down	hands down	hands raised	hands raised
Validity:	invalid	valid	invalid	valid	invalid	valid	invalid	valid
Mean RT:	904.66	846.10	973.43	921.57	1007.17	991.52	903.78	902.88
StdDev RT:	136.47	170.61	160.32	132.32	176.12	191.47	156.43	172.11
Mean ACC:	97.22	99.31	91.67	94.44	90.97	94.10	92.71	95.49
StdDev ACC:	4.71	2.35	8.51	8.39	9.17	7.57	10.22	8.14

Table 3: Design implications in Experiment 3.

Posture of: Participant	Posture of: Stimulus	Posture congruence	Stimulus movement direction	Direction congruence	Stimulus movement type	Type congruence
palms up	palms up	congruent	up	congruent	tap	congruent
palms up	palms up	congruent	down	incongruent	lift	incongruent
palms up	crossed	incongruent	up	congruent	lift	incongruent
palms up	crossed	incongruent	down	incongruent	tap	congruent
crossed	crossed	congruent	down	congruent	tap	congruent
crossed	crossed	congruent	up	incongruent	lift	incongruent
crossed	palms up	incongruent	down	congruent	lift	incongruent
crossed	palms up	incongruent	up	incongruent	tap	congruent

Note. Participants always responded with a key press, i.e. with a tapping movement. This tap could be either directed downwards or upwards depending on the posture of the participant on a given trial. This implies that Experiment 3 did not include all possible factor combinations.

Table 4: Cell means and standard deviations in Experiment 3.

Participant posture:	hands crossed	hands crossed	hands crossed	hands crossed	palms up	palms up	palms up	palms up
Stimulus posture:	hands crossed	hands crossed	palms up	palms up	hands crossed	hands crossed	palms up	palms up
Validity:	invalid	valid	invalid	valid	invalid	valid	invalid	valid
Mean RT:	1123.10	1083.69	1162.42	1131.79	1003.02	933.88	939.40	899.88
StdDev RT:	222.53	218.96	246.00	256.49	186.37	151.80	218.90	175.07
Mean ACC:	82.29	86.72	81.77	82.03	91.67	96.61	95.31	93.75
StdDev ACC:	12.18	13.20	13.28	14.66	10.04	6.11	5.29	7.14

Figure Captions

Figure 1: Stimuli and participants' hand posture in Experiment 1. The left stimulus configuration implies a posture as if the person was carrying a log, while the right stimulus implies a posture as if the person was preparing for an underhand chin-up. In this experiment participants responded with their palms directed upwards (using an upside-down keyboard), thus, optimally matching the implied posture of the left stimulus ('as if carrying a log'). Both stimuli show examples of valid gaze cues. The stimuli also show a deflection of the little finger in the left and a deflection of the index finger in the right configuration. Further explanations in the text.

Figure 2: Response Time (RT) and Accuracy (ACC) data in Experiment 1. Panel A shows RT data (in ms) and Panel B shows ACC data (in % correct). The numbers 17 and 500 refer to SOA in ms. Error bars are the standard error of mean and * denote significant differences (*: $p < .05$; **: $p < .01$). Further explanations in the text.

Figure 3: Stimuli and participants' postures in Experiment 2. Stimuli were identical to Experiment 1, but the participant's posture was changed randomly across trials between the two alternatives shown in the middle section of the Figure. The left posture of the participant with palms upwards and hands resting on the table was identical to the participants' posture throughout Experiment 1. Here a second posture was introduced where participants were instructed to raise their arms onto an armrest while keeping their hands in a palms-up position. The right section of the Figure shows one of the custom-made response pads. Two keys were used on each of two Targa® wireless number-pads for recording responses of the participant's left and

right hand, respectively. Wireless connections to the PC allowed for maximum flexibility of movement without the danger of cables becoming entangled. As shown in the Figure, Velcro® straps were used to attach the number-pad and the participants' hand to the same plywood support platform (rigid, but light and easy to move), enabling free movements of the hands and arms, while keeping the two response fingers of each hand (index and little finger) on top of the correct response keys at all times. This was essential for quick changes of posture (randomly across trials) while preserving the speed of responses. Further explanations in the text.

Figure 4: Response Time (RT) and Accuracy (ACC) data in Experiment 2. Panel A shows RT data (in ms) and Panel B shows ACC data (in % correct). Error bars denote the standard error of mean and * denote significant differences (*: $p < .05$; **: $p < .01$; ***: $p < .001$). Further explanations in the text.

Figure 5: Stimuli and participants' hand postures in Experiment 3. The left stimulus shows the same implied posture 'as if carrying a log' as previously employed in Experiments 2 and 3. The right stimulus implies a posture with crossed hands. Participants were instructed to adopt corresponding postures in a randomised fashion before each trial, thus, either creating a posture match with the first stimulus ('carrying a log'), or a posture match with the second stimulus ('hands crossed'). Additionally, in this experiment the direction of finger movements onscreen had to be controlled. On the far right the Figure shows the direction of the participants' key presses for each of their two postures: Always 'downwards' when their hands were crossed, and always 'upwards' when their hands were resting palms up on the table (the same response pads were used as in Expt. 2; see Fig. 3). Note that onscreen

fingers could move in either direction (downwards or upwards) within both stimulus configurations (executing lifts as well as taps). Thus, disregarding posture, the direction of the participant's key presses was congruent to the direction of the finger movement onscreen on half of the trials, while it was incongruent on the other half. Examples are visualised by means of grey arrows in the "Stimuli" section of the Figure. Note that only one finger moved on any given trial in the experiment. Also note that participants always tapped their finger, while onscreen fingers executed lifts as well as taps. Consequently, Experiment 3 did not include all possible factor combinations in the design. Please refer to Table 3 for full details of the employed design. Further explanations in the text.

Figure 6: Response Time (RT) and Accuracy (ACC) data in Experiment 3. Panel A shows RT data (in ms) and Panel B shows ACC data (in % correct). Error bars denote the standard error of mean and * denote significant differences (*: $p < .05$; **: $p < .01$; ***: $p < .001$). Further explanations in the text.

Figure 1

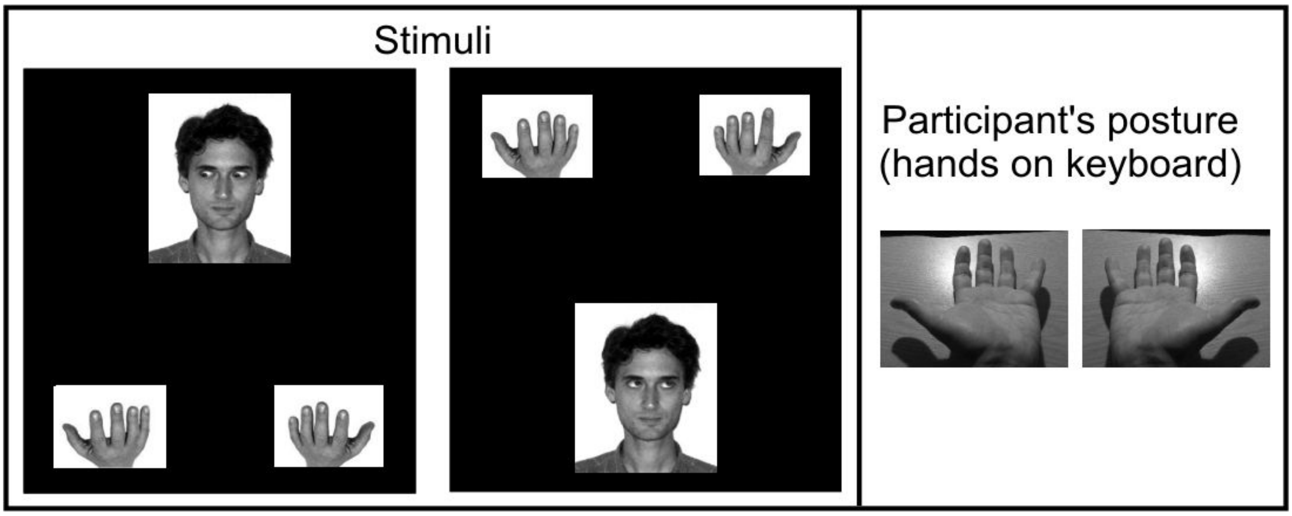


Figure 2:

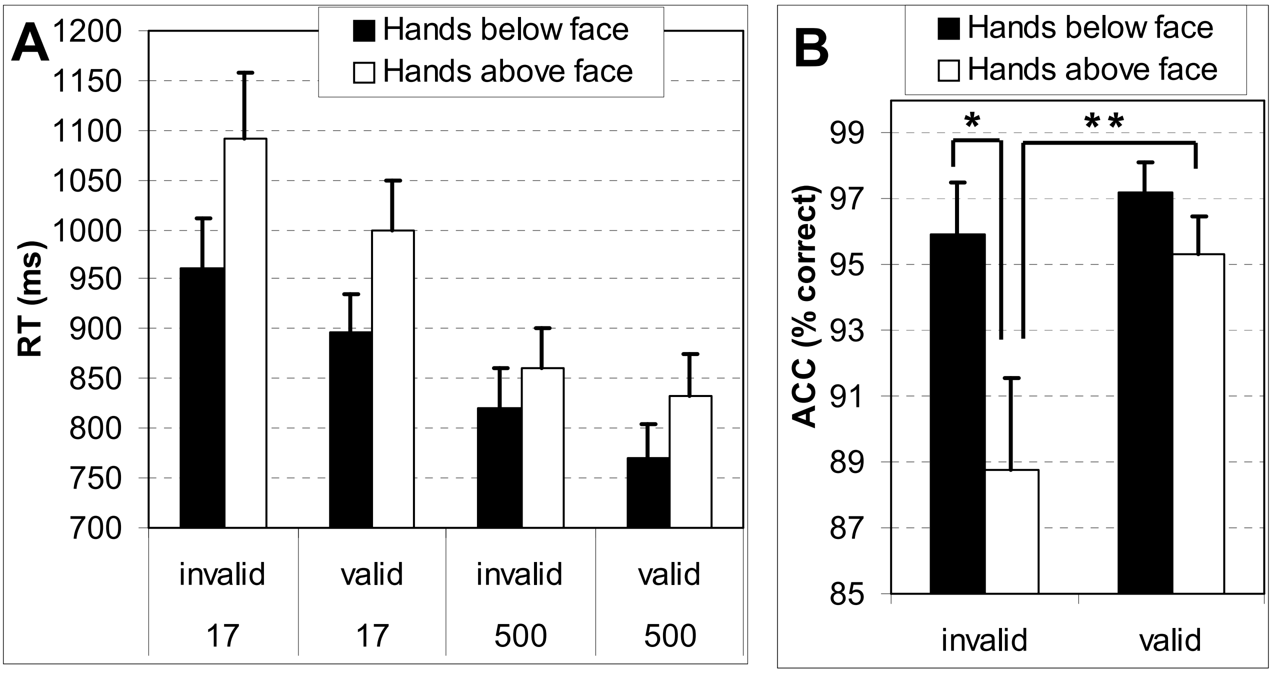

A
RT (ms)
1200
1150
1100
1050
1000
950
900
850
800
750
700
Hands below face
Hands above face
invalid
17
valid
17
invalid
500
valid
500
B
ACC (% correct)
Hands below face
Hands above face
99
97
95
93
91
89
87
85
*
**
invalid
valid

Figure 3:

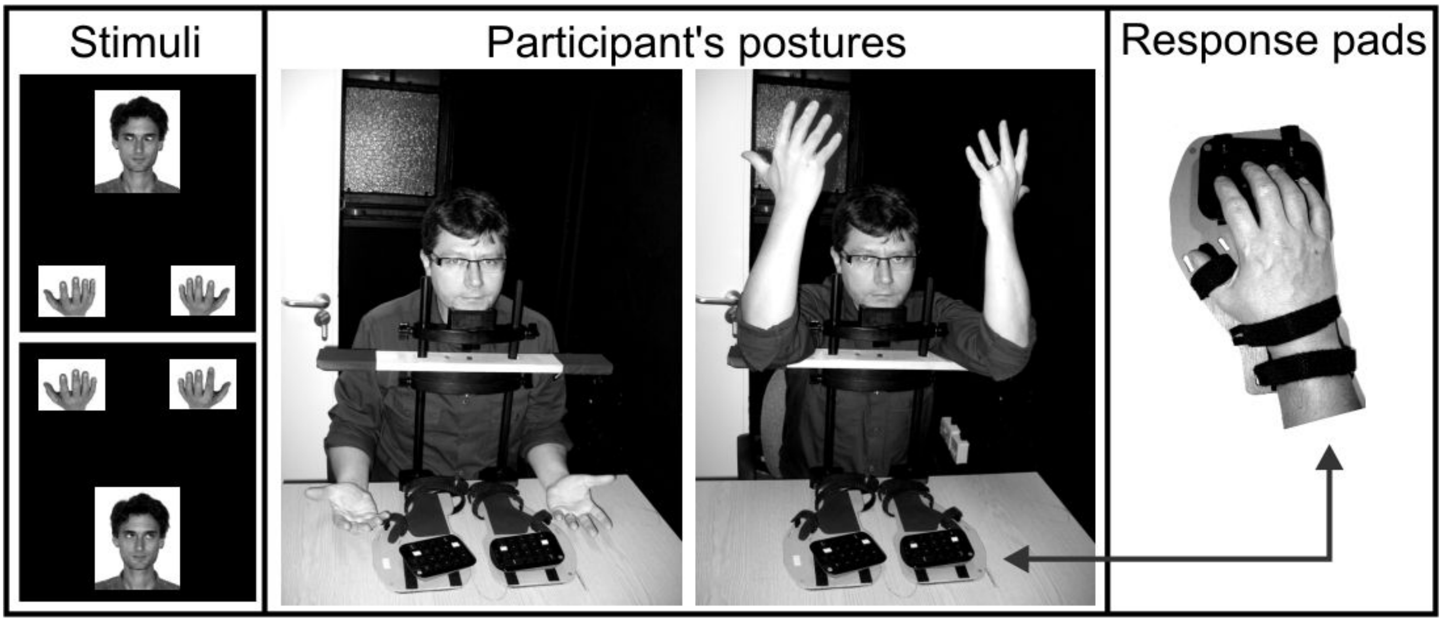


Figure 4:

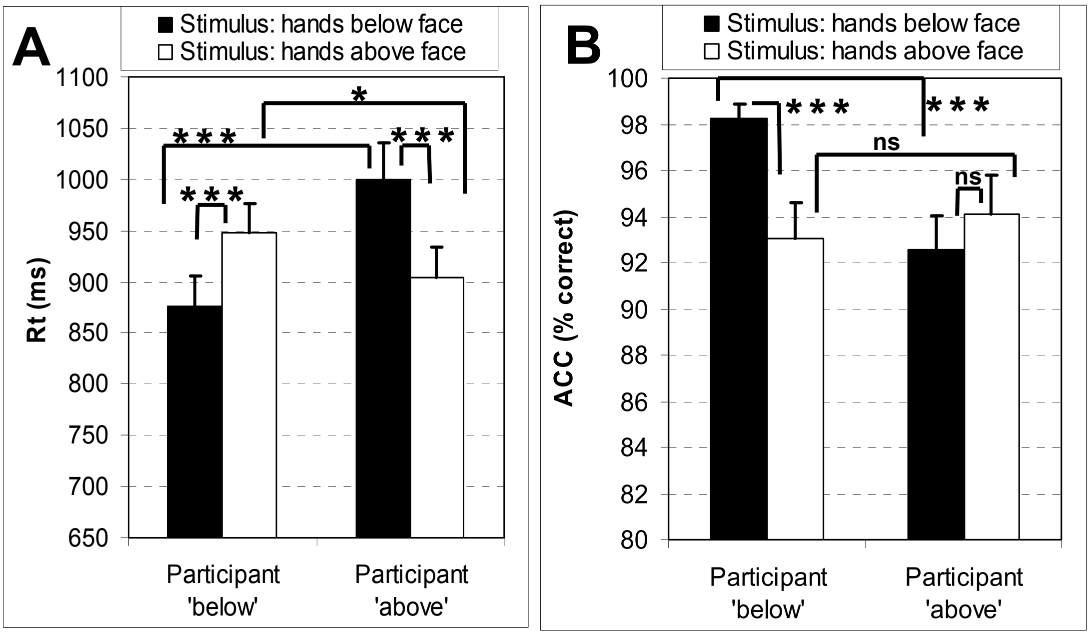

A
■ Stimulus: hands below face
□ Stimulus: hands above face
Rt (ms)
1100
1050
1000
950
900
850
800
750
700
650
***
***
*
***
Participant 'below'
Participant 'above'
B
■ Stimulus: hands below face
□ Stimulus: hands above face
ACC (% correct)
100
98
96
94
92
90
88
86
84
82
80
***
***
ns
ns
Participant 'below'
Participant 'above'

Figure 5:

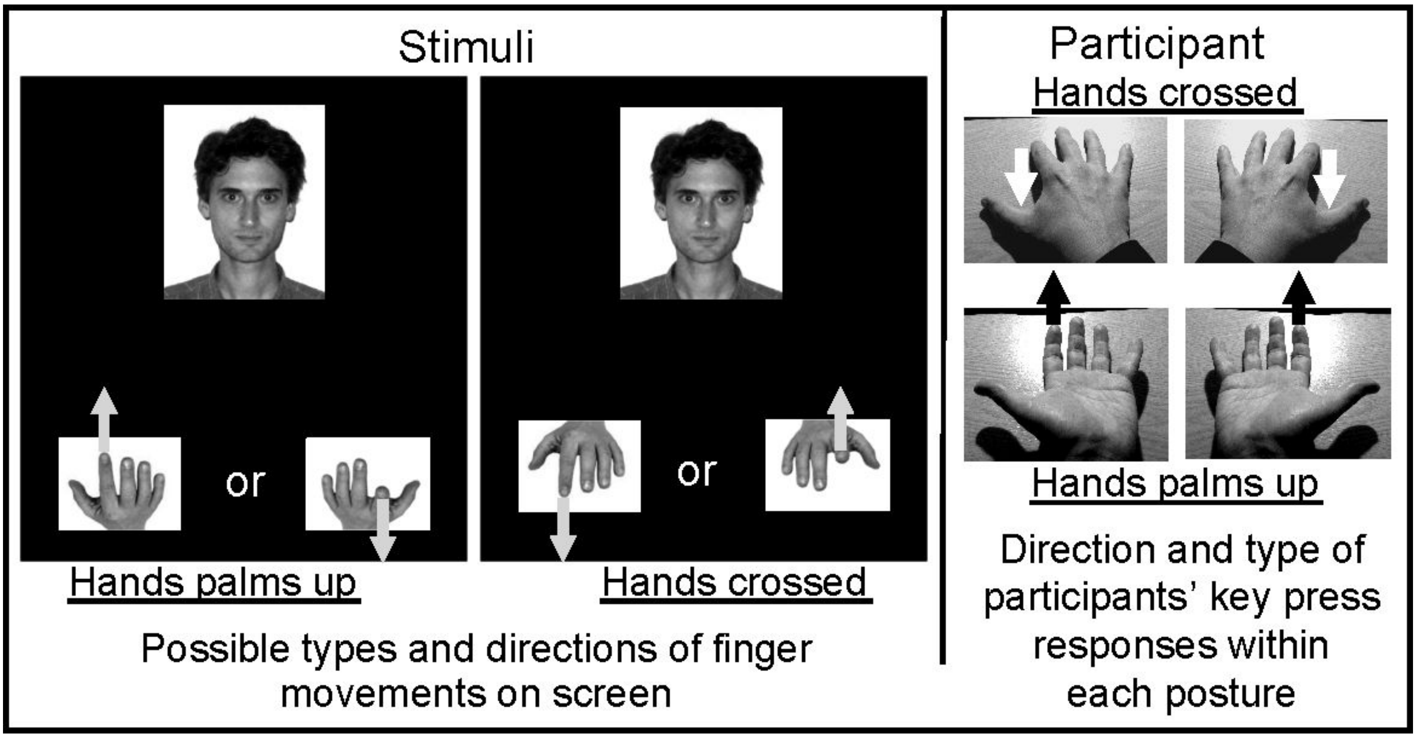
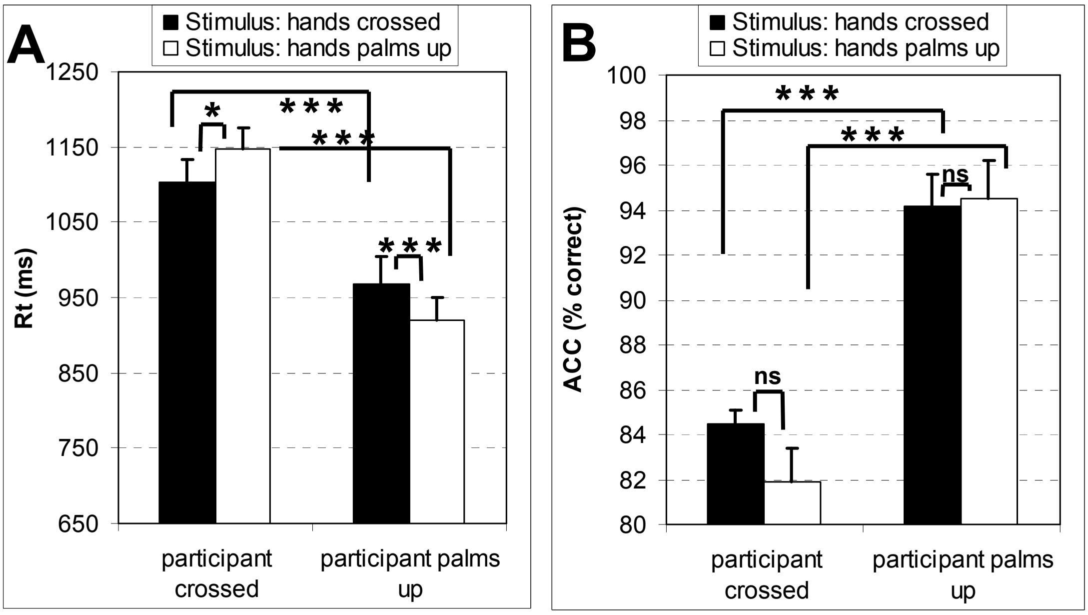

Stimuli
or
or
Hands palms up
Hands crossed
Possible types and directions of finger movements on screen
Participant
Hands crossed
Hands palms up
Direction and type of participants' key press responses within each posture

Figure 6:


A
Stimulus: hands crossed
Stimulus: hands palms up
Rt (ms)
1250
1150
1050
950
850
750
650
*
***
***
***
participant crossed
participant palms up
B
Stimulus: hands crossed
Stimulus: hands palms up
ACC (% correct)
100
98
96
94
92
90
88
86
84
82
80
***
***
ns
ns
participant crossed
participant palms up

Appendix: Implied Posture Ratings

We conducted a rating study, where we asked 14 right-handed Glasgow University students (7 females; average age 22) to indicate the plausibility of the postures implied by the stimuli used in the three reported Experiments and an unreported Pilot on a scale from 0 to 100. We also included anatomically impossible control stimuli that were not employed in any Experiment.

Methods

The rated stimulus configurations are shown in Figure A and comprised all stimuli employed throughout Experiments 1-3 (Postures 3-5 in top row/black bars of Fig. A) and an unreported Pilot (Postures 1 and 2) as well as anatomically impossible counterparts (bottom row/grey bars in graph of Fig. A). Anatomically impossible implied postures showed two left or two right hands and were not used in our experiments since the finger-mapping between stimulus and participant would have mismatched for one of the hands. The stimuli in the Rating study were presented in random succession (using E-prime® 2.0) and participants could look at the static stimulus (no gaze, no finger movements) for as long as they wished. Upon pressing a mouse button the stimulus disappeared and a horizontal slider was presented. Participants dragged the slider by means of the mouse towards the lower (0) or the upper end of the scale (100), indicating whether the displayed posture had been plausible (100 max) or implausible (0 min). Each stimulus configuration was presented once and it was stressed that response speed was not of the essence. Participants were neither encouraged nor discouraged to change their own posture during the stimulus presentation. Eleven out of the fourteen participants reported afterwards that they had moved their arms/hands at least once to try to align themselves with the implied posture of the displayed stimulus configuration in order to optimally appraise plausibility.

Results

The ratings are presented in Figure A and Wilcoxon non-parametric signed-ranks tests for related samples were employed for comparisons to accommodate for the different usage of the scale across participants. Generally the pattern consistently revealed that participants were well able to tell the difference between possible (in principle) and anatomically impossible postures (black vs. grey bars). In addition Figure A also shows which (possible) stimulus configurations were employed in which experiment and how their ratings differed or not. Further explanations are provided in the text of the respective experiments.

Table A: Posture Ratings, descriptive statistics.

Posture:	1		2		3		4		5	
Condition:	possible	impossible	possible	impossible	possible	impossible	possible	impossible	possible	impossible
Mean:	98.71	9.07	81.36	11.21	88.93	5.75	80.71	7.73	62.96	7.82
Stddev:	1.54	11.53	27.12	13.74	25.41	9.95	34.50	12.31	40.33	12.27

Posture numbers refer to the numbers in Figure A.

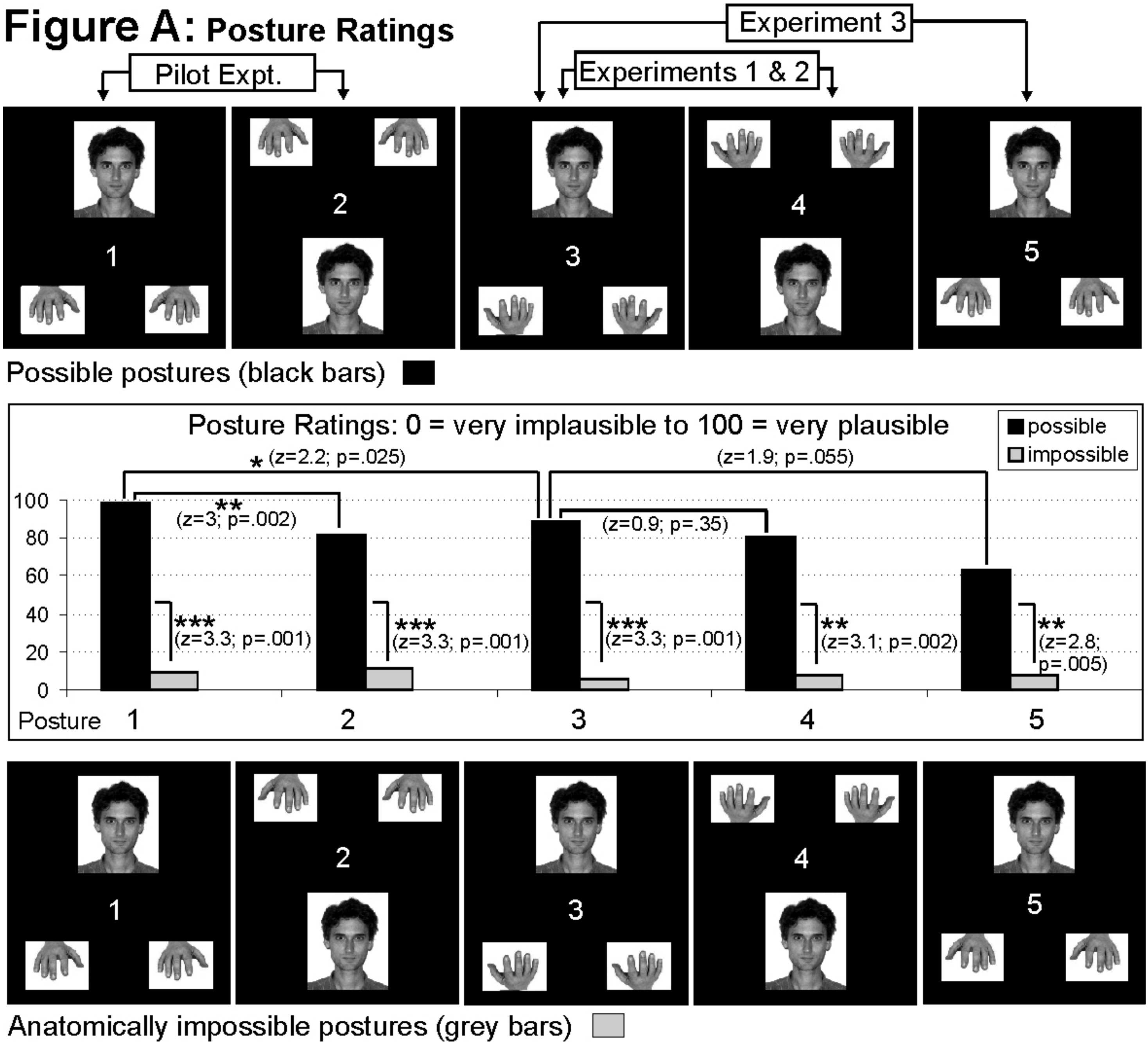


Figure A: Plausibility Ratings for implied postures. The rated material comprised the five stimulus configurations employed throughout Experiments 1-3 (configurations 3-5 in the top row) plus an unreported Pilot (configurations 1 and 2 in the top row), which implied postures that were 'in principle' possible (top row/black bars in graph). In addition anatomically impossible counterparts showing either two left or two right hands together with the face were also rated (bottom row/grey bars in graph). The impossible postures were not employed in the experiments reported here. The pictures of the implied postures are numbered and these numbers correspond to the x-axis in the graph, thus, linking stimuli to their ratings. At the top of the bar graph the main statistical comparisons are visualised as lines and significance values are provided. Comparisons were conducted by means of Wilcoxon signed-ranks tests for related samples.