



## The Neuroscience of Understanding the Emotions of Others

Robert P. Spunt and Ralph Adolphs

California Institute of Technology

### Abstract

We cannot help but impute emotions to the behaviors of others, and constantly infer not only *what* others are feeling, but also *why* they feel that way. The comprehension of other people's emotional states is computationally complex and difficult, requiring the flexible, context-sensitive deployment of cognitive operations that encompass rapid orienting to, and recognition of, emotionally salient cues; classification of emotions into culturally-learned categories; and using an abstract theory of mind to reason about what caused the emotion, what future actions the person might be planning, and what we should do next in response. This review summarizes what neuroscience data - primarily functional neuroimaging data - has so far taught us about the cognitive architecture enabling emotion understanding in its various forms.

### What is Emotion Understanding?

The essence of emotion understanding is the organization of information around discrete emotion categories. We regularly infer emotions from seeing somebody's facial expression, from hearing an animal's cries, from observing the way a person behaves in a crowd, or even just from reading a situation that a character is facing in a novel. These are all very different types of information, yet all can be made sense of by being classified into an emotion category. We refer to information of this kind as *emotion-relevant*. Typically, we have multiple such types of information available at the same time. To assign a category, such information must somehow be aggregated, with conflicts resolved, to arrive at an emotion category that is most consistent with the sensory evidence available.

We distinguish in our review processes for detecting and orienting attention to emotion-relevant sensory cues in the environment (*detection*); processes for classifying information in emotional terms (*categorization*); and processes for attributing the categorized emotion to a cause (*causal attribution*), often through combination with other forms of social cognition, such as the representation of place-specific social norms and reasoning about mental-states (Premack & Woodruff, 1978; Gallagher & Frith, 2003; Saxe, Carey, & Kanwisher, 2004). As displayed in Figure 1, these three kinds of process can be conceived as yielding emotion understanding at different levels of abstraction, from a relatively shallow representation of

---

**Contact:** Ralph Adolphs, radolphs@caltech.edu; Robert Spunt, spunt@caltech.edu; California Institute of Technology, 1200 E California Blvd, HSS 228-77, Pasadena, CA 91125.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

concrete nonverbal cues (e.g., “eyebrows raised toward midline”), to a relatively higher-level representation that combines these cues to infer the emotion category describing *what* the target is feeling (e.g., “fear”); and finally, to an even deeper appreciation of *why* the target feels that way (e.g., “delivering a speech in an hour”) (Vallacher & Wegner, 1987). The rather heterogeneous literature germane to emotion understanding, and the correspondingly heterogeneous and unreliable neuroanatomy it has identified, can profit by acknowledging this representational diversity in how humans come to understand stimuli with emotional content. This scheme also makes specific predictions about how this might differ in animals, who may share with us the ability for low-level representations, but lack some of the high-level inferences.

The serialized scheme in Figure 1 is of course in reality considerably more complicated in at least two ways. For one, all three of the stages depicted in fact typically occur in parallel and overlap in time. Indeed, emotion categorization and attribution can occur extremely rapidly: electrophysiological responses in prefrontal cortex (Kawasaki et al., 2001), and in the amygdala (Oya, Kawasaki, Howard, & Adolphs, 2002), that differentiate emotion categories inferred from visual stimuli have been reported with latencies around 100ms, as fast or faster than latencies required for object recognition. Relatedly, there is evidence for multisensory integration at even some of the earliest processing stages—for instance, subcortical structures like the superior colliculus and thalamus already show neuronal responses that integrate visual and auditory information (Ghose, Maier, Nidiffer, & Wallace, 2014; Miller, Pluta, Stein, & Rowland, 2015), and integration of emotional information can again be seen as early as 100ms with ERPs (Pourtois, de Gelder, Vroomen, Rossion, & Crommelinck, 2000). This raises the second complication: all arrows in Figure 1 should be bidirectional, and there is good evidence that even high-level attributional representations (e.g., in prefrontal cortex) can feed back down to regions involved in detection and early perception (e.g., in visual cortices). This feedback is in line with computational architectures that emphasize predictive coding and Bayesian inference in perceptual cognition. This is a large subfield of computational neuroscience, in which several different models are being explored for their explanatory power in describing how the brain minimizes sensory errors given prior expectations – and one strategy may be to generate rough prior expectations very rapidly, and then refine them once sensory processing is complete.

Although we focus on emotion understanding from observing other people, such as seeing their facial expressions, as we noted above we regularly infer emotions also from mere descriptions of the situations in which people might find themselves. Several of the studies we mention below (e.g., Skerry & Saxe, 2015) in fact use such stimuli—lexical descriptions of scenarios from which we would infer an emotion. For example, imagine being told that a child badly wanted a puppy, and one day after school there is a new puppy at home that the child meets for the first time. There is no mention of any observable behavior on the part of child at all, yet this situation is as prototypical as a facial expression and we have no difficulty inferring “happiness” as the emotion. In this example, we do not engage all of the levels of processing shown in Figure 1: we are able to achieve emotion understanding even in the absence of detection of observable sensory cues. Nonetheless, information about the emotion that can be inferred from this description is represented in the same network of brain structures as for inferring emotions from visual stimuli that show people’s behavior

(Skerry & Saxe, 2014). Given the diversity of naturalistic settings in which humans interact and communicate their emotional states to one another, the actual cognitive processes involved in any given instance of emotion understanding will thus vary as a function of the stimulus.

### **Is there emotion? Detecting emotion-relevant cues**

People and animals have evolved mechanisms for rapidly detecting the presence of emotional information in the environment. While many of these cues signal information most relevant for the observer (such as potential threats), there is also evidence for detecting a broader array of sensory cues that can serve as building blocks for constructing an emotion understanding of other people. For example, a recent study used multivoxel pattern analysis (MVPA), a multivariate classification approach that distinguishes stimulus categories by patterns of activation measured across many voxels, and found that the right posterior superior temporal sulcus (pSTS) contained reliable information about the presence of different facial actions (such as raising the inner corners of the eyebrows) that provide the basic action units that, when combined, form distinct emotional expressions (Srinivasan, Golomb, & Martinez, 2016).

Similarly, there is a line of work encompassing lesions (Adolphs et al., 2005), single-unit recordings (Rutishauser et al., 2013), and fMRI (Gamer & Büchel, 2009) suggesting that the amygdala may be particularly sensitive to wide eyes in faces that may signal fear. The eye region of the face tends to draw the observer's visual attention; this attentional bias correlates with amygdala activity in healthy adults and is reduced in patients with amygdala lesions, who also show a corresponding deficit in recognizing fear in facial expressions (Adolphs et al., 2005). There are analogs in other sensory modalities; for instance, screams and cries possess a specific acoustic ingredient—power modulations in the 30–150Hz range—that imbues them with the emotional quality of “roughness” and also activates the amygdala (Arnal, Flinker, Kleinschmidt, Giraud, & Poeppel, 2015). There is ongoing debate about the extent to which such cue detection, and in particular the amygdala's role in it, is automatic and non-conscious (e.g., Pessoa & Adolphs, 2010).

### **What are they feeling? Identifying specific emotions**

Specific cues can serve to direct our attention to information that is salient for emotion understanding, yet by themselves they typically underconstrain the interpretations that are possible. Somebody might also be screaming or wide-eyed after having won a big prize, rather than in anticipation of being eaten by a tiger. Whereas emotion-relevant cues may thus map onto coarser dimensions such as valence or arousal, classifying emotions into specific categories of individual emotions requires more inference and generally involves cortical sensory regions in addition to subcortical structures. The cortical regions involved are generally those that represent that sensory modality and its specialized processing aspects. For instance, categorizing emotion from facial expressions involves cortical regions in the temporal lobe known to represent faces or biological motion; categorizing emotion from voices involves primary and higher-order auditory cortices. MVPA has shown that there is information about emotion categories in the activation pattern seen in face-selective areas of

fusiform gyrus (Skerry & Saxe, 2015), the medial prefrontal cortex (Peelen, Atkinson, & Vuilleumier, 2010) (to which we return below), and the superior temporal sulcus (Wegrzyn et al., 2015) (involved in face and biological motion processing). Yet it is important to note that strong evidence for fine-grained representation of different emotions is still lacking, in good part because most studies only contrast one or two emotions, but not a large number of emotions.

Related to difficulties in investigating representations for specific emotions are two further important points. First, one might think that the sparse sampling of different emotions in any single neuroimaging study could be overcome simply by meta-analyses that look across a large number of studies. This has in fact been done in a number of meta-analyses, which conclude either that there is no specificity in how the brain represents individual emotions at all (Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012), that such representation is overlapping and distributed (Wager et al., 2015), or that it is lower-dimensional than for specific emotions (e.g., just for valence, as supported also by some individual research studies; Chikazoe, Lee, Kriegeskorte, & Anderson, 2014). One problem with all these is that the studies going into the meta-analyses are of mixed quality and highly variable in stimuli and tasks, making it difficult to extract specific conclusions. Another problem is that the resolution of fMRI, especially univariate fMRI, may well be too low to resolve representations of distinct emotions. Indeed, as we saw above, the best emerging evidence for specific emotion representations comes from MVPA, not from univariate analyses.

But a second important problem for investigating the neural basis of recognizing specific emotions is to decide what those emotions should be in the first place. There is a commonly used set of about six “basic” emotions (happiness, surprise, fear, anger, disgust and sadness), which are the set that psychologists like Paul Ekman have claimed are recognized universally across cultures (Ekman, 1994). But there is also good emerging evidence that in fact emotions are not categorized similarly across all cultures at all (from either face or voice; Gendron, Roberson, van der Vyver, & Barrett, 2014a; Gendron, Roberson, van der Vyver, & Barrett, 2014b) making it unclear what should be the “ground truth” for how to divide up the categories. Even if there were consensus on the categories to include, issues remain with the common practice of using prototypic, exaggerated facial expressions to represent them. This is because people rarely encounter the prototypic form of some emotions, such as fear (Somerville & Whalen, 2006). In addition to introducing problematic confounds in prior experience between emotion categories, such a lack of realism in the stimuli used in these studies imposes strong limits on the ecological validity of the psychological processes being elicited (Zaki & Ochsner, 2009).

Following a historical proposal by William James (James, 1884), several modern theories of the conscious experience of emotion (feelings) hypothesize that we feel emotions in virtue of our brain representing the body states associated with the emotion (Damasio, 2003; Craig, 2008). In light of studies showing that observers often mimic (i.e., mirror) the emotional behavior of others, James’ mechanism suggests that observers may understand *what* others are feeling because they can vicariously experience the same feeling in themselves i.e., that we *simulate* the body state of another person by representing that state in somatosensory-related regions (S1, S2, insula) in our own brain (Goldman & Sripada, 2005). In support of

this hypothesis, lesions of somatosensory-related cortices (particularly in the right hemisphere) disrupt people's ability to recognize facial emotion (Adolphs, Damasio, Tranel, Cooper, & Damasio, 2000). Most interestingly, fMRI studies have found that MVPA of regions within right somatosensory cortex represented information about specific facial expressions, and that this representation showed some somatotopy such that specific emotions could be decoded from those regions of S1 that represent the facial features most diagnostic for that emotion (such as wide eyes for fear; Kragel & LaBar, 2015).

## Why are they feeling it? Causal attribution of emotion

Imagine witnessing your close colleague pound their fists while seated at their desk, mouth agape as they glare at their computer monitor. To fully understand this emotional situation, it would be insufficient to merely recognize that your colleague is angry. Instead, observers need to generate a coherent account, an explanation, of what it is that they observe: they need also to understand *why* the other person feels the way they do, and this is accomplished by making causal attributions. This final, richest, aspect of emotion understanding thus involves two components that each map onto somewhat distinct neural regions: (1) an understanding of the emotion in terms of semantic knowledge about the emotion (often encoded in language); (2) an understanding of the emotion in terms of its causes. The first involves brain regions that store semantic conceptual knowledge; and the second brain regions broadly involved in Theory of Mind (Figure 2).

For component (1), regions within the anterior temporal lobe (ATL) are believed to be critically involved in maintaining conceptual knowledge about the social world. Such knowledge includes both specific facts about familiar others (e.g., their first name, favorite foods) (Sugiura et al., 2006; Imaizumi et al., 1997), as well as more general abstract beliefs about the world, such as place-specific behavioral norms and mental-state and trait concepts (Zahn et al., 2007; Lambon Ralph, Pobric, & Jefferies, 2009).

For component (2), the attribution of identified emotions to abstract causes is likely supported by regions in the dorsomedial and ventrolateral prefrontal cortices. These regions are thought to implement executive functions for retrieving relevant abstract knowledge and selecting among competing interpretations of complex stimuli (Satpute, Badre, & Ochsner, 2013; Fairhall & Caramazza, 2013; Goldberg, Perfetti, Fiez, & Schneider, 2007; Green, Fugelsang, Kraemer, Shamosh, & Dunbar, 2006), functions that, in the domain of emotional and social inference, typically take the form of mental-state inferences (Spunt, Ellsworth, & Adolphs, 2016; Spunt, Kemmerer, & Adolphs, 2016; Skerry & Saxe, 2015; Kim et al., 2015; Skerry & Saxe, 2014). Indeed, meta-analyses of neuroimaging studies of theory-of-mind reasoning in its various forms reliably implicate these same regions (Van Overwalle & Baetens, 2009; Amodio & Frith, 2006; Gallagher & Frith, 2003), and the specific association with tasks used to assess theory-of-mind is particularly reliable for the dmPFC (Schurz, Radua, Aichhorn, Richlan, & Perner, 2014).

## Outstanding Issues

By way of conclusion, we briefly identify several questions and issues that would be important to tackle in future neuroscience research on emotion understanding. Two general issues that face current studies are about details: to provide a more detailed account of the different processes schematized in Figure 1 and to understand how they interact; and to provide a more detailed, and possibly different, set of emotion categories. Addressing both of these issues will require careful attention to task design, and will likely require obtaining behavioral performance measures as well as experimental designs informed by computational models. It will also require a close consideration of individual differences (e.g., goals, beliefs, prior experience) that might lead different observers to different interpretations of an emotional stimulus (e.g., Spunt & Adolphs, 2015; Spunt, Ellsworth, & Adolphs, 2016).

Another rich domain will be to understand the automatic or controlled nature of emotion understanding (for further discussion, see Spunt and Lieberman, 2014). While we believe that both processing aspects are involved, the assumption has generally been that emotion inference is relatively automatic. Can we control how we see emotions in other people? If so, can we train people to see others differently? These important open questions are also relevant for disorders such as autism, in which emotion inference is atypical.

A final topic for the future would be to bring the investigation of self-attribution of emotions closer to the literature we have reviewed here. Do we engage some of the same neural systems when we attribute an emotion to ourselves as we do when we attribute emotions to others? While self-relevant processing has also prominently highlighted one of the nodes we noted above—the dorsomedial PFC—it remains quite opaque what to put into the bottom two levels of processing in Figure 1, since we do not know what normally starts an emotion understanding process when it is applied to oneself.

## Acknowledgments

Funded in part by NIMH grant 2P50MH094258.

## References

- Adolphs R. Emotion. *Current Biology*. 2010; 20(13):R549–52. DOI: 10.1016/j.cub.2010.05.046 [PubMed: 20619803]
- Adolphs R, Damasio H, Tranel D, Cooper G, Damasio AR. A role for somatosensory cortices in the visual recognition of emotion as revealed by 3-D lesion mapping. *The Journal of Neuroscience*. 2000; 20:2683–2690. [PubMed: 10729349]
- Adolphs R, Damasio H, Tranel D, Cooper G, Damasio AR. A role for somatosensory cortices in the visual recognition of emotion as revealed by 3-D lesion mapping. *The Journal of Neuroscience*. 2000; 20:2683–2690. [PubMed: 10729349]
- Adolphs R, Gosselin F, Buchanan TW, Tranel D, Schyns P, Damasio AR. A mechanism for impaired fear recognition after amygdala damage. *Nature*. 2005; 433(7021):68–72. DOI: 10.1038/nature03086 [PubMed: 15635411]
- Amodio DM, Frith CD. Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews. Neuroscience*. 2006; 7(4):268–277. DOI: 10.1038/nrn1884 [PubMed: 16552413]

- Arnal LH, Flinker A, Kleinschmidt A, Giraud AL, Poeppel D. Human screams occupy a privileged niche in the communication soundscape. *Current Biology*. 2015; 25(15):2051–2056. DOI: 10.1016/j.cub.2015.06.043 [PubMed: 26190070]
- Chikazoe J, Lee DH, Kriegeskorte N, Anderson AK. Population coding of affect across stimuli, modalities and individuals. *Nature Neuroscience*. 2014; 17(8):1114–1122. DOI: 10.1038/nn.3749 [PubMed: 24952643]
- Craig, AD. Interoception and emotion: a neuroanatomical perspective. In: Lewis, M.Jeanette, M., Barrett, L.F., editors. *Handbook of emotions*. 3rd. New York, NY: Guilford Press; 2008. p. 272-292.
- Damasio, A. *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*. Orlando, FL: Harvest; 2003.
- Ekman P. Strong evidence for universals in facial expressions: a reply to Russell's mistaken critique. *Psychol Bull*. 1994; 115(2):268–287. [PubMed: 8165272]
- Fairhall SL, Caramazza A. Brain regions that represent amodal conceptual knowledge. *Journal of Neuroscience*. 2013; 33(25):10552–10558. DOI: 10.1523/JNEUROSCI.0051-13.2013 [PubMed: 23785167]
- Gallagher HL, Frith CD. Functional imaging of 'theory of mind'. *Trends Cogn Sci*. 2003; 7(2):77–83. DOI: 10.1016/S1364-6613(02)00025-6 [PubMed: 12584026]
- Gamer M, Büchel C. Amygdala activation predicts gaze toward fearful eyes. *Journal of Neuroscience*. 2009; 29(28):9123–9126. DOI: 10.1523/JNEUROSCI.1883-09.2009 [PubMed: 19605649]
- Gendron M, Roberson D, van der Vyver JM, Barrett LF. Cultural relativity in perceiving emotion from vocalizations. *Psychol Sci*. 2014a; 25(4):911–920. DOI: 10.1177/0956797613517239 [PubMed: 24501109]
- Gendron M, Roberson D, van der Vyver JM, Barrett LF. Perceptions of emotion from facial expressions are not culturally universal: evidence from a remote culture. *Emotion*. 2014b; 14(2): 251.doi: 10.1037/a0036052 [PubMed: 24708506]
- Ghose D, Maier A, Nidiffer A, Wallace MT. Multisensory response modulation in the superficial layers of the superior colliculus. *Journal of Neuroscience*. 2014; 34(12):4332–4344. DOI: 10.1523/JNEUROSCI.3004-13.2014 [PubMed: 24647954]
- Goldberg RF, Perfetti CA, Fiez JA, Schneider W. Selective retrieval of abstract semantic knowledge in left prefrontal cortex. *Journal of Neuroscience*. 2007; 27(14):3790–3798. DOI: 10.1523/JNEUROSCI.2381-06.2007 [PubMed: 17409243]
- Goldman AI, Sripada CS. Simulationist models of face-based emotion recognition. *Cognition*. 2005; 94(3):193–213. DOI: 10.1016/j.cognition.2004.01.005 [PubMed: 15617671]
- Green AE, Fugelsang JA, Kraemer DJ, Shamosh NA, Dunbar KN. Frontopolar cortex mediates abstract integration in analogy. *Brain Research*. 2006; 1096(1):125–137. DOI: 10.1016/j.brainres.2006.04.024 [PubMed: 16750818]
- Imaizumi S, Mori K, Kiritani S, Kawashima R, Sugiura M, Fukuda H, Nakamura K. Vocal identification of speaker and emotion activates different brain regions. *NeuroReport*. 1997; 8(12): 2809–2812. DOI: 10.1097/00001756-199708180-00031 [PubMed: 9295122]
- James W. What is an emotion? *Mind*. 1884; 9(34):188–205. Retrieved from <http://mydocs.strands.de/MyDocs/07066/07066.pdf>.
- Kawasaki H, Kaufman O, Damasio H, Damasio AR, Granner M, Bakken H, Adolphs R. Single-neuron responses to emotional visual stimuli recorded in human ventral prefrontal cortex. *Nature Neuroscience*. 2001; 4:15–16. DOI: 10.1038/82850 [PubMed: 11135639]
- Kim J, Schultz J, Rohe T, Wallraven C, Lee SW, Bühlhoff HH. Abstract representations of associated emotions in the human brain. *Journal of Neuroscience*. 2015; 35(14):5655–5663. DOI: 10.1523/JNEUROSCI.4059-14.2015 [PubMed: 25855179]
- Kragel PA, LaBar KS. Multivariate neural biomarkers of emotional states are categorically distinct. *Soc Cogn Affect Neurosci*. 2015; 10(11):1437–1448. DOI: 10.1093/scan/nsv032 [PubMed: 25813790]
- Lambon Ralph MA, Pobric G, Jefferies E. Conceptual knowledge is underpinned by the temporal pole bilaterally: convergent evidence from rTMS. *Cereb Cortex*. 2009; 19(4):832–838. DOI: 10.1093/cercor/bhn131 [PubMed: 18678765]

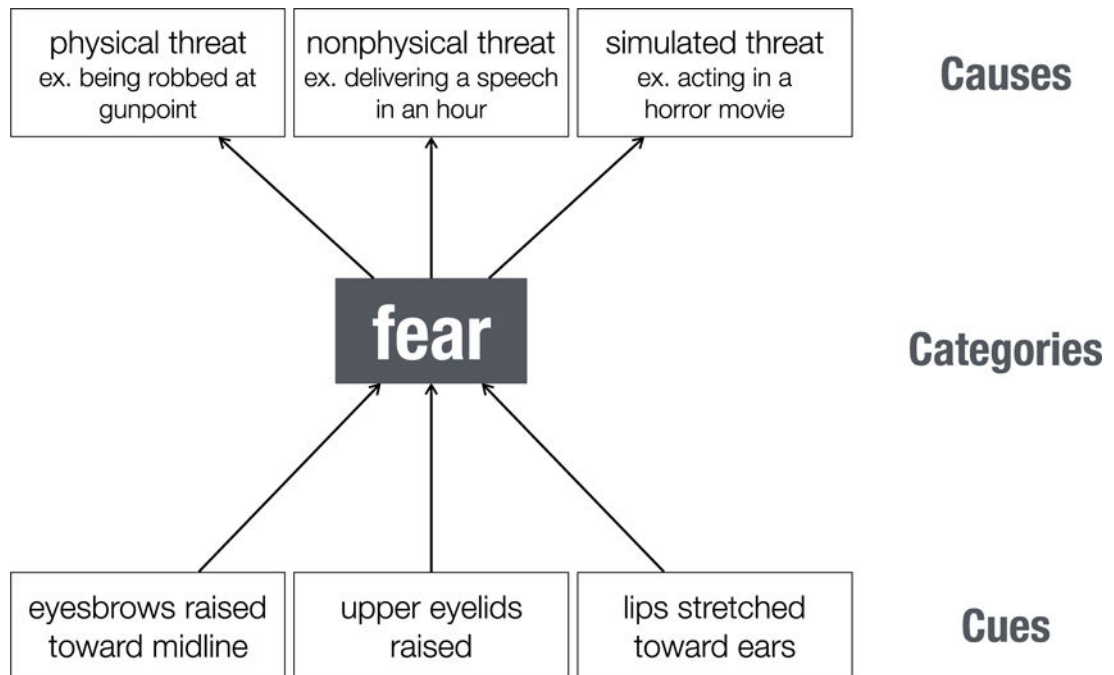
- Lindquist KA, Wager TD, Kober H, Bliss-Moreau E, Barrett LF. The brain basis of emotion: a meta-analytic review. *Behav Brain Sci.* 2012; 35(3):121–143. DOI: 10.1017/S0140525X11000446 [PubMed: 22617651]
- Miller RL, Pluta SR, Stein BE, Rowland BA. Relative unisensory strength and timing predict their multisensory product. *Journal of Neuroscience.* 2015; 35(13):5213–5220. DOI: 10.1523/JNEUROSCI.4771-14.2015 [PubMed: 25834047]
- Oya H, Kawasaki H, Howard MA, Adolphs R. Electrophysiological responses in the human amygdala discriminate emotion categories of complex visual stimuli. *Journal of Neuroscience.* 2002; 22(21):9502–9512. Retrieved from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&opt=Citation&list\\_uids=12417674](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&opt=Citation&list_uids=12417674). [PubMed: 12417674]
- Peelen MV, Atkinson AP, Vuilleumier P. Supramodal representations of perceived emotions in the human brain. *Journal of Neuroscience.* 2010; 30(30):10127–10134. DOI: 10.1523/JNEUROSCI.2161-10.2010 [PubMed: 20668196]
- Pessoa L, Adolphs R. Emotion processing and the amygdala: from a ‘low road’ to ‘many roads’ of evaluating biological significance. *Nature Reviews. Neuroscience.* 2010; 11:773–783. DOI: 10.1038/nrn2920 [PubMed: 20959860]
- Pourtois G, de Gelder B, Vroomen J, Rossion B, Crommelinck M. The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport.* 2000; 11(6):1329–1333. [PubMed: 10817616]
- Premack D, Woodruff G. Does the chimpanzee have a theory of mind. *Behav Brain Sci.* 1978; 1(04):515. doi: 10.1017/S0140525X00076512
- Rutishauser U, Tudusciuc O, Wang S, Mamelak AN, Ross IB, Adolphs R. Single-neuron correlates of atypical face processing in autism. *Neuron.* 2013; 80(4):887–899. DOI: 10.1016/j.neuron.2013.08.029 [PubMed: 24267649]
- Satpute AB, Badre D, Ochsner KN. Distinct regions of prefrontal cortex are associated with the controlled retrieval and selection of social information. *Cereb Cortex.* 2013; doi: 10.1093/cercor/bhs408
- Saxe R, Carey S, Kanwisher N. Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology.* 2004; 55:87–124. DOI: 10.1146/annurev.psych.55.090902.142044
- Schurz M, Radua J, Aichhorn M, Richlan F, Perner J. Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neurosci Biobehav Rev.* 2014; 42C:9–34. DOI: 10.1016/j.neubiorev.2014.01.009
- Skerry AE, Saxe R. A common neural code for perceived and inferred emotion. *Journal of Neuroscience.* 2014; 34(48):15997–16008. DOI: 10.1523/JNEUROSCI.1676-14.2014 [PubMed: 25429141]
- Skerry AE, Saxe R. Neural representations of emotion are organized around abstract event features. *Current Biology.* 2015; 25(15):1945–1954. DOI: 10.1016/j.cub.2015.06.009 [PubMed: 26212878]
- Somerville LH, Whalen PJ. Prior experience as a stimulus category confound: An example using facial expressions of emotion. *Soc Cogn Affect Neurosci.* 2006; 1(3):271–274. DOI: 10.1093/scan/nsl040 [PubMed: 18985113]
- Spunt RP, Adolphs R. Folk explanations of behavior: A specialized use of a domain-general mechanism. *Psychol Sci.* 2015; 26(6):724–736. DOI: 10.1177/0956797615569002 [PubMed: 25911123]
- Spunt RP, Ellsworth E, Adolphs R. The neural basis of understanding the expression of the emotions in man and animals. *Soc Cogn Affect Neurosci.* 2016; doi: 10.1093/scan/nsw161
- Spunt RP, Kemmerer D, Adolphs R. The neural basis of conceptualizing the same action at different levels of abstraction. *Soc Cogn Affect Neurosci.* 2016; 11(7):1141–1151. DOI: 10.1093/scan/nsv084 [PubMed: 26117505]
- Spunt, RP., Lieberman, MD. Automaticity, control, and the social brain. In: Sherman, JW, Gawronski, B., Trope, Y., editors. *Dual-Process Theories of the Social Mind.* 2. New York: Guilford Press; 2014. p. 279-299.



- Srinivasan R, Golomb JD, Martinez AM. A neural basis of facial action recognition in humans. *Journal of Neuroscience*. 2016; 36(16):4434–4442. DOI: 10.1523/JNEUROSCI.1704-15.2016 [PubMed: 27098688]
- Sugiura M, Sassa Y, Watanabe J, Akitsuki Y, Maeda Y, Matsue Y, Kawashima R. Cortical mechanisms of person representation: recognition of famous and personally familiar names. *Neuroimage*. 2006; 31(2):853–860. DOI: 10.1016/j.neuroimage.2006.01.002 [PubMed: 16478667]
- Vallacher RR, Wegner DM. What do people think they're doing? Action identification and human behavior. *Psychol Rev*. 1987; 94(1):3–15. DOI: 10.1037/0033-295X.94.1.3
- Van Overwalle F, Baetens K. Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *NeuroImage*. 2009; 48(3):564–584. DOI: 10.1016/j.neuroimage.2009.06.009 [PubMed: 19524046]
- Wager TD, Kang J, Johnson TD, Nichols TE, Satpute AB, Barrett LF. A Bayesian model of category-specific emotional brain responses. *PLoS Computational Biology*. 2015; 11(4):e1004066.doi: 10.1371/journal.pcbi.1004066 [PubMed: 25853490]
- Wegrzyn M, Riehle M, Labudda K, Woermann F, Baumgartner F, Pollmann S, Bien CG, Kissler J. Investigating the brain basis of facial expression perception using multi-voxel pattern analysis. *Cortex*. 2015; 69:131–140. DOI: 10.1016/j.cortex.2015.05.003 [PubMed: 26046623]
- Zahn R, Moll J, Krueger F, Huey ED, Garrido G, Grafman J. Social concepts are represented in the superior anterior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*. 2007; 104(15):6430–6435. DOI: 10.1073/pnas.0607061104 [PubMed: 17404215]
- Zaki J, Ochsner K. The need for a cognitive neuroscience of naturalistic social cognition. *Ann N Y Acad Sci*. 2009; 1167:16–30. DOI: 10.1111/j.1749-6632.2009.04601.x [PubMed: 19580548]

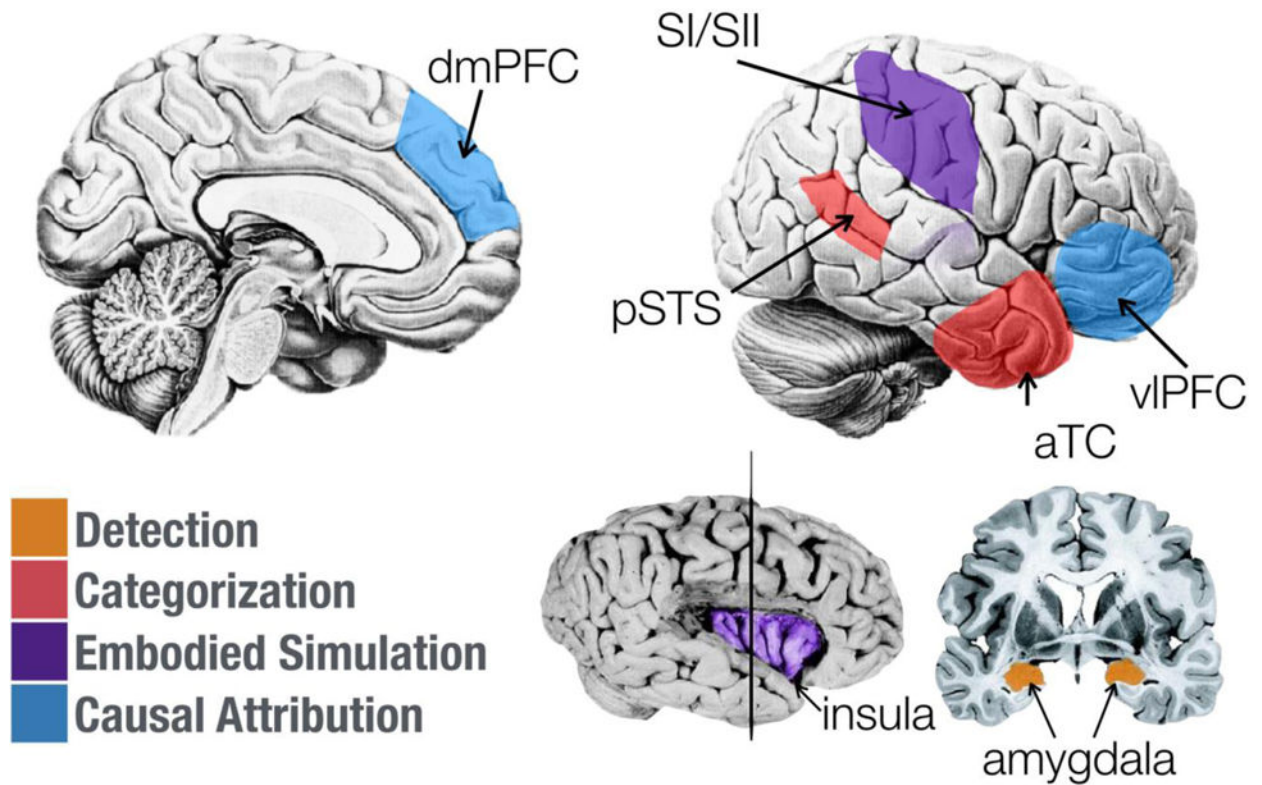
### Highlights

- emotion understanding consists of detecting cues, inferring emotion categories, and attributing causes
- a network of brain structures is involved for each of these
- the dmPFC is the most prominent brain regions involved in inferring emotions and causes



**Figure 1.**

An illustration of the hierarchical and multifaceted nature of understanding other people's behavior in emotional terms. We use fear as an example, but note that the distinctions we make can be applied to any emotion category. To simplify, we demonstrated how a fearful facial expression might be identified at three levels of abstraction. At the lowest level are the different elements of the stimulus (e.g., facial and bodily actions, object of gaze) that in combination can be understood in terms of relatively higher-level behavioral categories, such as fear. Using our capacity for reasoning about the causes of behavior (e.g., using a Theory of Mind), even higher levels of understanding become possible: We understand not just that the person is expressing fear, but why they are expressing it. To organize the literature review shown in the main text, we compartmentalize component processes as contributing primarily to either early detection of emotion-relevant cues; intermediate categorization of the behavior in terms of emotion and action categories; and, finally, further inferences that serve to build a coherent narrative that identifies the cause(s) of the emotion. In actuality, processing at these three levels is typically concurrent and involves substantial feedback, not shown in this figure.



**Figure 2.**

Selected brain structures implicated in the different functional components of emotion understanding discussed in the main text. Each structure is color-coded to indicate, based on existing evidence, the component(s) to which it contributes. To enable a clear visualization, a number of structures known to be involved in emotion understanding are discussed in the main text but not shown in this figure. aTC, anterior temporal cortex; pSTS, posterior superior temporal sulcus; dmPFC, dorsomedial prefrontal cortex; vIPFC, ventrolateral prefrontal cortex. Portions of this figure adapted from Adolphs (2010).