Journal of Experimental Psychology:
Human Perception and Performance
2014, Vol. 40, No. 2, 457–464

0096-1523/14/$12.00 DOI: 10.1037/a0035348

OBSERVATION

Can't Touch This: The First-Person Perspective Provides Privileged Access to Predictions of Sensory Action Outcomes

Patric Bach
Plymouth University

Wendy Fenton-Adams
Bangor University

Steven P. Tipper
University of York

Previous studies have shown that viewing others in pain activates cortical somatosensory processing areas and facilitates the detection of tactile targets. It has been suggested that such shared representations have evolved to enable us to better understand the actions and intentions of others. If this is the case, the effects of observing others in pain should be obtained from a range of viewing perspectives. Therefore, the current study examined the behavioral effects of observed grasps of painful and nonpainful objects from both a first- and third-person perspective. In the first-person perspective, a participant was faster to detect a tactile target delivered to their own hand when viewing painful grasping actions, compared with all nonpainful actions. However, this effect was not revealed in the third-person perspective. The combination of action and object information to predict the painful consequences of another person's actions when viewed from the first-person perspective, but not the third-person perspective, argues against a mechanism ostensibly evolved to understand the actions of others.

Keywords: viewing perspective, first person, third person, action observation, action prediction, mirror neurons

Supplemental materials: http://dx.doi.org/10.1037/a0035348.supp

Viewing others perform actions (e.g., Oosterhof, Wiggett, Diedrichsen, Tipper, & Downing, 2010; Rizzolatti & Craighero, 2004), display emotions (Wicker et al., 2003), encounter touch (Keysers et al., 2004; Bufalari, Aprile, Avenanti, Di Russo, & Aglioti, 2007), and pain (Morrison, Peelen, & Downing, 2007; Singer et al., 2004) elicits activation of neuronal ensembles that are similarly recruited when we directly experience these phenomena. A prominent view is that these "shared representations" are a product of specialized brain mechanisms that give people direct insights into the internal states of others (e.g., Ramachandran, 2000; Oberman & Ramachandran, 2008; Schütz-Bosbach, Mancini, Aglioti, & Haggard, 2006). On this view, shared representations have evolved as adaptations to the requirement of having to understand the behavior of conspecifics and are assumed to confer substantial adaptive advantages: they help people to empathize with one another, to coordinate and predict future actions, and to detect deception. Consistent with such a primarily social role, a large number of mirror neurons are viewpoint independent or respond selectively to actions from a third-person perspective (Caggiano et al., 2011; Oosterhof, Tipper, & Downing, 2012), and there is increasing evidence that disrupting shared representation also disrupts social understanding (see Avenanti, Candidi, & Urgesi, 2013, for a review).

Recently, however, the view that shared representations evolved specifically to facilitate social understanding has been challenged (cf. Heyes, 2010; Brass & Heyes, 2005; Keysers & Perrett, 2004). These theories do not deny that shared representations are computed in a wide range of circumstances or that they play a crucial role in action understanding and empathy. However, instead of emerging from specifically evolved brain systems for social understanding, shared representations are seen as (very useful) by-products of the processes that monitor and control the individual's own actions (e.g., Brass & Heyes, 2005; Heyes, 2010; Keysers & Perrett, 2004; Gallese, 2001). Prediction processes—and the internal models they rely on—have taken center stage in such accounts. Humans constantly predict how their bodies will affect the environment, and how the environment will affect them (e.g., Friston, 2010; Schütz-Bosbach & Prinz, 2007). These predictions allow them to take evasive action, to make course corrections, or to stop actions altogether if negative outcomes are expected. They emerge

Patric Bach, School of Psychology, Plymouth University, Devon, England; Wendy Fenton-Adams, School of Psychology, Bangor University, Gwynedd, Wales; and Steven P. Tipper, Psychology Department, University of York, York, England.

Supported by the Economic and Social Research Council (Grant ES/J019178/1).

Correspondence concerning this article should be addressed to Steven P. Tipper, Psychology Department, University of York, Heslington, York, YO10 5DD, England. E-mail: steven.tipper@york.ac.uk

from sophisticated processes combining multiple sources of information. During reaching, for example, actors combine information about their action with the internal model of the goal object—its anticipated weight, softness, and texture—to predict the specific consequences of grasping it, such that grips can be adjusted and future actions can be planned before contact is made (see Johansson & Flanagan, 2009, for a review).

We and others have argued that shared representations could emerge naturally from such prediction mechanisms (e.g., Morrison, Tipper, Fenton-Adams, & Bach, 2013; Bach, Bayliss, & Tipper, 2011; Kilner, Friston, & Frith, 2007; Miall, 2003; Gallese, 2001). Because the visual input to these mechanisms is very similar during both action and observation, the same integration processes can take place and yield the same predictions of action outcomes. A recent study (Morrison et al., 2013) provided initial evidence for this idea. Participants watched hands either grasp or withdraw from painful and nonpainful objects, and judged the appropriateness of these actions. We tested whether observers' tactile processing systems would represent the painful sensory consequences associated with grasping the painful object. No direct cues to pain were shown, such as skin damage or negative emotional expressions. Any sensory expectation of pain could therefore not be directly extracted from the stimulus, but—similar to action execution—had to be predicted by combining action information (whether it involved hand-object contact) with the internal model of the goal object (whether it was painful to touch). We found that observers' somatosensory cortices showed higher activations for painful grasps than in any other hand−object interaction that did not cause pain, suggesting that participants made such predictions. Moreover, subsequent psychophysical experiments revealed that seeing painful grasps also increased participants' readiness to detect tactile stimulation on their own fingers, but not auditory stimulation. The lack of effect with auditory stimuli ruled out more general attention or arousal related explanations of the effects, and revealed that the predictions of sensory action outcomes specifically affected sensory-tactile representation systems.

These findings show that people make sophisticated predictions about the outcome of others' actions, but it remains an open issue whether these predictions emerge from processes evolved to enable social understanding or to enable the prediction of consequences for the self. The present study tested these alternative hypotheses, following a logic introduced by Oberman and Ramachandran (2008; see Gallese, 2001; Schütz-Bosbach et al., 2006, for related approaches). It rests on the notion that a mechanism, which has evolved for monitoring one's own actions, should be driven most directly by visual input that matches the first-person view one has of one's own actions. The sensory consequences of others actions should therefore be derived effectively when seen from this first-person perspective, but less so when they are seen from a third-person view, which captures the typical viewpoint when watching the actions of others. The opposite pattern is predicted if these mechanisms are specialized for social understanding (cf. Oberman & Ramachandran, 2008). That is, simulation of the sensory consequences of an action should be activated when observing the actions of another person, which are typically viewed from the third-person perspective, rather than one's own actions typically viewed from a first-person perspective. Such a specialization for the actions of others has been demonstrated for the case of automatic imitation. Motor responses during movement observation are stronger if the action is attributed to another person rather than to oneself (Schütz-Bosbach et al., 2006).

To test these predictions, we adapted the psychophysical paradigm of Morrison et al. (2013). As before, participants watched hands grasp or not grasp objects that could be painful or nonpainful, and, at the same time, pressed a button whenever they felt supraliminal tactile stimulation on their own fingers. The readiness to detect such stimulation—measured by response times—when participants viewed grasps of painful objects (relative to grasps of neutral objects or misses of either type) served as a measure of the extent to which participants inferred the sensory-tactile consequences of these actions.

Two important changes were made to the original design. First, in the original study, participants judged whether the actions were appropriate to the object (e.g., grasps were appropriate for neutral objects but not painful ones, and vice versa for withdrawals), a task that by itself encouraged deriving the sensory consequences of the actions (Morrison et al., 2013). To test natural biases in prediction systems, it is crucial to eliminate such top-down task influences and to tap into more automatic modes of processing. Thus, after participants were familiarized with the painful and nonpainful objects, they merely reported whether the hand made contact with the object; whether this contact would cause pain was not relevant.

Second, in the original study, participants saw the actions from the side (Morrison et al., 2013). To be able to manipulate the similarity of the visual input to either one's own or other people's actions, the actions were now presented from a bird's-eye perspective. This allowed us to generate first-person and third-person views by simply mirroring the displays along the horizontal axis, but showing otherwise identical stimuli. Thus, in the *first-person perspective* condition, the stimuli were rotated such that they matched the input one receives from one's own actions: during the reach, the hand moved away from the participant toward an object, while the arm pointed backward to the approximate location of the participant's body. In contrast, in the *third-person perspective* condition, the stimuli were rotated to match the view one has of the action of another person. During the reach, the hand moved toward an object and the participant, while the arm pointed away from the participant's body.

This paradigm allowed us to test, first, whether observers' tactile representation systems would predict the sensory consequences of grasping painful objects even when the nature of the object was irrelevant to the task. If this is the case, tactile stimulation should again be detected more quickly when viewing painful grasps, compared with any other type of hand-object interaction that does not cause pain (i.e., grasps of neutral objects or misses of either type of object). Second, it allowed us to test how viewing perspective affects these automatic predictions. If sensory predictions emerge from mechanisms for representing one's own actions, these effects should be stronger for actions seen from the first-person compared with the third-person perspective. In contrast, if they emerge from a dedicated system for social understanding, any effects should be stronger in the third- than the first-person perspective.

Note that prior studies have demonstrated impressive effects of third-person information, such that the observer's internal state was affected by what another person attempted to ignore (Frischen, Loach, & Tipper, 2009), attempted to avoid (Griffiths & Tipper, 2009), or what they could see (Samson, Apperly, Braithwaite, Andrews, & Bodely Scott, 2010). However, the internal states of interest were typically (a) directly discernible from the visual stimulation, (b) did not require internal inference or combination of sources of information, or (c) were observed in tasks for which representing the others' perspective was encouraged by task or stimuli. Moreover, (d) none of these studies implemented a first-person control condition. Ours is the first study to dissociate predictive and social understanding views of shared representations, where the outcome of more sophisticated prediction processes can be compared across perspectives.

Method

Participants

Forty-eight participants (14 men, three left-handed) were recruited through the participation panel at the School of Psychology at Bangor University in Wales. All participants were aged 18 years or older ($M = 20.23$ years, $SD = 3.14$), had normal or corrected-to-normal vision, and were first language English speakers. They received course credits to compensate them for their time. The procedures were approved by the School of Psychology Ethics Committee at Bangor University.

Apparatus

Visual and tactile stimuli were presented using Presentation (http://www.neurobs.com) on a 3.2-GHz Pentium computer running Windows XP. The tactile stimulator (an Oticon BC462 bone conductor) was attached to the tip of the participant's right index finger with adjustable tape. The stimulation was a 200-Hz sine wave, overlaid with white noise, lasting 50 ms. The first and last 10 ms were faded in and out to prevent sharp transients. Participants wore earplugs and ear protectors to prevent them from hearing the stimulation device.

Stimuli

Participants viewed two frame-action sequences. Each sequence was shot from a birds-eye perspective and showed a hand interacting with one of seven painful and seven nonpainful objects (Figure 1a). The first frame always showed a hand in a neutral position near an object (for 750 ms). The second frame (500 ms) showed the same hand either grasp or miss the object. The two frames followed each other without a gap, creating the impression of apparent motion (see Figure 1b).

Each sequence could be shown in the first-person or the third-person perspective (see Figure 2). The same photographs of hands and objects were used for each perspective by flipping and rotating the images. In the *first-person perspective* condition, the stimuli were rotated such that they matched the input one receives from one's own actions: during the reach, the hand moved away from the participant toward an object. In the *third-person perspective* condition, the stimuli were rotated to match the view one has of the action of another person: during the reach, the hand moved toward an object and the participant.

In the third-person perspective, as a between-subjects factor, participants either viewed an anatomical match of the first-person perspective hand (i.e., a right hand in the first-person perspective and a right hand in the third-person perspective) or viewed a mirror image of the first-person perspective hand (i.e., a right hand in the first-person perspective and a left hand in the third-person perspective) (see Figure 2). This factor of no interest was included to account for the possible differential effects of specular and mirror image forms or third-person perspective stimuli (e.g., Bertenthal, Longo, & Kosobud, 2006).

Design and Procedure

To familiarize participants with the painfulness of the objects, they first completed a computer-based 28-item rating scale questionnaire (see Morrison et al., 2013). Using a 5-point Likert scale (1 = *not at all* and 5 = *very much*), participants rated each of the 14 objects they would see during the experiment on how painful they imagined it would be to grasp the object, and to what degree they judged this from their own experience. To ensure participants would recognize the stimuli, a side view of the object (see supplemental Figure 1) was displayed for 2,000 ms. The object was then presented in the bird's-eye view for painfulness ratings, which

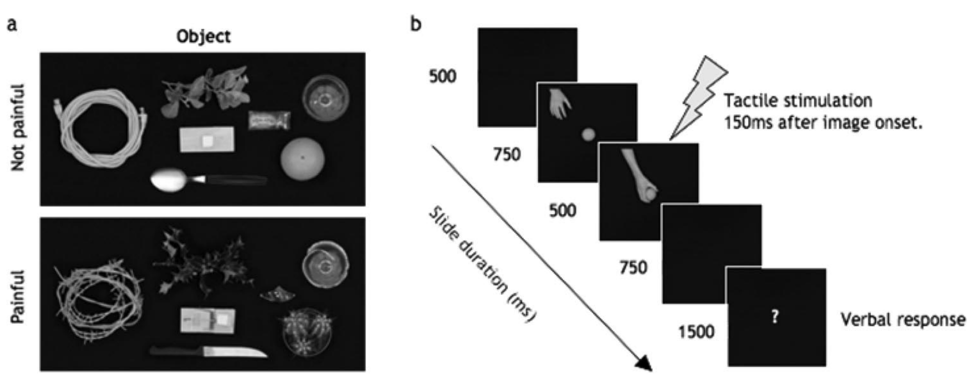

Figure 1. Examples of object stimuli. Nonpainful objects included a cable, plant, cheese on a wooden board, a spoon, a tomato sauce sachet, a wine glass, and an orange (a). Painful objects included barbed wire, holly, a loaded mousetrap, a serrated sharp knife, a shard of glass, a broken wine glass, and a cactus. Schematic of tactile detection task (total trial duration: 4,000 ms) (b). Tactile stimulation (duration: 50 ms duration) occurred 150 ms after the onset of the frame in which the hand interacted with the object. The delayed verbal response occurred during the final frame of the trial sequence, when the question mark was on screen.

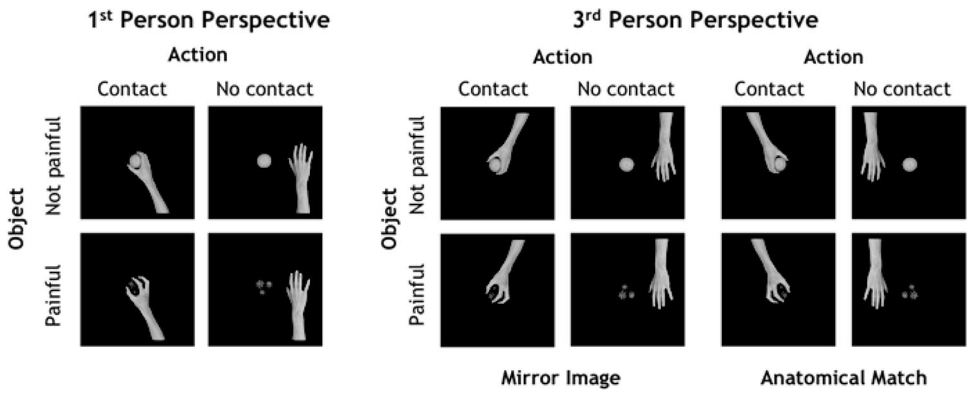

Figure 2. Example of experimental conditions. Within participants: viewpoint (first-person perspective, third-person perspective), action (contact, no contact), and object (painful, not painful). Between participants: third-person mapping (mirror image, anatomical match).

was the form in which it was seen throughout the rest of the experiment.

On completion of the rating scale, participants inserted the earplugs and wore the ear protectors. Participants attached the stimulation device to their right index finger. The stimulation device was switched on to familiarize the participant with the tactile stimulation and the tactile detection task. Participants were asked to press the spacebar on the keyboard with their left hand as quickly as possible whenever they felt the tactile stimulation. The tactile target was the same as during the experiment proper.

When the experimenter was confident the participant could perform the task correctly, participants completed 16 practice trials that were randomly selected from the main experimental trials. Participants viewed two frame sequences of a hand either approach and grasp, or miss, painful and nonpainful objects, from a first-person perspective and a third-person perspective (see Figure 1 and 2). Tactile stimulation occurred 150 ms after the onset of the second frame (where the hand performed the action) on 80% of the trials. At the end of each trial, whether there was tactile stimulation or not, participants had to make a verbal response about whether the action involved "contact" or "no contact."

The experiment proper consisted of 280 trials in total and was subdivided into four blocks. Trials were equally distributed over the eight different conditions (Object Painful/Nonpainful × Contact/No Contact × Perspective), with 35 trials in each condition, 27 of which were trials with tactile stimulation. Each of these four blocks was preceded by a shorter block of 16 trials that served to remind participants about the painfulness of the objects. In these blocks, participants saw the same actions and performed the same task, but reported whether the object they had seen was potentially "painful" or "not painful." These data were not analyzed.

Results

Three additional participants were excluded for not performing the task correctly. One participant did not report whether the hand made contact, but reported object painfulness instead. The other two did not respond quickly enough to the tactile stimulations (mean reaction times [RTs] > 1,500 ms). Preemptive detections (< 100 ms) and RTs greater than 1,500 ms were removed from the data (0.49%). The data for RTs (see Figure 3), hits, and false alarms (see Figure 4) were entered into separate 2 × 2 × 2 repeated-measures analyses of variance (ANOVA) with the factors viewpoint (first- and third-person perspective), action (contact, no contact), and object (painful, not painful) and the between-groups factor of third-person mapping (anatomical match, mirror image).

Reaction Times

The analysis of RTs (see Figure 3) revealed a main effect of action, $F(1, 46) = 58.37$, $p < .001$, $\eta_p^2 = .56$, and object, $F(1,$

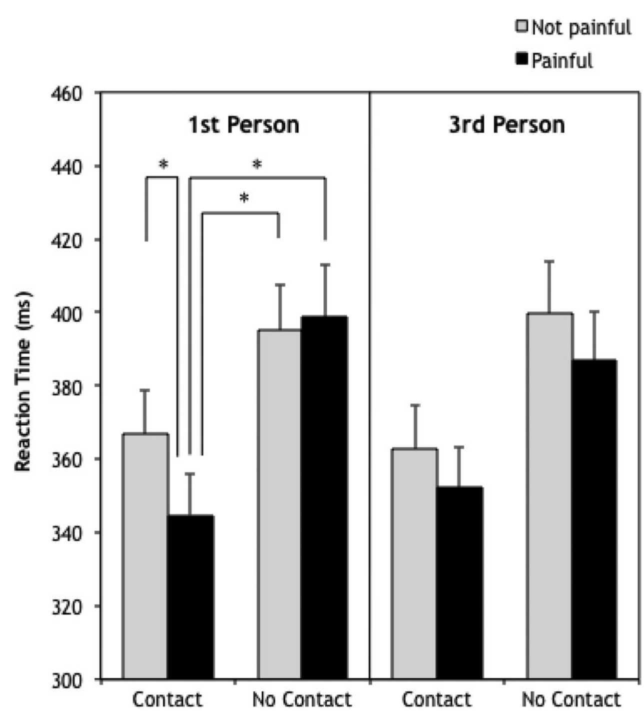

Figure 3. Mean reaction time (in milliseconds) to the detection of a tactile stimulus, while observing hands from the first-person perspective (left panel) and the third-person perspective (right panel) grasp or miss painful and nonpainful objects. Error bars are +1 *SEM*.

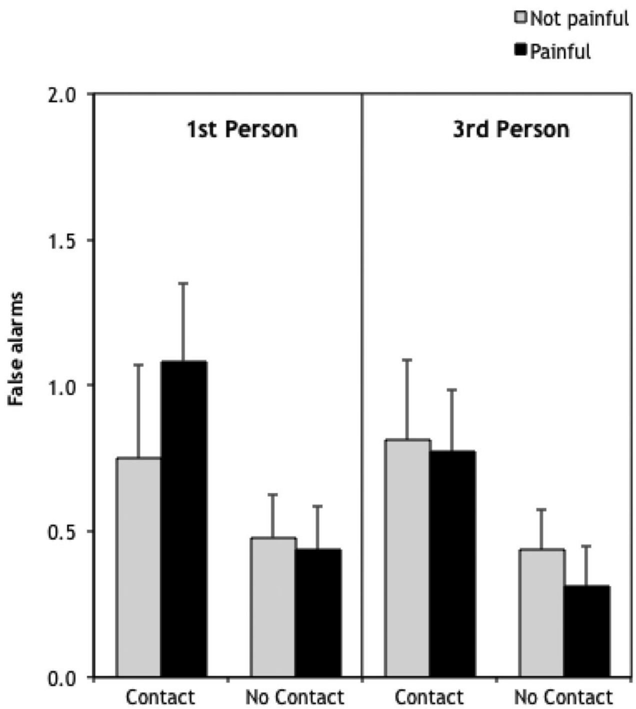

Figure 4. Mean number of false alarms (erroneous detections of tactile stimulation) while observing hands grasp or miss painful and nonpainful objects from the first-person perspective (left panel) or the third-person perspective (right panel). Error bars are +1 *SEM*.

46) = 10.61, p = .002, η_p^2 = .19. Overall, participants detected tactile stimulation more quickly when they viewed contact compared with no contact, and when they viewed painful objects compared with nonpainful objects. There was a trend toward an Action × Object interaction, $F(1, 46) = 3.64$, $p = .063$, $\eta_p^2 = .07$, and, crucially, a significant three-way Viewpoint × Action × Object interaction, $F(1, 46) = 6.14$, $p = .017$, $\eta_p^2 = .12$. No other main effects or interactions were significant (Fs < 1.92).

To better understand the three-way interaction, RTs in the first- and the third-person perspective were analyzed in separate 2 × 2 ANOVAs, with the factors of action (contact, no contact) and object (painful, not painful). Analysis of the first-person perspective trials confirmed the main effects of action, $F(1, 47) = 52.67$, $p < .001$, $\eta_p^2 = .53$, and of object, $F(1, 47) = 4.75$, $p = .034$, $\eta_p^2 = .09$, and the critical Action × Object interaction, $F(1, 47) = 7.65$, $p = .008$, $\eta_p^2 = .14$. Indeed, planned comparisons indicated that in the first-person perspective, participants responded to touch on their own fingers more quickly when seeing painful grasps than in any other condition (ps < .002), fully replicating the results of Morrison et al. (2013). In contrast, analysis of the third-person perspective only revealed the known main effects of action, $F(1, 47) = 41.31$, $p < .001$, $\eta_p^2 = .47$, and object, $F(1, 47) = 9.72$, $p = .003$, $\eta_p^2 = .17$, but, importantly, and in contrast to the first-person perspective, no evidence for an interaction, $F(1, 47) = 0.11$, $p = .737$, $\eta_p^2 < .01$.

The supplemental material provide robustness analyses of the effects testing their sensitivity to the inclusion of the third-person mapping factor, a signal detection analysis of the data, as well as an RT distribution analysis (de Jong, et al., 1994), which shows that the three-way interaction developed from the fastest to the slowest responses.

False Alarms

In a version of the current task using close-to-threshold stimulation, Morrison et al. (2013) reported a bias to report touch when observing painful grasps even in the absence of stimulation (false alarms). The current experiment was an above-threshold stimulation task, and, as such, did not prompt participants to make false alarms in the same way an ambiguous stimulus might. Nevertheless, to investigate whether a similar pattern existed in the current data, false alarms were entered into the same three-way ANOVA as the RTs. False alarms were recorded when participants pressed the spacebar to report they had felt stimulation even though no stimulation was delivered.

The number of false alarms (see Figure 4) did not differ between viewpoint ($F < 2$) or object painfulness ($F < 1$), and no Viewpoint × Action interaction ($F < 1$) was observed. There was no between-group effect of third-person mapping (anatomical match, mirror image), $F(1, 46) = 1.59$, $p = .213$, $\eta_p^2 = .03$, and no interaction with this factor. The ANOVA revealed a main effect of action, $F(1, 46) = 17.34$, $p < .001$, $\eta_p^2 = .27$, such that more false alarms were made when participants viewed hands making contact with objects ($M = 0.85$) than when the hands missed the objects ($M = 0.42$). There also was a Viewpoint × Object interaction, $F(1, 46) = 4.08$, $p = .049$, $\eta_p^2 = .08$, suggesting that the first-person perspective generally increased the likelihood for false alarms when viewing actions toward painful objects, compared with nonpainful objects. Numerically, this effect appeared to be driven by the painful grasp condition and, therefore, mirrored the RT data. However, the relevant three-way Viewpoint × Action × Object interaction was not significant, $F(1, 46) = 1.04$, $p = .312$, $\eta_p^2 = .02$.

Hits

Hits reflected the percentage of correctly detected stimulations. Participants detected a mean of 27.43 (SE = 0.14) of the 28 stimulation trials. There was no between-group effect of third-person mapping (anatomical match, mirror image), $F < 1$, $p = .642$, $\eta_p^2 < .01$. The three-way ANOVA revealed only one significant interaction between action and the between-group effect of third-person mapping, $F(1, 46) = 6.98$, $p = .01$, $\eta_p^2 = .13$. Participants who viewed anatomically matching actions in the third-person perspective made more hits when viewing grasps compared with misses, $t(23) = 2.52$, $p = .019$, but not participants who viewed mirror images. However, this effect was not relevant for the key research question.

Discussion

Actors constantly predict the consequences of their own actions, based on an integration of action and object information (Johansson & Flanagan, 2009). The present study revealed that a similar integration happens during action observation. Participants reported tactile stimulation on their own fingers while watching hands grasp or not grasp painful or neutral objects. We found that tactile stimulation was detected more quickly when participants simultaneously viewed actions with painful consequences, compared with actions that did not cause pain. This happened even though sensory action consequences were not task relevant, and no direct cues to pain were given. Our data therefore reveal that sensory consequences are predicted "on the fly" during action observation, and—similar to action execution—emerge from a combination of object knowledge (whether it causes pain) with information about the observed action (whether it makes contact with the object), rather than from either of these aspects alone.

Of note, this specific effect of observing action with painful consequences was restricted to the first-person perspective, in which the visual input matched the input one would receive from one's own actions. In the third-person perspective, the typical perspective we have on the actions of others, tactile responses only showed the more basic effects of whether the hand generally made contact with the object (whether the object was painful or not), and whether the object was generally painful (whether it was touched or not). In contrast to the first-person perspective, the two aspects were not combined to predict the sensory consequences of the actions. These differences emerged even though the visual stimulation was identical in both conditions (i.e., the images were merely mirrored), and overall RTs did not differ.

Our data therefore reveal that the first-person perspective, but not the third-person perspective, provides privileged access to mechanisms that predict an action's sensory consequences by combining action and object knowledge. This finding challenges theories that assume that sophisticated shared representations of self and other emerge from specialist brain networks that have evolved to support the understanding of the actions of others (rather than the actions of oneself). Without further assumptions, such theories predict stronger effects in third-person settings. This is the typical viewpoint from which the actions of others are observed and, therefore, presents the adaptive challenge for which such mechanisms should be specialized (cf. Oberman & Ramachandran, 2008; Schütz-Bosbach et al., 2006). Instead, our findings are in line with the idea that sensory predictions emerge from basic mechanisms that have evolved for monitoring the observer's own actions. The primary purpose of such mechanisms is predicting the consequences of these actions, such that negative outcomes are detected and course corrections can be made (Johansson & Flanagan, 2009; Csibra, 2007; Kilner et al., 2007; Miall, 2003). As found by the current study, they should therefore be specifically tuned to actions from the first-person perspective, in which the visual stimulation matches the typical input from one's own actions, but less so for the third-person perspective, which captures the typical viewpoint on the actions of others.

As predicted, in the first-person perspective, anticipation of painful action consequences sped up tactile detection on participants' own fingers. This is consistent with the proposal that sensory consequences of others' actions are represented in the observers' own tactile representation systems. It specifically supports the view that somatosensory systems combine physical stimulation and stimulation predicted from the visual input in an additive manner, such that any response threshold can be reached more quickly when both are available (Morrison et al., 2013; see Roach, McGraw, & Johnston, 2011, for evidence for a similar summation in the visual domain).

The data are unlikely to reflect more general contributions of heightened attention or sped up motor responses. To explain the specific effect of painful grasps, even such general effects must result from a prediction of painful action outcomes, based on the integration of object and action knowledge (rather than from either factor alone). In addition, several findings have supported the view that the effects reflect changes in sensory-tactile systems. First, in our prior work, which established the current procedure (Morrison et al., 2013), the effect of painful grasps was observed only for the detection of tactile—but not auditory—targets, ruling out a general attention or arousal interpretation of the effect. Second, functional magnetic resonance imaging data has confirmed that the effect was specific to the somatosensory cortices, rather than to other neural systems (e.g., relating to visual perception or motor output). Third, in the current study, the RT distribution analysis (see supplemental data) ruled out a mere priming of fast responses (i.e., due to an alerting response). Together, these results support the notion that our first-person perspective effects reflect a prediction of the sensory-tactile action consequences.

An important question is how predictive models that arose for the control of one's own action can account for effects of shared representation in third-person perspectives, which have been observed in a variety of studies, and which may form one basis of social understanding and empathy (e.g., Avenanti et al., 2013). Our findings predict that, in the absence of biasing task factors, such effects should be restricted to basic components of other people's action, such as whether an object is generally painful or whether a hand makes contact or misses. More sophisticated effects may require tasks or stimuli that encourage perspective taking and, therefore, make the third-person stimuli usable as input for first-person prediction processes. For example, our prior work has revealed sophisticated sensory−tactile prediction effects even though actions were presented in a third-person perspective (Morrison et al., 2013). In that study, however, participants were required to judge whether the action was appropriate to the object or not (grasping a painful object was inappropriate and grasping a nonpainful object was appropriate), a task that by itself biased participants to infer the actor's sensations.

Similar distinctions are also evident in the prior literature. Studies reporting automatic generation of shared representations in the third-person perspective have typically used internal states that could be directly gleaned from the stimuli without requiring integration across sources of information. In the context of tactile processing, for example, studies have manipulated whether a body part was touched, without manipulating the type of object (e.g., Keysers et al., 2004; Schaefer, Xu, Flor, & Cohen, 2009), or they varied the painfulness of an object, without manipulating whether it was touched (Meyer, Kaplan, Essex, Damasio, & Damasio, 2011). In contrast, many of the more sophisticated effects have been shown to be modulated by either encouraging or disrupting perspective taking. For example, automatic imitation (e.g., Brass, Bekkering, Wohlschläger, & Prinz, 2000) is disrupted if participants are made aware that the stimuli are virtual and do not belong to a sentient agent (Longo & Bertenthal, 2009), but is enhanced by prior social interaction (Hogeveen & Obhi, 2012). Pain empathy and prediction of others' future behavior only engages self-related brain areas when the observed people were similar to the participants or belonged to the same social group (e.g., Mitchell, Macrae, & Banaji, 2006; Avenanti, Sirigu, & Aglioti, 2010). Finally, an important modulatory factor is whether the observed behavior is relevant for one's own action planning. The curvature of an observed reach over an obstacle only affected the observers' reach when the two actors shared a workspace (Griffiths & Tipper, 2012) or when the viewed action was in the observer's own peripersonal space (Griffiths & Tipper, 2009). Thus, together, these observations are in line with our findings and suggest that deriving sophisticated shared representation of other people's internal states from a third-person perspective is not automatic, but depends on a motivation for perspective taking.

Conclusion

Shared representations have been demonstrated for almost all aspects of others' internal states, and they have been shown to form a basis of social understanding and empathy, but the origin of these mechanisms—and the boundary conditions for their activation—are still unclear. Here, we found that automatic predictions of the tactile−sensory consequences of others' actions only occurred in the first-person perspective, suggesting a specific role for mechanisms predicting the outcome of one's own actions. In contrast, our data suggest that the prediction of sensory action consequences in third-person perspectives is under cognitive control, and may only happen when perspective taking promotes deeper encoding of the actions.

References

Avenanti, A., Candidi, M., & Urgesi, C. (2013). Vicarious motor activation during action perception: Beyond correlational evidence. *Frontiers in Human Neuroscience, 7,* 185. doi: 10.3389/fnhum.2013.00185

Avenanti, A., Sirigu, A., & Aglioti, S. M. (2010). Racial bias reduces empathic sensorimotor resonance with other-race pain. *Current Biology, 20,* 1018–1022. doi:10.1016/j.cub.2010.03.071

Bach, P., Bayliss, A. P., & Tipper, S. P. (2011). The predictive mirror: Interactions of mirror and affordance processes during action observation. *Psychonomic Bulletin & Review, 18,* 171–176. doi:10.3758/s13423-010-0029-x

Bertenthal, B. I., Longo, M. R., & Kosobud, A. (2006). Imitative response tendencies following observation of intransitive actions. *Journal of Experimental Psychology: Human Perception and Performance, 32*(2), 210–225.

Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial and imitative cues. *Brain and Cognition, 44,* 124–143.

Brass, M., & Heyes, C. (2005). Imitation: Is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Science, 9,* 489–495.

Bufalari, I., Aprile, T., Avenanti, A., Di Russo, F., & Aglioti, S. M. (2007). Empathy for pain and touch in the human somatosensory cortex. *Cerebral Cortex, 17,* 2553–2561.

Caggiano, V., Fogassi, L., Rizzolatti, G., Pomper, J. K., Thier, P., Giese, M. A., & Casile, A. (2011). View-based encoding of actions in mirror neurons of Area F5 in Macaque premotor cortex. *Current Biology, 21,* 144–148. doi:10.1016/j.cub.2010.12.022

Csibra, G. (2007). Action mirroring and action understanding: An alternative account. *Sensorimotor Foundations of Higher Cognition, 22,* 427–451. doi:10.1093/acprof:oso/9780199231447.003.0020

de Jong, R., Liang, C. C., & Lauber, E. (1994). Conditional and unconditional automaticity: A dual-process model of effects of spatial stimulus-response correspondence. *Journal of Experimental Psychology: Human Perception and Performance, 20*(4), 731.

Frischen, A., Loach, D., & Tipper, S. P. (2009). Seeing the world through another person's eyes: Simulating selective attention via action observation. *Cognition, 111,* 212–218. doi:10.1016/j.cognition.2009.02.003

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience, 11,* 127–138. doi:10.1038/nrn2787

Gallese, V. (2001). The "shared manifold" hypothesis. From mirror neurons to empathy. *Journal of Consciousness Studies, 8*(5–7), 33–50.

Griffiths, D., & Tipper, S. P. (2009). Priming of reach trajectory when observing actions. Hand-centred effects. *The Quarterly Journal of Experimental Psychology, 62,* 2450–2470. doi:10.1080/17470210903103059

Griffiths, D., & Tipper, S. P. (2012). When far becomes near: Shared environments activate action simulation. *The Quarterly Journal of Experimental Psychology, 65,* 1241–1249. doi:10.1080/17470218.2012.688978

Heyes, C. (2010). Where do mirror neurons come from? *Neuroscience and Biobehavioral Reviews, 34,* 575–583. doi:10.1016/j.neubiorev.2009.11.007

Hogeveen, J., & Obhi, S. S. (2012). Social interaction enhances motor resonance for observed human actions. *The Journal of Neuroscience, 32,* 5984–5989. doi:10.1523/JNEUROSCI.5938-11.2012

Johansson, R. S., & Flanagan, J. R. (2009). Coding and use of tactile signals from the fingertips in object manipulation tasks. *Nature Reviews Neuroscience, 10,* 345–359. doi:10.1038/nrn2621

Keysers, C., & Perrett, D. I. (2004). Demystifying social cognition: A Hebbian perspective. *Trends in Cognitive Sciences, 8*(11), 501–507.

Keysers, C., Wicker, B., Gazzola, V., Anton, J. L., Fogassi, L., & Gallese, V. (2004). A touching sight: SII/PV activation during the observation and experience of touch. *Neuron, 42,* 335–346.

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.

Longo, M. R., & Bertenthal, B. I. (2009). Attention modulates the specificity of automatic imitation to human actors. *Experimental Brain Research, 192,* 739–744. doi:10.1007/s00221-008-1649-5

Meyer, K., Kaplan, J. T., Essex, R., Damasio, H., & Damasio, A. (2011). Seeing touch is correlated with content-specific activity in primary somatosensory cortex. *Cerebral Cortex, 21,* 2113–2121. doi:10.1093/cercor/bhq289

Miall, R. C. (2003). Connecting mirror neurons and forward models. *NeuroReport: For Rapid Communication of Neuroscience Research, 14*(17), 2135–2137.

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron, 50,* 655–663.

Morrison, I., Peelen, M. V., & Downing, P. E. (2007). The sight of others' pain modulates motor processing in human cingulate cortex. *Cerebral Cortex, 17,* 2214–2222. doi:10.1093/cercor/bhi129

Morrison, I., Tipper, S. P., Fenton-Adams, W. L., & Bach, P. (2013). "Feeling" others' painful actions: The sensorimotor integration of pain and action information. *Human Brain Mapping, 34,* 1982–1998. doi:10.1002/hbm.22040

Oberman, L. M., & Ramachandran, V. S. (2008). How do shared circuits develop? *Behavioral and Brain Sciences, 31,* 34–35.

Oosterhof, N. N., Tipper, S. P., & Downing, P. E. (2012). Viewpoint (in)dependence of action representations: An MVPA study. *Journal of Cognitive Neuroscience, 24,* 975–989. doi:10.1162/jocn_a_00195

Oosterhof, N. N., Wiggett, A. J., Diedrichsen, J., Tipper, S. P., & Downing, P. E. (2010). Surface-based information mapping reveals crossmodal vision–action representations in human parietal and occipitotemporal cortex. *Journal of Neurophysiology, 104*(2), 1077–1089.

Ramachandran, V. S. (2000). Mirror neurons and imitation learning as the driving force behind "the great leap forward" in human evolution. *Edge*. Retrieved from http://www.edge.org/3rd_culture/ramachandran/ramachandran_index.html

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27,* 169–192. doi:10.1146/annurev.neuro.27.070203.144230

Roach, N. W., McGraw, P. V., & Johnston, A. (2011). Visual motion induces a forward prediction of spatial pattern. *Current Biology, 21,* 740–745. doi:10.1016/j.cub.2011.03.031

Samson, D., Apperly, I. A., Braithwaite, J., Andrews, B. J., & Bodely Scott, S. E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1255–1266. doi:10.1037/a0018729

Schaefer, M., Xu, B., Flor, H., & Cohen, L. G. (2009). Effects of different viewing perspectives on somatosensory activations during observation of touch. *Human Brain Mapping, 30,* 2722–2730. doi:10.1002/hbm.20701

Schütz-Bosbach, S., Mancini, B., Aglioti, S. M., & Haggard, P. (2006). Self and other in the human motor system. *Current Biology, 16*(8), 1830–1834.

Schütz-Bosbach, S., & Prinz, W. (2007). Prospective coding in event representation. *Cognitive Processes, 8,* 93–102.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science, 303,* 1157–1162.

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. *Neuron, 40,* 655–664.

Received December 2, 2012
Revision received September 4, 2013
Accepted October 1, 2013 ■