Consciousness and Cognition 36 (2015) 434–437

Contents lists available at ScienceDirect

Consciousness and Cognition

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

Domain-specific and domain-general processes in social perception – A complementary approach

John Michael [a,*], Alessandro D'Ausilio [b]

[a] Central European University, Budapest, Hungary
[b] Italian Institute of Technology, Italy

ARTICLE INFO

Article history:
Received 1 December 2014
Accepted 11 December 2014
Available online 30 December 2014

Keywords:
Social perception
Perspective-taking
Action understanding
Spatial cueing
Gaze cueing
Associative learning

ABSTRACT

In this brief discussion, we explicate and evaluate Heyes and colleagues' deflationary approach to interpreting apparent evidence of domain-specific processes for social perception. We argue that the deflationary approach sheds important light on how functionally specific processes in social perception can be subserved at least in part by domain-general processes. On the other hand, we also argue that the fruitfulness of this approach has been unnecessarily hampered by a contrastive conception of the relationship between domain-general and domain-specific processes. As an alternative, we propose a complementary conception: the identification of domain-general processes that are engaged in instances of social perception can play a positive, structuring role by adding additional constraints to be accounted for in modelling the domain-specific processes that are also involved in such instances.

In an important series of recent papers, Heyes and colleagues (Cook, Bird, Catmur, Press, & Heyes, 2014; Heyes, 2014; Santiesteban, Catmur, Hopkins, Bird, & Heyes, 2014; Catmur, Walsh, & Heyes, 2007) have developed a deflationary approach to interpreting apparent evidence of domain-specific processes for social perception. Briefly, the approach consists in demonstrating how the apparent evidence can be accounted for by domain-general processes, i.e. without postulating any additional, domain-specific processes that track mental states. If successful, the resulting deflationary explanations would in each case show how findings observed in paradigms that seem to tap specifically *social* perception can in fact be brought about directly by domain-general processes – i.e. bypassing any need for dedicated social-perceptual to identify others' mental states. In the following, we explicate and evaluate this deflationary approach. Focusing on two separate areas of research to which the approach has been applied (visual perspective-taking and action understanding), we argue that the deflationary approach has been successful in illuminating how functionally specific processes for social perception can be subserved at least in part by domain-general processes, but that its potential to contribute to ongoing research has been limited by an artificial dichotomy between domain-general and domain-specific processes. As an alternative to this *contrastive* conception of the relationship between domain-general and domain-specific processes, we propose a *complementary* conception: the identification of domain-general processes that are engaged in instances of social perception can play a positive, structuring role by adding additional constraints to be accounted for in modelling the domain-specific processes that are also involved in such instances.

* Corresponding author at: Department of Cognitive Science, Central European University, Frankel Leo Ut 30-34, 1023 Budapest, Hungary. Fax: +36 305 399 329.
E-mail address: michaelj@ceu.hu (J. Michael).

http://dx.doi.org/10.1016/j.concog.2014.12.009
1053-8100/

Let us begin by considering a current dispute about the results from a paradigm purportedly tapping an automatic process for calculating other agents' visual perspectives. In the basic paradigm (Samson, Apperly, Braithwaite, Andrews, & Bodley Scott, 2010), participants view an image of a room in which an avatar is standing in the middle and facing either to the left or to the right (this is varied from one trial to the next). On each trial, anywhere from zero to three red discs are displayed on the walls of the room, such that on some trials the avatar can see all of them (if she, for example, is facing to the left, and three discs appear on the wall to the left but none appear on the wall to the right), but on other trials the avatar can only see some but not all of the discs (e.g. if she is facing to the left, and one disc appears on the wall to the right). Participants have the task of either calculating how many discs the avatar can see or calculating how many discs they themselves see (this is varied from one trial to the next). The basic finding on the latter type of trial is that participants' performance is impaired when they themselves can see a different number of discs than the avatar can see. In other words, participants appear to have been calculating how many discs the avatar could see even though it was irrelevant to, and indeed interfering with, their task (Samson et al., 2010). In a follow-up study using the same paradigm, Qureshi, Apperly, and Samson (2010) presented participants with an additional cognitive load during the experiment, and found that the interference from the avatar's perspective increased. The authors interpret this as evidence that participants calculated what the avatar could see (level-1 perspective taking) automatically, in parallel to the calculation of what the subject herself could see. As the authors put it: 'This is the first direct evidence of a cognitively efficient process for "theory of mind" in adults that operates independently of executive function' (Qureshi et al., 2010: 230). In contrast, the selection of which perspective to draw upon in judging the number of discs is a controlled process requiring executive resources, and is therefore impaired by the cognitive load manipulation.

This pattern of results is highly suggestive of a dedicated and automatic process for level-1 visual perspective-taking. Nevertheless, it is important to be cautious. As Heyes and colleagues have pointed out (Heyes, 2014; Santiesteban et al., 2014), it is possible that the effect is at least partially driven by domain-general spatial-cueing, with the avatars serving as cues to trigger attention either to the left or to the right. In fact, this interpretation is supported by the findings of Santiesteban et al. (2014), who replicated Samson et al.'s effect using arrows instead of avatars. According to Heyes (2014) and Santiesteban et al. (2014), we should conclude that the paradigm in fact does not tap any specifically *social*-perceptual or *social*-cognitive process at all (i.e. perspective-taking). Instead, Heyes argues, we should conclude that the effect is driven by 'submentalizing processes', i.e. by 'domain-general cognitive processes that do not involve thinking about mental states but can produce in social contexts behaviour that looks as if it is controlled by thinking about mental states' (2014: 132). In other words, the effect observed in the paradigm could be brought about by a domain-general attentional process, bypassing any dedicated mechanism for tracking other agents' perspectives.

But while Santiesteban et al.'s findings do indeed suggest that a domain-general attentional process may be playing an important role in this paradigm, it would be overly hasty to conclude from this that the paradigm does not also involve domain-specific processes for identifying agentic features (e.g. a human-like body, the gaze direction of the avatar). Indeed, human or human-like bodies are uniquely salient and biologically significant cues. It is therefore unsurprising that when presented with images of scenes containing human faces as well as arrows, people are far more likely to fixate on the human faces and to follow the gaze direction of those faces than to fixate on the arrows – even if the arrows are much larger and more prominently positioned. Moreover, there is ample research to motivate the conjecture that Samson and colleagues' paradigm engages a medley of domain-specific processes, and that the effect tapped in the paradigm is not only similar to spatial cueing using arrows, but also, and in important ways, different from it. First of all, it has been shown that gaze cueing, unlike spatial cueing with arrows, is automatic in the sense that faces (but not arrows) trigger spatial cueing even if the gaze direction of the face has very low cue validity (20%), and participants are informed of this (Driver et al., 1999; Friesen & Kingstone, 1998). Secondly, participants tend to evaluate faces in quasi-moral terms (i.e. as trustworthy or untrustworthy) depending on the cue validity of their gaze direction (Bayliss & Tipper, 2006). Thirdly, by systematically varying not only the locations of targets, but also the objects (i.e. rectangular figures) in which those targets appeared, Marotta, Lupiánez, Martella, and Casagrande (2012) were able to show that faces, unlike arrows, trigger a pure location-based cueing effect, whereas arrows, unlike faces, trigger a pure object-based cueing effect. They interpret this finding as evidence that faces and arrows engage qualitatively different (i.e. location-based versus object-based) orienting mechanisms.

These findings strongly suggest that social cues, such as human-like bodies and faces, function differently from non-social cues – even though they may also engage a common attentional process. This highlights the importance of investigating similarities and differences between avatars and other types of cue (such as arrows) in the Samson paradigm. In other words, a moderate application of Santiesteban and colleagues' deflationary strategy points in the direction of interesting new hypotheses for further research rather than to mere debunking. Specifically, a complementary approach that incorporates a moderate version of the deflationary strategy generates the prediction – contra Santiesteban and colleagues – that reducing the validity of the avatar's gaze direction to 20% would have little effect upon participants' performance. On the other hand, contra Samson and colleagues' interpretation of the finding as evidence of a mechanism for automatic perspective-taking, a complementary approach predicts that various other systems can cooperate with the systems engaged in this paradigm. Thus, we predict that performance in the Samson paradigm could be modulated by manipulating participants' social knowledge, such as their beliefs about the avatar, e.g. about whether the avatar is sighted or blind, whether a pair of goggles that s/he is wearing is transparent or opaque, etc. Intriguingly, this latter prediction is motivated by the findings of Teufel et al. (2009), who reported that participants' processing of gaze direction was facilitated when a subject believed that a person wearing goggles was able to see through them (as opposed to the goggles being opaque). In sum, the identification of a domain-general attentional process (spatial cueing) that is engaged in a social context raises important further questions about how that domain-general process functions in this specific context, and about what other domain-specific social-perceptual and social-cognitive processes may also be at work.

Let us now turn to a different example, namely to the ongoing dispute about a possible contribution of the motor system in identifying the goals, or underlying intentions, of others' actions (sometimes termed 'action understanding'. As is by now well known, this dispute was initiated by the discovery of so-called 'mirror neurons' in macaque monkeys, i.e. neurons that are active when an animal is either performing a particular action or observing another agent performing the same or a similar action (Di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). Since it is only rare that single-cell activity can be recorded in humans (but see Mukamel, Ekstrom, Kaplan, Iacoboni, & Fried, 2010), most research involving human participants has been performed with imaging techniques that measure activity in brain regions rather than individual cells. This research has revealed a network, comprising regions of premotor cortex (PMC), inferior parietal lobule (IPL) and somatosensory areas (Gazzola & Keysers, 2009; Vogt et al., 2007; Buccino et al., 2001; Iacoboni et al., 2005; Rizzolatti & Craighero, 2004), that is 'activated during performance of [an] action as well as during observation of the same action' (Frith & Singer, 2008). On the basis of these findings, it has been proposed that this network, sometimes called the 'mirror neuron system' (MNS), may constitute a dedicated, and perhaps evolved, mechanism for action understanding (i.e. for identifying the goals, or underlying intentions, of other agents' actions) (Gallese, Keysers, & Rizzolatti, 2004; Gazzola & Keysers, 2009; Iacoboni et al., 2005; Kilner, Neal, Weiskopf, Friston, & Frith, 2009). These claims have remained controversial, however, with competing models offering conflicting accounts of the function of MNS activation. To some extent, this is because a great deal of the research in humans has involved correlational methods, making it difficult to rule out various deflationary explanations that do not attribute a causal role to the MNS in action understanding (but see Cattaneo, Sandrini, & Schwarzbach, 2010; Michael et al., 2014; Pobric & Hamilton, 2006; Tidoni, Borgomaneri, di Pellegrino, & Avenanti, 2013; Urgesi, Candidi, Ionta, & Aglioti, 2006).

Against this background, Cook et al. (2014) have proposed a deflationary explanation according to which the activation of motor areas during passive observation may result from domain-general associative learning (Cook et al., 2014). In other words, there is a high degree of association between motor commands, on the one hand, and the sight of the movements initiated by those motor commands, on the other – e.g. the experience of moving one's hand is tightly associated with the sight of a moving hand. If this is the case, then the findings in question may not reflect a dedicated mechanism for action understanding but may instead be a mere byproduct of domain-general associative learning – and activation of the MNS would result directly from the visual perception of movements without being shaped by or contributing to the representation of intentions or any other mental states. Consistent with this hypothesis, earlier work by the same research group has demonstrated that the response profiles of neural circuits in the MNS can be modulated, or even reversed, by relatively short training periods (Catmur et al., 2007).

In this case, as in the case of visual perspective-taking, it is important to be cautious in reflecting on the consequences of the deflationary approach. First of all, it must be emphasized that the associationist account is compatible with the possibility that neural circuits with both motor and perceptual properties, though shaped by a domain-general learning process, may nevertheless take on domain-specific functions, such as the function of contributing to action understanding. Thus, an account of the origins of the MNS does not by any means save us the trouble of carrying out targeted research to address the separate issue of function. Next, it must also be emphasized that while the associationist account may provide at least part of an account of the origins of the MNS, there is no reason whatsoever to expect that it is incompatible with other, more domain-specific, factors also playing a role in the learning process. On the contrary, a fruitful way to make use of the insights generated by the associationist approach would be to attempt to pinpoint limitations of associationist learning, and to articulate and test hypotheses about processes that may subserve functions for which associationist learning is not sufficient. For example, the associationist account is not well-suited to explaining instances in which agents are able to use their own motor capacities in identifying other agents' goals in the absence of associative links between those motor capacities and visual information about the effects of those movements. Thus, D'Ausilio et al. (2014) were able to induce a version of the McGurk effect without participants actually seeing the face of a speaker but, instead, a visual representation of the speaker's tongue. Given that the participants had presumably not had the occasion to observe their own or others' tongues while producing the relevant sounds, it is highly implausible to appeal to associative learning in explaining this finding. Instead, it is much more likely that the participants' perception was biased by their specific motor knowledge about how they move their tongue when producing speech.

More generally speaking, the associationist account is not well-suited to account for cases in which there is no clear association between an action and its visual appearance. While D'Ausilio and colleagues' finding presents an extreme example of such a case (the absence of any association), there may in general be a gradient of association strength – i.e. some body parts, such as the hands, may afford the formation of far stronger associations than other parts that are less accessible to vision (e.g. the shoulders and back; faces, if it were not for mirrors). But there is no reason to believe that a strong visuo-motor association is a necessary prerequisite for the actions and/or expressions produced by a body part to be intelligible for an observer. If that were the case, it would mean that our ancestors living in cultures without mirrors must have been unable to understand facial expressions.

Moreover, even in cases in which associative learning is likely to play a role, there is no reason to assume that it should be the only or the most important process at work. In most instances in which we produce or observe actions (or both), there are a myriad possible associations that could be made. It is therefore important to consider how these possibilities could be constrained. For example, one possible way is to identify the goals that specific actors are likely to be pursuing in specific situations, and the means that are most efficacious in bringing about those goals. By ignoring such factors, a strict associationist account is forced to endorse the prediction that stronger visuomotor associations correlate in a linear fashion with stronger mirror-like effects. However, the strength of the latter seems instead to be linked to the relevance of a given effector (i.e. hands, face) for social interaction. In fact, such social-relevance weighting maps nicely to the motor and somatosensory homunculus geometry, with hands and face being represented most prominently. The 'body distortions' observed in the homunculus are based on the density of sensory receptors as well as on our capacity for fine motor control of that body part. In brief, those body parts are critical in all our (social and non-social) activities, and this might be the metric upon which mirror-like activities are built – not pure visuomotor association.

In both of the cases we have touched upon in this brief commentary, the deflationary approach developed by Heyes and colleagues sheds important light on how functionally specific social cognitive processes can be subserved at least in part by domain-general processes. On the other hand, we have also argued that the fruitfulness of this approach has been unnecessarily hampered by a contrastive conception of the relationship between domain-general and domain-specific processes. We have therefore proposed a complementary conception: the identification of domain-general processes that are engaged in instances of social perception can lead to specific new research questions about the domain-specific processes that are also engaged in such instances.

References

Bayliss, A. P., & Tipper, S. P. (2006). Predictive gaze cues and personality judgments should eye trust you? *Psychological Science, 17*(6), 514–520.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience, 13*, 400–404.

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

Cattaneo, L., Sandrini, M., & Schwarzbach, J. (2010). State-dependent TMS reveals a hierarchical representation of observed acts in the temporal, parietal, and premotor cortices. *Cerebral Cortex, 20*(9), 2252–2258.

Cook, R., Bird, G., Catmur, C., Press, C., & Heyes, C. M. (2014). Mirror neurons: From origin to function. *Behavioural and Brain Sciences, 37*, 177–241.

D'Ausilio, A., Bartoli, E., Maffongelli, L., Berry, J. J., & Fadiga, L. (2014). Vision of tongue movements bias auditory speech perception. *Neuropsychologia, 63*, 85–91.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research, 91*, 176–180.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition, 6*(5), 509–540.

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin and Review, 5*, 490–495.

Frith, C., & Singer, T. (2008). The role of social cognition in decision-making. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*, 3875–3886.

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

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *TiCS, 8*(9), 396–403.

Gazzola, V., & Keysers, C. (2009). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: Single-subject analyses of unsmoothed fMRI data. *Cerebral Cortex, 19*, 1239–1255.

Heyes, C. M. (2014). Submentalizing: I'm not really reading your mind. *Perspectives on Psychological Science, 9*, 131–143.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology, 3*, 529–535.

Kilner, J., Neal, A., Weiskopf, N., Friston, K., & Frith, C. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *Journal of Neuroscience, 29*(32), 10153–10159.

Marotta, A., Lupiánez, J., Martella, D., & Casagrande, M. (2012). Eye gaze versus arrows as spatial cues: Two qualitatively different modes of attentional selection. *Journal of Experimental Psychology: Human Perception and Performance, 38*(2), 326–335.

Michael, J., Sandberg, K., Skewes, J., Wolf, T., Blicher, J., Overgaard, M., et al (2014). TMS (cTBS) demonstrates a causal role of premotor homunculus in action understanding. *Psychological Science, 25*, 963–972.

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). *Current Biology, 20*, 750–756.

Pobric, G., & Hamilton, A. (2006). Action understanding requires the left inferior frontal cortex. *Current Biology, 16*, 524–529.

Qureshi, A. W., Apperly, I. A., & Samson, D. (2010). Executive function is necessary for perspective selection, not Level-1 visual perspective calculation: Evidence from a dual-task study of adults. *Cognition, 117*, 230–236.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27*, 169–192.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3*, 131–141.

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley 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.

Santiesteban, I., Catmur, C., Hopkins, S., Bird, G., & Heyes, C. M. (2014). Avatars and arrows: Implicit mentalizing or domain general processing? *Journal of Experimental Psychology: Human Perception and Performance, 40*(3), 929–937.

Teufel, C., Alexis, D. M., Todd, H., Lawrance-Owen, A. J., Clayton, N. S., & Davis, G. (2009). Social cognition modulates the sensory coding of observed gaze direction. *Current Biology, 19*(15), 1274–1277.

Tidoni, E., Borgomaneri, S., di Pellegrino, G., & Avenanti, A. (2013). Action simulation plays a critical role in deceptive action recognition. *The Journal of Neuroscience, 33*(2), 611–623.

Urgesi, C., Candidi, M., Ionta, S., & Aglioti, S. (2006). Representation of body identity and body actions in extrastriate body area and ventral premotor cortex. *Nature Neuroscience, 10*(1), 30–31.

Vogt, S., Buccino, G., Wohlschlager, A. M., Canessa, N., Shah, N. J., Zilles, K., et al (2007). Prefrontal involvement in imitation learning of hand actions: Effects of practice and expertise. *Neuroimage, 37*(4), 1371–1383.