

Dissociating the Detection of Intentionality from Animacy in the Right Posterior Superior Temporal Sulcus

Tao Gao, Brian J. Scholl, and Gregory McCarthy

Department of Psychology, Yale University, New Haven, Connecticut 06520-8205

Certain motion patterns can cause even simple geometric shapes to be perceived as animate. Viewing such displays evokes strong activation in temporoparietal cortex, including areas in and near the (predominantly right) posterior superior temporal sulcus (pSTS). These brain regions are sensitive to socially relevant information, but the nature of the social information represented in pSTS is unclear. For example, previous studies have been unable to explore the perception of shifting intentions beyond animacy. This is due in part to the ubiquitous use of complex displays that combine several types of social information, with little ability to control lower-level visual cues. Here we address this challenge by manipulating intentionality with parametric precision while holding cues to animacy constant. Human subjects were exposed to a “waving wolf” display, in which one item (the wolf) chased continuously, but its goal (i.e., the sheep) frequently switched among other shapes. By contrasting this with three other control displays, we find that the wolf’s changing intentions gave rise to strong selective activation in the right pSTS, compared with (1) a wolf that chases with a single unchanging intention, (2) very similar patterns of motion (and motion change) that are not perceived as goal-directed, and (3) abrupt onsets and offsets of moving objects. These results demonstrate in an especially well controlled manner that right pSTS is involved in social perception beyond physical properties such as motion energy and salience. More importantly, these results demonstrate for the first time that this region represents perceived intentions beyond animacy.

Introduction

Vision recovers not only the physical structure of the external world, but also its causal and social structure. For example, even simple geometric shapes can give rise to robust percepts of animacy and intentionality when they move in certain ways (Heider and Simmel, 1944; for review, see Scholl and Tremoulet, 2000). That our visual system is adapted to extract socially relevant information from even such impoverished stimuli demonstrates the fundamental role that perception plays in social cognition, and vice versa. Viewing such displays evokes strong activation in temporoparietal cortex, including areas in and near the posterior superior temporal sulcus (pSTS) and angular gyrus, particularly in the right hemisphere (Castelli et al., 2000; Blakemore et al., 2003; Martin and Weisberg, 2003; Schultz et al., 2004, 2005; Heberlein, 2008). The pSTS is also strongly activated by biological motion of human avatars (Pelphrey et al., 2003) and point-light displays (Bonda et al., 1996; Grossman et al., 2000). These converging results suggest that pSTS plays an important role in representing information related to social agency.

But what is the nature of this information? Despite the cartoon-like motion and the use of simple geometric shapes, displays such as those of Heider and Simmel (1944) and their modern counterparts are visually and cognitively complex. In particular, they confound animacy detection with the detection of related social properties (e.g., intentions, goals, and emotions), and often also with asocial properties (such as salience and attentional capture). Our present goal is to isolate a role for the perception of shifting intentions, while holding cues to perceived animacy constant. Animacy and intentions are separate but interdependent constructs. For example, one can readily perceive animacy without perceiving intentionality, as when a cue such as self-propulsion leads one to perceive a moving shape as alive, even when one cannot divine any particular goal. (This seems to occur, for example, when watching swarms of insects fly around in seemingly chaotic patterns—clearly alive, but without any obvious intention.)

Here we created displays in which simple discs appeared to move in a self-propelled manner, a robust cue for perceived animacy (Tremoulet and Feldman, 2000). We then exploited a powerful cue for perceived intentionality: chasing behavior (Gao et al., 2009; Gao and Scholl, 2011), wherein one moving disc (the sheep) is pursued by another disc (the wolf). In the simplest case, the wolf pursued a single sheep for the duration of the display. In the “waving wolf” display, the wolf chased two sheep in alternation. Here, cues to animacy (such as self-propulsion and chasing) are held constant, even though the perceived intentions of the wolf may shift dramatically (now pursuing one target, now another). With the inclusion of additional control displays, we were thus able to dissociate the role of temporoparietal cortex,

Received Feb. 6, 2012; revised June 13, 2012; accepted July 26, 2012.

Author contributions: T.G., B.J.S., and G.M. designed research; T.G. performed research; T.G. and G.M. analyzed data; T.G., B.J.S., and G.M. wrote the paper.

This work was supported by NIH Grants NS41328 and MH05286 to G.M. We thank William Walker and Dr. George He for their help in data acquisition and analysis.

The authors declare no financial conflicts of interest.

Correspondence should be addressed to Gregory McCarthy, Department of Psychology, Yale University, PO Box 208205, New Haven, CT 06520-8205. E-mail: gregory.mccarthy@yale.edu.

DOI:10.1523/JNEUROSCI.0562-12.2012

Copyright © 2012 the authors 0270-6474/12/3214276-05\$15.00/0

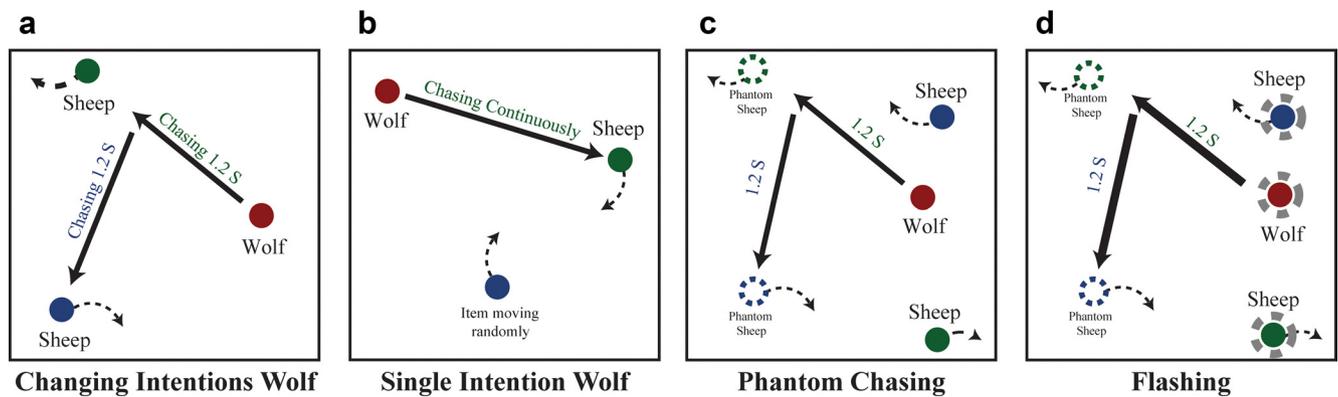


Figure 1. *a*, An illustrative static frame from the dynamic Changing Intentions display in which the red wolf alternated chasing the blue and green sheep every 1.2 s. *b*, In the Single Intention trial, the wolf chased one sheep continuously. *c*, In the Phantom Chasing display, the wolf alternated chasing the mirrored location of each sheep every 1.2 s. This display preserved the abrupt motion changes of the wolf and the correlation of its motion with the sheep. However, the target of its chase was invisible. *d*, The Flashing display was identical to Phantom Chasing except that the three discs irregularly flashed on and off during their motion.

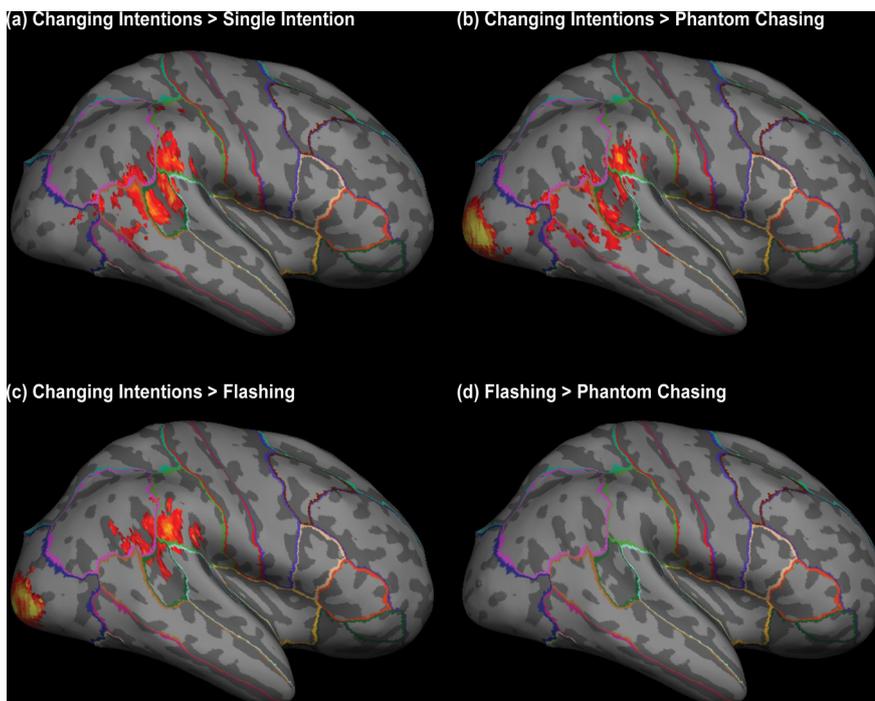


Figure 2. *a–d*, Key contrasts from our GLM analysis displayed upon an inflated representation of the right MNI152 hemisphere: Changing Intentions > Single Intention (*a*), Changing Intentions > Phantom Chasing (control) (*b*), Changing Intentions > Flashing (*c*), Flashing > Phantom Chasing (*d*). All contrasts have been corrected for multiple comparisons ($p < 0.05$).

and in particular the pSTS, in both animacy detection (using more rigorous displays than in past studies) and the perception of shifting intentions (isolated here for the first time).

Materials and Methods

Subjects. Fifteen subjects (6 female, 9 male; mean age = 23.7 years) with normal or corrected-to-normal vision and no history of neurological or psychiatric illness participated in the study. The Yale University Human Investigations Committee approved the study and all subjects provided written informed consent. The experiment consisted of four runs, each consisting of 24 randomly ordered trials, with six trials for each of four display types. There was a 2 s fixation period between trials.

Stimulus displays. There were four display types (Fig. 1): (1) Changing Intentions (or wavering wolf), (2) Single Intention, (3) Phantom Chasing, and (4) Flashing displays. The Changing Intentions display (Fig. 1*a*) contained three colored discs (0.63°) that moved about on an 11.32° by

11.32° back-projected display that the subject viewed with a mirror mounted on the head coil of the MRI system. The discs were randomly assigned the colors red, green, and blue on each trial. At the beginning of each trial, each disc immediately began moving at a constant speed of 9.49°/s. The two discs designated as sheep randomly changed motion direction within a 90° window (centered on the current heading) approximately every 167 ms. (On each 16.7 ms frame of motion, each disc had a 10% chance of changing its direction; in practice, this led to direction changes every 167 ms on average.) The disc designated as the wolf also adjusted its direction of motion every 167 ms, so that it was displaced approximately in the direction of the currently chased sheep. This pursuit was not perfectly direct or “heat-seeking,” but occurred with a random deviation $< 15^\circ$ (clockwise or counterclockwise). The motion trajectories of each trial were algorithmically generated with the constraint that the distance between any two items always exceeded 2° on each frame. Every 1.2 s, the wolf switched the target of its pursuit: instead of being displaced each frame in the rough direction of the current sheep (which then became the old sheep), the displacements began occurring in the direction of the other sheep (which then became the current sheep).

The Single Intention display was identical to the Changing Intentions display with the important exception that the wolf chased a single sheep for the entire 10 s display (Fig. 1*b*). The third (unchased) sheep moved along a random trajectory of the same type as the chased sheep, but its motion was not correlated (i.e., it did not interact) with either the wolf or the chased sheep. Although not forewarned about the identity (color) of the wolf, its intentional, goal-directed, chasing behavior was easily ascertained in both the Changing Intentions and Single Intention Wolf displays, yielding a robust cue to perceived animacy (Gao et al., 2009; Gao and Scholl, 2011). However, the perceived intentions of the wolf are held constant in the Single Intention display but are seen to shift regularly and dramatically in the Changing Intentions display, and thus the difference in activation between these two conditions constitutes the comparison of primary interest.

Because changing intentions were signaled by abrupt changes in motion direction in the Changing Intentions display, we included a Phantom Chasing control condition in which identical abrupt changes in

motion occurred, while also maintaining all correlations between the motions of the wolf and sheep (Fig. 1c). The Phantom Chasing display was identical to the Changing Intentions display, except that the wolf chased the mirror image of the sheep (i.e., the reflection of the sheep's position across the center of the display) rather than the sheep itself. However, since this target of the wolf's pursuit was invisible, the wolf's motion was no longer perceived as goal-directed (Gao et al., 2009, their Experiment 2), even though it was moving in the same intrinsic manner. Thus, if the pSTS represents perceived intentions rather than correlated motion and/or abrupt changes of motion direction, then activation should be lower during the Phantom Chasing display than in the Changing Intentions display.

Finally, and like so many other studies in this domain, tests for social properties such as animacy and intentionality could also be confounded if visual attention was differentially attracted to changing visual events. We controlled for this possibility by including a Flashing control display (Fig. 1d). The Flashing display was identical to the Phantom Chasing display, except that at any moment each disc had a 10% chance of disappearing for 83.3 ms and then reappearing. Subsequent disappearances of the same disc were separated by at least 833 ms. Thus, each disc appeared to flash on and off unpredictably several times during the 10 s display. Sudden onsets and offsets are among the most powerful cues for attentional capture (Yantis and Jonides, 1984).

To verify that observers perceived the Changing Intentions trials (but not Phantom Chasing trials) in terms of shifting intentions, we collected both free descriptions of the displays and ratings of both animacy and intentionality. When asked informally to describe Changing Intentions displays, naive observers frequently invoked notions of chasing and shifting intentions—e.g., “. . . the red ball goes back and forth chasing the green ball and then switches and chases the blue ball” and “. . . the red ball tries to tag the green and when the blue ball was close, it would chase the blue and keeps going back and forth.” In contrast, no independent observer ever described the Phantom Chasing display in such terms—but instead described the underlying physical kinematics as in “There were three dots hitting walls and bouncing off of them” and “A series of ping pong balls, a group of balls, bouncing off walls.” We also ran a short study wherein 14 naive observers (per condition) simply viewed a single Changing Intentions or Phantom Chasing trial (between subjects), and then immediately used a seven-point scale to answer two questions: (1) “To what degree did the discs move as if they were ‘alive’? (1 = not at all alive, 7 = seemed very alive)” and (2) “To what degree did it look as if the red disc's motion was goal-directed and intentional? (1 = not at all, 7 = definitely).” The ratings did not differ for the first question (5.14 vs 4.36; $t_{(26)} = 1.34, p = 0.193$), but ratings for the second question were reliably higher for Changing Intentions trials compared with Phantom Chasing trials (5.86 vs 3.21; $t_{(26)} = 4.36, p < 0.001$).

Image acquisition. Brain images were acquired at the Magnetic Resonance Research Center at Yale University using a 3.0 tesla TIM Trio Siemens scanner with a 12-channel head coil. Functional images were acquired using an echo planar pulse sequence (TR = 2 s, TE = 25 ms, flip angle $\alpha = 90^\circ$, matrix = 64^2 , in-plane resolution = 3.75 mm^2 , slice thickness = 3.5 mm, 37 slices). High-resolution T1-weighted structural images were acquired using a 3D MPRAGE sequence (matrix = 256^2 , in-plane resolution = 1 mm^2 , slice thickness = 1 mm, 176 slices) and a second series of structural images were acquired that was coplanar with the functional images (matrix = 256^2 , in-plane resolution = 1 mm^2 , slice thickness = 3.5 mm, 37 slices).

fMRI data analysis. Data analysis for both experiments used the FMRIB Software Library (FSL) (Smith et al., 2004). All images were skull-stripped using FSL's brain extraction tool, supplemented when necessary by manual masking. The first three volumes (6 s) of each functional dataset were removed to diminish MR equilibration effects. Data were temporally realigned to correct for interleaved slice acquisition, and spatially realigned to correct for head motion using FSL's MCFLIRT linear realignment tool. Images were spatially smoothed with a 5 mm FWHM isotropic Gaussian kernel. Each time series was high-pass filtered (0.01 Hz cutoff) to remove low-frequency drift. Functional images were registered to structural coplanar images, which were in turn registered to high-resolution anatom-

Table 1. Summary of clusters identified by our GLM analysis and the intersection of contrasts depicted in Figure 3

Contrast	Volume (cm^3)	Peak z	x, y, z
a) Changing Intentions > Single Intention			
Right pSTS	12.0	3.5	56, -38, 12
b) Changing Intentions > Phantom Chasing			
Right pSTS	9.7	3.6	58, -38, 24
Right occipital pole	23.8	5.0	22, -98, -2
Left occipital pole	10.2	4.8	-20, -98, -4
Left superior parietal lobule	3.6	3.3	-42, -50, 54
c) Changing Intentions > Flashing			
Right pSTS	9.7	3.7	58, -36, 30
Bilateral occipital pole	43.6	5.2	-26, -92, -8
Left supramarginal gyrus	18.1	3.9	-46, -48, 50
Left precentral gyrus	6.8	3.5	-26, -8, 48
d) Flashing > Phantom Chasing			
None			
Intersection of contrasts a, b, c			
Right pSTS	2.7	3.4	66, -42, 26
Right STG	0.1	2.9	50, -38, 10

An anatomical label for each cluster is provided that corresponds to its approximate center of mass. The volume of each cluster, the peak z statistic within the cluster, and the MNI coordinates of that peak z statistic are also provided. For the intersection analysis, the peak z statistic represents the mean of the z-statistics for the contrasts represented in Figure 2, a–c.

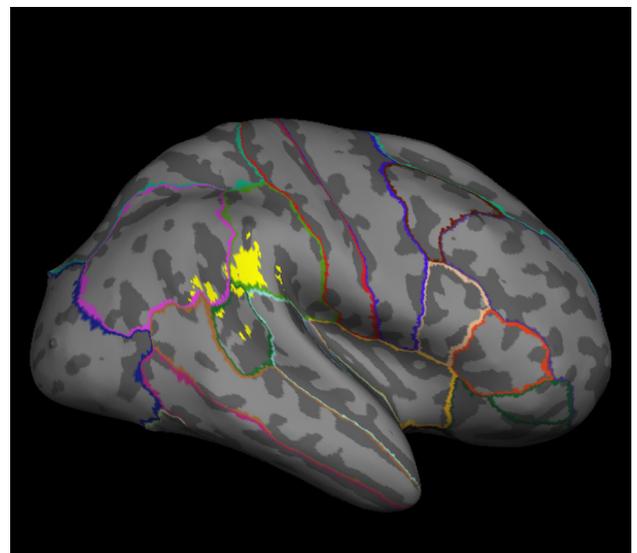


Figure 3. The voxels depicted in yellow represent the intersection of voxels identified as activated in the contrasts depicted in Figure 2a–c.

ical images, which were then normalized to the Montreal Neurological Institute's MNI152 template.

Whole brain voxelwise regression analyses were performed using FSL's fMRI expert analysis tool (FEAT). Each condition within each preprocessed run was modeled with a boxcar function convolved with a single-gamma hemodynamic response function. Regressors were constructed for the four display conditions (Changing Intentions, Single Intention, Phantom Chasing, and Flashing). In each task, the first-level analyses for each run for each subject were combined into a second-level analysis for each subject using a fixed-effects model. The individual subject analyses were then combined into a third-level group analysis. For group-level analyses, parameter estimates were assessed with a mixed-effects model, with the random-effects component of variance estimated using FSL's FLAME stage 1 + 2 procedure. For the third-level analyses, the activation maps were thresholded using FSL's two-stage cluster-correction procedure. Voxels with $z \geq 2.3$ were retained in the first stage and the resulting clusters were then evaluated at a corrected $p < 0.05$ using Gaussian random-field theory. All group-level analyses are presented here on a

flattened, inflated representation of the cortical surface, derived by FreeSurfer (Fischl et al., 2002) from the MNI152 brain.

Results

As per our hypotheses, we focused our analysis upon the right pSTS and adjacent lateral occipitotemporal cortex (LOTc). The contrasts ($p < 0.05$, corrected for multiple comparisons, as above) are shown in Figure 2. Table 1 presents a summary of all significant clusters identified in our analyses. In the test of our main hypothesis, the Changing Intentions display gave rise to stronger activation in the right pSTS and adjacent right LOTc than did the Single Intention display (Fig. 2*a*). In our control contrasts, the Changing Intentions display also gave rise to stronger activation in the right pSTS than did either the Phantom Chasing display (Fig. 2*b*) or the Flashing display (Fig. 2*c*). A comparison of the Flashing and Phantom Chasing control displays revealed no significant differences (i.e., no selectively activated voxels) in any brain region (Fig. 2*d*). For the main contrasts depicted in Figure 2, *a–c*, the pSTS activation occurred exclusively in the right hemisphere; i.e., no clusters were identified in the left pSTS region that survived correction for multiple comparisons. However, activations of smaller spatial extent were observed in the left pSTS when we examined *z*-statistics uncorrected for multiple comparisons. Figure 3 presents the intersection of the activations for the three main contrasts of interest depicted in Figure 2, *a–c*. The common activation occurred in the right pSTS, including part of the angular gyrus and extending into the posterior supramarginal gyrus.

Discussion

Chasing is a powerful cue to the perception of animacy (Gao et al., 2009), and was exploited here to hold animacy constant in the Changing Intentions and Single Intention displays: in both cases, the wolf was always chasing a target according to the same rules. The particular target being chased, however, varied frequently in the Changing Intentions displays but not in the Single Intention displays. The resulting fMRI analysis showed robust selective right pSTS engagement by the Changing Intentions displays.

Moreover, the additional control conditions confirmed that this activation was due to the perception of shifting intentions *per se*, and not to other lower-level factors. In particular, this activation could not have been due to any differences related to motion energy (e.g., when the wolf more frequently shifted directions suddenly in the Changing Intentions displays). Similar selective activity was observed when contrasting the Changing Intentions displays with the Phantom Chasing displays, in which all such properties were preserved but the perception of animacy and intentionality was eliminated (Gao et al., 2009). This contrast could not reflect any differences in prediction error (Hampton et al., 2008), since the timing of the trajectory changes in both conditions was always constant (occurring every 1.2 s) and since the motion of the wolf was carefully constructed to be equally (un)predictable—always being an equally deterministic function of a randomly moving sheep item. Similarly, this activation could not have been due to differences in attentional capture, since similar selective activity was observed when contrasting the Changing Intentions displays with the Flashing displays, in which sudden onsets and offsets were used to capture attention in a maximally powerful fashion (Yantis and Jonides, 1984).

On the basis of these results, we draw two primary conclusions. First, but less importantly, we conclude that the right pSTS is engaged in social perception, as distinct from other nonsocial factors. This same conclusion has been drawn from many other studies, but it has been challenging to support such claims given

that the relevant social versus nonsocial conditions have always differed in so many other correlated ways. Here, we submit that this contrast in the present study is especially compelling, as our use of the chasing displays allowed us to rule out these competing factors, varying social information while holding nearly every lower-level property constant.

Our second and more important conclusion is that the right pSTS is selectively engaged by the perceptual analysis of shifting intentions, beyond animacy. This novel conclusion was in no way mandated by previous studies. Cues for detecting animacy include the apparent violation of Newtonian laws, self-propulsion, and abrupt changes in motion direction and speed (Tremoulet and Feldman, 2000). These cues share a common property, in that physical forces in the environment cannot fully explain an object's motion. Its motion must therefore be attributed to forces that are internal to the object, yielding the perception of animacy. If the engagement of the pSTS in social perception was limited to the detection of animacy in this way, though, it might still fail to implement any routines that are social in nature. Instead, the pSTS would only need to represent a physical model of objects' movements, such that any motion that violated this physical model would be categorized as animate.

In contrast to the detection of animacy, perceived intentionality is governed by a rationality principle, which is a social principle above and beyond low-level motion properties. According to this principle, agents will tend to choose actions that achieve their desires most efficiently, given their beliefs about the world. This property was initially discussed in the context of philosophical considerations (Dennett, 1987) and cognitive development (Gergely et al., 1995; Csibra, 2008), but it has since been incorporated into studies of perceived intentionality (Baker et al., 2009; Gao and Scholl, 2011). This principle also applies to the wavering wolf display that was used in our Changing Intentions trials. In previous studies of chasing, the rationality principle was violated by having the wolf move randomly during certain intervals—a manipulation that simply disrupts the ability to detect chasing in the first place (Gao and Scholl, 2011). Here, in contrast, the visual system faces a dilemma. On one hand, cues to chasing are always equally present and robust; but on the other hand, these cannot be unified, since the rationality principle is constantly being violated for any particular target. We suggest that the visual system reconciles this dilemma by generating percepts of shifting intentions, and that the right pSTS is selectively engaged by this process. In this way, the right pSTS may play a critical role in the neural realization of the rationality principle in online visual perception.

References

- Baker CL, Saxe R, Tenenbaum JB (2009) Action understanding as inverse planning. *Cognition* 113:329–349.
- Blakemore SJ, Boyer P, Pachot-Clouard M, Meltzoff A, Segebarth C, Decety J (2003) The detection of contingency and animacy from simple animations in the human brain. *Cereb Cortex* 13:837–844.
- Bonda E, Petrides M, Ostry D, Evans A (1996) Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *J Neurosci* 16:3737–3744.
- Castelli F, Happé F, Frith U, Frith C (2000) Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage* 12:314–325.
- Csibra G (2008) Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition* 107:705–717.
- Dennett DC (1987) *The intentional stance*. Cambridge, MA: MIT.
- Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM (2002) Whole brain segmentation: automated label-

- ing of neuroanatomical structures in the human brain. *Neuron* 33:341–355.
- Gao T, Scholl BJ (2011) Chasing vs. stalking: interrupting the perception of animacy. *J Exp Psychol Hum Percept Perform* 37:669–684.
- Gao T, Newman GE, Scholl BJ (2009) The psychophysics of chasing: A case study in the perception of animacy. *Cogn Psychol* 59:154–179.
- Gergely G, Nádasdy Z, Csibra G, Bíró S (1995) Taking the intentional stance at 12 months of age. *Cognition* 56:165–193.
- Grossman E, Donnelly M, Price R, Pickens D, Morgan V, Neighbor G, Blake R (2000) Brain areas involved in perception of biological motion. *J Cogn Neurosci* 12:711–720.
- Hampton AN, Bossaerts P, O’Doherty JP (2008) Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc Natl Acad Sci U S A* 105:6741–6746.
- Heberlein A (2008) Animacy and intention in the brain: neuroscience of social event perception. In: *Understanding events: from perception to action* (Shipley T, Zacks J, eds.), pp. 363–388. New York: Oxford UP.
- Heider F, Simmel M (1944) An experimental study of apparent behavior. *Am J Psychol* 57:243–259.
- Martin A, Weisberg J (2003) Neural foundations for understanding social and mechanical concepts. *Cogn Neuropsychol* 20:575–587.
- Pelphrey KA, Mitchell TV, McKeown MJ, Goldstein J, Allison T, McCarthy G (2003) Brain activity evoked by the perception of human walking: controlling for meaningful coherent motion. *J Neurosci* 23:6819–6825.
- Scholl BJ, Tremoulet PD (2000) Perceptual causality and animacy. *Trends Cogn Sci* 4:299–309.
- Schultz J, Imamizu H, Kawato M, Frith CD (2004) Activation of the human superior temporal gyrus during observation of goal attribution by intentional objects. *J Cogn Neurosci* 16:1695–1705.
- Schultz J, Friston KJ, O’Doherty J, Wolpert DM, Frith CD (2005) Activation in posterior superior temporal sulcus parallels parameter inducing the percept of animacy. *Neuron* 45:625–635.
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM (2004) Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23 [Suppl 1]:S208–S219.
- Tremoulet PD, Feldman J (2000) Perception of animacy from the motion of a single object. *Perception* 29:943–951.
- Yantis S, Jonides J (1984) Abrupt visual onsets and selective attention: evidence from visual search. *J Exp Psychol Hum Percept Perform* 10:601–621.