Cognition 166 (2017) 201–206

Contents lists available at ScienceDirect

Cognition

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

Original Articles

Action recognition is sensitive to the identity of the actor

Ylva Ferstl *, Heinrich Bülthoff *, Stephan de la Rosa *

Max Planck Institute for Biological Cybernetics, Department for Perception, Cognition and Action, Spemanstr. 38, 72076 Tübingen, Germany

ARTICLE INFO

Article history:
Received 15 February 2017
Revised 24 May 2017
Accepted 26 May 2017
Available online 3 June 2017

Keywords:
Action recognition
Social cognition
Embodiment
Identity recognition

ABSTRACT

Recognizing who is carrying out an action is essential for successful human interaction. The cognitive mechanisms underlying this ability are little understood and have been subject of discussions in embodied approaches to action recognition. Here we examine one solution, that visual action recognition processes are at least partly sensitive to the actor's identity. We investigated the dependency between identity information and action related processes by testing the sensitivity of neural action recognition processes to clothing and facial identity information with a behavioral adaptation paradigm. Our results show that action adaptation effects are in fact modulated by both clothing information and the actor's facial identity. The finding demonstrates that neural processes underlying action recognition are sensitive to identity information (including facial identity) and thereby not exclusively tuned to actions. We suggest that such response properties are useful to help humans in knowing who carried out an action.

1. Introduction

Influential embodied theories of action recognition (Theory of event coding (Hommel, 2010, 2011; Hommel, Müsseler, Aschersleben, & Prinz, 2001); common coding (Prinz, 1990, 1992, 1997); direct matching (Rizzolatti, Fogassi, & Gallese, 2001)) suggest that actions are represented by their sensory consequences. Essential to these theories is that one's own and another person's actions are encoded by the same cognitive representation (hereafter referred to as common code), thereby suggesting an equivalence between the perception and execution of actions. Although these theories provide an elegant explanation for how an observer can efficiently appreciate the consequences of an observed action, they raise the fundamental question of how the brain knows who is carrying out an action: is it the observer or the other person? This problem arises because an activation of a common action code could be due to the execution of the action (i.e. oneself) or due to action observation (i.e. another person). While previous research provided empirical evidence against own and other actions being represented in the exactly same way (Schütz-Bosbach, Mancini, Aglioti, & Haggard, 2006), the underlying mechanisms for the differentiation are still poorly understood.

An answer to this problem was recently provided by Dolk (2011), Dolk, Hommel, Prinz, and Liepelt (2013), Dolk et al. (2014) within their referential coding account. According to this account, action representations also encode information about the spatial location of an action. Because own actions are associated with a different spatial location than other persons' actions, the spatial location can be used to determine which actor (i.e. oneself or another person) caused the activation of an action representation. This theory provides an explanation for how the cognitive system could distinguish between one own and another person's action.

It is important to note that the encoding of the spatial location of an action provides the answer to where an action was carried out but not to who carried out an action. This latter question is, however, of great importance for a many social cognitive processes, especially when one observes the actions of a group of people. Take for example, the Sally Ann theory of mind test (Baron-Cohen, Leslie, & Frith, 1985; Wimmer & Perner, 1983). The successful passing of this test requires participants to know who left the room and who placed the marble in the box. Hence, successful functioning in a social environment requires the observer to associate an action with person.

How does the cognitive system associate an action with an actor? Interestingly, current neuroscientific evidence suggests a loose coupling between a person's identity and bodily actions. Research about identity has so far focused on the most prominent cue to a person's identity, facial identity. For example, facial identity is processed in cortical areas (FFA, ATL, OFA) (Anzellotti & Caramazza, 2014; Haxby, Hoffman, & Gobbini, 2000; Nestor, Plaut, & Behrmann, 2011) that have been argued to be selectively sensitive for faces but not bodies (Peelen & Downing, 2005; Schwarzlose, Baker, & Kanwisher, 2005). Moreover, biological

* Corresponding authors.
E-mail addresses: yferstl@tcd.ie (Y. Ferstl), heinrich.buelthoff@tuebingen.mpg.de (H. Bülthoff), delarosa@tuebingen.mpg.de (S. de la Rosa).

http://dx.doi.org/10.1016/j.cognition.2017.05.036
0010-0277/

motion processing has been associated with posterior regions in the superior temporal sulcus (Beauchamp, Lee, Haxby, & Martin, 2003; Pelphrey et al., 2003) that were shown not to be correlated with facial identity (Peelen & Downing, 2007; Peelen, Wiggett, & Downing, 2006). Some evidence even argues for body identity being processed in areas dissociate from action discrimination (Urgesi, Candidi, Ionta, & Aglioti, 2007). Overall, these results suggest that actions and a person's identity are represented in separate neural substrates.

Nevertheless, a person's identity is undoubtedly relevant in action processing and identity recognition is crucial for interpreting an action. Philipp and Prinz (2010) have for example shown that task-irrelevant identity information can influence action behavior. Furthermore research about social top-down influences on motor processing acknowledges that an actor's identity can modulate action perception (e.g. Shiffrar & Freyd, 1990; Pelphrey, Morris, & Mccarthy, 2004; Knoblich & Sebanz, 2006). These accounts, however, do not provide an understanding of the underlying mechanisms of the actor-action association.

We revisited the view about the independence of bottom-up processing of action cues from identity by examining the sensitivity of neural action recognition processes to the actor's identity using an action adaptation paradigm. Adaptation paradigms are a powerful tool to behaviorally examine tuning characteristics of neural processes (Barraclough, Ingham, & Page, 2012; Barraclough & Jellema, 2011; Barraclough, Keith, Xiao, Oram, & Perrett, 2009; Dinstein, Hasson, Rubin, & Heeger, 2007; Grill-Spector & Malach, 2001; Webster, 2011; Webster & MacLeod, 2011). Importantly, recent evidence showed that behavioral action adaptation effects similar to the ones employed in the present study correlate with response changes in right pSTS (Thurman, van Boxtel, Monti, Chiang, & Lu, 2016), demonstrating the usefulness of behavioral adaptation for exploring neural processes underlying action recognition.

In brief, adaptation refers to the change in perception of an ambiguous test stimulus (e.g. a morph between a hug and a push) after the prolonged exposure to an unambiguous adaptor (e.g. a hug or a push). This change in perception is attributed to a transient response change of the neural population underlying the probed visual task during the adaptation period. If the perception of the test stimulus relies at least in part on the same neural populations as the adaptor stimulus, then the response change induced during the adaptation period is assumed to affect the perception of the test stimulus. By examining which visual differences between test and adaptor stimulus are able to modulate the adaptation effect, one can assess the tuning properties of the involved neural populations.

In previous research, we have demonstrated that action adaptation effects are not merely the result of a response change on an abstract non-visual decision level. Rather, action adaptation effects are bound to visual action information. For example, attending to different kinds of visual information of the same action (e.g. to the type of action or the movement direction) alters the magnitude and direction of adaptation effects (De La Rosa, Ekramnia, & Bülthoff, 2016). Moreover, substituting visual action information with linguistic action information (e.g. using action words instead of action images) completely abolishes the action adaptation effect (De La Rosa, Streuber, Giese, Bülthoff, & Curio, 2014). Taken together these results suggest that visual action adaptation effects can be reliably linked to visual action information.

Here, we took advantage of the adaptation paradigm to measure the sensitivity of action recognition processes to several potential sources of identity information of the actor. Specifically, the two persons typically differ with respect to their clothing and facial identity, whereby the latter is surely a more reliable cue an actor's identity. In the current study we probed action recognition processes underlying the human ability to tell two actions apart (action categorization) for their sensitivity to clothing and facial identity information. We reasoned that if visual action recognition processes are sensitive to the actor's identity, then action adaptation effects should also depend on the actor's identity. Specifically, in this case we expected action adaptation effects to be stronger when the same actor carries out adaptor and test actions (same identity condition) compared to when different actors carry out adaptor and test action (different identity condition). To increase the external validity of our results, we used a novel augmented reality setup in which participants observed actions carried out by a life-size, human-looking, three dimensional avatar. After participants repeatedly observed the avatar carrying out an adaptor action (hug or push) they subsequently categorized an ambiguous morphed action carried out by either the same or a different avatar as either hug or push.

To this end, we conducted three experiments in which we explored the effect of clothing and facial identity (Experiment 1), facial identity (Experiment 2), and clothing (Experiment 3) of the actor on action adaptation effects.

2. Material and methods

2.1. Participants

In each of the three experiments we tested a new set of 24 participants from the local community in Tübingen who gave written informed consent regarding their participation. The study was approved by the ethics review board of the University of Tübingen and all experiments were performed in accordance with relevant guidelines and regulations.

2.2. Apparatus & stimuli

We used a large screen augmented reality setup in which two three-dimensional, life-sized, male avatars carried out the actions. The use of avatars allowed us to change the identity (as defined by face texture, body texture, and mesh shape) of the avatar without altering the motion. Using this setup, participants visually adapted to a hug or a push action and reported their subjective impression about a morphed action. The morphed actions were calculated as a weighted average of corresponding joint angles between the motion-captured hug and push actions. We used seven different weights, hence as a result there were seven different ambiguous test actions each containing elements of both a hug and a push. Participants' task was to report whether they perceived a test action more as a push or a hug. Participants adapted to a hug and to a push action in separate conditions and we measured participants' proportion of 'push' responses in all experimental conditions. To measure the adaptation effect we subtracted the proportion of 'push' responses in the hug adaptation condition from the ones in the push condition (adaptation difference). Hug and push conditions were tested in four separate experimental conditions in which we varied the actor's identity between the adaptor and test stimuli independently. Hence, we compared the adaptation effects between same (Actor $A_{adaptor}$-Actor A_{test}; Actor $B_{adaptor}$-Actor B_{test}) and different (Actor $A_{adaptor}$-Actor B_{test}; Actor $B_{adaptor}$-Actor A_{test}) actor conditions, each for both adaptor actions, resulting in a total of eight experimental conditions.

2.3. Procedure

At the very beginning of an experiment, the participant saw each of the seven different test stimuli three times in random order without any adaptation to measure action categorization in the