Understanding Affective Cognition: Frontiers in modeling reasoning about others' emotions

Desmond C. Ong (dco@stanford.edu)
Jamil Zaki (jzaki@stanford.edu)
Noah D. Goodman (ngoodman@stanford.edu)
Department of Psychology, Stanford University, Stanford CA, USA

Keywords: Affective cognition; Social cognition; Neuroscience; Computational Modeling; Developmental Approaches

Motivation

Social life constantly requires us to decipher information about others into inferences about their emotional states: for example, we have to reason about what makes our romantic partners happy (a surprise gift?) or angry (not doing one's chores?), and what they would do in those emotional states, in order to plan our upcoming interactions. Such *affective cognition*, or our ability to reason about others' emotions, scaffolds everything from cooperation to the maintenance of social relationships. Affective cognition lies in the intersection of two foundational social cognitive topics, Theory of Mind (ToM; the ability to reason about others' mental states) and Empathy (the ability to feel and understand others' emotions). Although the past decade has seen much progress in understanding ToM and empathy using neuroscience (Koster-Hale & Saxe, 2013; Zaki & Ochsner, 2012), developmental (Meltzoff, 2011), and computational (Baker, Saxe, & Tenenbaum, 2009; Goodman & Stuhlmuller, 2013) approaches, somewhat less attention has been paid specifically to affective cognition and some of its foundational cognitive questions. How do we represent (cognitively and neurally) others' emotional states? How do we reason with those representations? How do we make predictions and inferences about others' future actions or desires based on their emotions? Finally, how does affective cognition shift across development? The aim of the symposium is to answer these, and other relevant questions, at the forefront of this field.

In this symposium, we will discuss recent work in affective cognition done by some of the above-cited researchers and their labs who approach the topic from different disciplines and backgrounds. There will be ample time set aside at the end of the symposium for the speakers and the audience to discuss future research directions. **Amy Skerry** will kick off the symposium by discussing recent work on neural representations of others' emotional states. **Desmond Ong** will then present a computational model of emotion attribution to other agents. **Yang Wu** will continue the discussion with a model of inference over an agent's desires and beliefs from observed emotions, in both adults and children. Our final speaker, **Andrew Meltzoff** will describe and discuss recent developmental work on how children learn and make inferences about others' emotional traits, and regulate their own behavior in response. Finally, open discussion, especially with regards to implications and future directions, will be led and moderated by **Noah Goodman**.

Affective cognition is a key topic that is growing increasingly relevant to the cognitive science community in the past few years, and it is the perfect time to bring together leaders in the field of affective cognition to discuss the exciting state of the art of this field, moving forward. In addition to gaining deeper understanding of the core computations that underlie such reasoning, there are many promising applications of this work. In particular, this symposium and its applications resonate well with the theme of this year's conference, *Cognitive Science Meets Artificial Intelligence: Human and Artificial Agents in Interactive Contexts*. First, having a better understand about basic human affective cognition will enable us to make progress in building interactive artificial agents that can reason about human emotions (Picard, 2000). Second, having models of working and dysfunctional affective cognition would enable us to build computational algorithms and agents to assist with diagnosing and treating clinical disorders (Montague, Dolan, Friston, & Dayan, 2012). Such applications of affective cognition are becoming more ubiquitous each year, and this dialogue between leaders in the cognitive science community will, we hope, spur future research in both the basic and applied disciplines.

A common neural code for perceived and inferred emotions

Amy Skerry (Harvard University) & Rebecca Saxe (Massachusetts Institute of Technology)

Attributing emotional states to others is crucial for predicting and explaining action, identifying intentions and attitudes, and successfully navigating social interactions. While the emotions of other people can often be perceived from overt reactions (e.g. facial or vocal expressions), they can also be inferred from situational information in the absence of observable expressions. How does the human brain make use of these diverse forms of evidence to generate a common representation of a target's emotional state? In the present research, we identify neural patterns that correspond to emotions attributed based on context, and find that these patterns generalize across different cues from which an emotion can be inferred. Specifically, we use fMRI to measure neural responses to dynamic facial expressions with positive and negative valence, and to short animations in which the valence of a character's emotion could be identified only from the situation. Using multi-voxel pattern analysis (MPVA),

we identify brain regions containing stimulus-specific and stimulus-independent information about the emotional state of the target. In the medial prefrontal cortex (MPFC), a classifier trained to discriminate emotional valence for one stimulus type (e.g. animated situations) could successfully discriminate valence for the remaining stimulus type (e.g. facial expressions), indicating a representation of valence that abstracts away from specific perceptual features and generalizes across different forms of evidence. These data provide a step towards understanding how the brain transforms stimulus-bound inputs into abstract representations of others' emotions.

Modeling lay theories of emotion attribution

Desmond C. Ong, Jamil Zaki, & Noah D. Goodman (Stanford University)

Humans are extremely skilled at reasoning about others' emotions; this ability is crucially important to forming and maintaining social relationships. Yet despite its importance, few formal or quantitative theories have successfully described how affective cognition operates. Here we address this gap by constructing a computational model, drawing on tools from Bayesian statistics. We use a simple gambling paradigm to quantitatively test this model, finding that emotion judgments are well explained in terms of a low-dimensional representation based on value-related computations. Further, we demonstrate that emotion inferences across multiple situations are tightly predicted by Bayes' rule. Our results speak to a deep structural relationship between emotions and cognitive inference, and suggest wide-ranging applications to basic psychological theory and psychiatry.

Joint inference of desire and belief from facial expression

Yang Wu, Chris L. Baker, Joshua B. Tenenbaum, & Laura Schulz (Massachusetts Institute of Technology)

Theory of mind research has looked at how learners infer unobservable mental states from observable actions and outcomes. However, such research has tended to neglect another observable source of data: the agent's reactions to actions and outcomes. In particular, the agent's facial expression might provide information about mental states (e.g., the agent's beliefs and desires) that are otherwise ambiguous given the action. Here we present a behavioral study and a computational model looking at how, given identical actions and outcomes, adults use facial expressions to reason backward to agents' underlying beliefs and desires. We find that participants can integrate the likelihood of the observed facial expression and the likelihood of the observed action with their prior over mental states to produce joint inferences of desire and belief, consistent with a Bayesian model analysis. In ongoing work, we are testing the predictions of this account with young children.

Infants' Emotion Attributions – Influences on Learning and Imitation

Andrew N. Meltzoff & Betty Repacholi (University of Washington)

One of the chief tasks of infancy is to understand others' emotional states and traits. Negative emotions, such as anger, are particularly salient to the young child. We report studies in which 15-month-old infants are bystanders and observe a social interchange between two adults. One of the adults becomes angry when the other adult performs an action (as if it was a 'forbidden act'). Infants are then presented with the object, and we measure whether they imitate the forbidden act in view of the previously angry adult (the adult now has a non-angry, neutral facial expression). Results show that infants regulate when to imitate, tapping inhibitory control to govern their own behavior. Interestingly, prior to language, infants have the social-cognitive foundations for making primitive trait-like inferences about angry people: If a person has a history of becoming angry about particular actions, infants treat that person as 'anger-prone' and likely to become angry about novel actions. Infants integrate information about a person's emotional history and their attentional states to make predictions about that person's emotional responses in new situations, and use this to govern their own behavior.

References

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, *113*(3), 329 - 349.

Goodman, N. D., & Stuhlmuller, A. (2013). Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science*, *5*(1).

Koster-Hale, J., & Saxe, R. (2013). Functional neuroimaging of theory of mind. In *Understanding other minds: Perspectives from developmental social neuroscience* (p. 132). Oxford University Press.

Meltzoff, A. N. (2011). Social cognition and the origins of imitation, empathy, and theory of mind. In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development* (Vol. 2, pp. 49–75). Wiley-Blackwell.

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in cognitive sciences*, *16*(1), 72-80.

Picard, R. W. (2000). *Affective computing*. MIT press.

Zaki, J., & Ochsner, K. N. (2012). The neuroscience of empathy: progress, pitfalls and promise. *Nature neuroscience*, *15*(5), 675–680.