Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Psychology

Formalizing emotion concepts within a Bayesian model of theory of mind☆

Rebecca Saxe and Sean Dae Houlihan

Sensitivity to others' emotions is foundational for many aspects of human life, yet computational models do not currently approach the sensitivity and specificity of human emotion knowledge. Perception of isolated physical expressions largely supplies ambiguous, low-dimensional, and noisy information about others' emotional states. By contrast, observers attribute specific granular emotions to another person based on inferences of how she interprets (or 'appraises') external events in relation to her other mental states (goals, beliefs, moral values, costs). These attributions share neural mechanisms with other reasoning about minds. Situating emotion concepts in a formal model of people's intuitive theories about other minds is necessary to effectively capture humans' fine-grained emotion understanding.

Address
Department of Brain and Cognitive Sciences, 43 Vassar St, MIT, Cambridge, MA 02139, USA

Corresponding author: Saxe, Rebecca (saxe@mit.edu)

Current Opinion in Psychology 2017, **17**:15–21

This review comes from a themed issue on **Emotion**

Edited by **Batja Gomes de Mesquita** and **Lisa Feldman Barrett**

For a complete overview see the Issue and the Editorial

Available online 27th April 2017

http://dx.doi.org/10.1016/j.copsyc.2017.04.019

Introduction

"I'd rather write an encyclopedia about common emotions," he admitted. "From A for 'Anxiety about picking up hitchhikers' to E for 'Early risers' smugness' through to Z for 'Zealous toe concealment, or the fear that the sight of your feet might destroy someone's love for you.' "

—*The little Paris bookshop by Nina George.*

If your friend is experiencing early risers' smugness, how would you know? From a quick glance at her face and posture, you see she is experiencing a low-arousal positive emotion. To refine this attribution, though, you would need knowledge of the context and cause of the emotion. She is more likely to feel smug, you know intuitively, if she chose to wake up early (rather than being woken involuntarily by a screaming baby) and if she used those extra hours to her relative advantage (rather than wasting them counting sheep). As this example illustrates, human observers can recognize and reason about highly-differentiated, or fine-grained, emotions. Here we propose that fine-grained emotion concepts are best captured in a Bayesian hierarchical generative model of the intuitive theory of other minds [1•].

The role of concepts in emotion has been much disputed [2–5]. This question is particularly hard for *first person* emotions: when I myself feel anxious, what is the role of my concept of 'anxiety' in the construction of my experience? Here, we selectively tackle an easier problem: the problem of other minds. We recognize anxiety in our friends, distinguish their anxiety from their disappointment or regret, and try to respond in appropriate ways [6]; but how do we make such specific and accurate emotion attributions to another person? In order to formally address that question, we situate emotion concepts in a computational model of the intuitive theory of mind [1•,7]. (Note that intuitive or lay theories are causally structured, but generally not explicit, declarative, or introspectively accessible [8].)

Situating emotion concepts within an intuitive theory of mind

Initial scientific descriptions of an 'intuitive theory of mind' focused on its application to predicting others' intentional actions [9]. Minimally, intentional actions can be predicted (and explained) as consequences of the agent's beliefs and desires, and modeled as inverse planning [10]. Subsequent models have considerably extended this basic premise to capture causal relations between other kinds of mental states. For example, Greg's choices additionally depend on (what he believes about) the costs of his actions [11]; his beliefs update in response to new evidence [7]; his actions are influenced by his habits [12]; and so on. A hierarchical Bayesian model of this intuitive causal theory can explain both observers' forward inferences (predicting Greg's actions

☆ We thank the NSF Center for Minds Brains and Machines (CCF-1231216) and NIH Grant 1R01 MH096914-01A1 (R.S.) for supporting the work; and Josh Tenenbaum, Laura Schulz and Stefano Anzellotti for vigorous discussion.

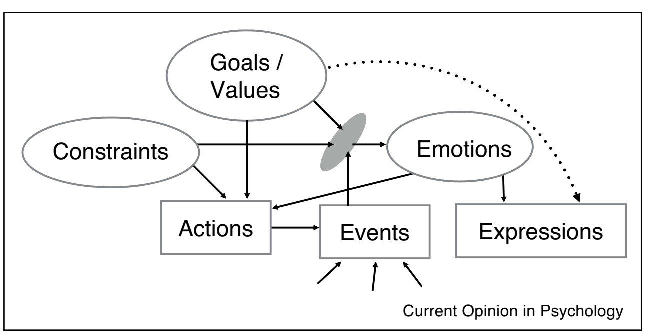
given his beliefs and desires) and inverse inferences (inferring Greg's beliefs and desires given his actions) [10].

People readily incorporate emotions in their intuitive reasoning about other minds [13–15] but only recently have computational models of theory of mind been elaborated to include emotion concepts. Minimally, in the intuitive theory, emotions (or emotional reactions) are caused by how the person interprets (or 'appraises') external events in relation to his other mental states (goals, beliefs, moral values, costs, traits, *etc.*; Figure 1). For example, Greg's emotional reactions will depend on whether (according to Greg) external events fulfill his goals, contradict his beliefs, reduce the constraints or costs of his preferred actions, violate his values, and so on. As with intentional actions, the same intuitive theory also supports inverse inferences. In the intuitive theory, emotions (which are internal mental states) cause emotional expressions (which are externally observable behaviors), so observers can use perceived emotional behaviors to infer underlying emotions (*i.e.*, perform an inverse inference from observed effects to unobserved cause). Situating emotion concepts within the intuitive theory of mind in this way may seem obvious, but has many ramifications, some of which we explore in the remainder of this article.

Specificity and development of emotion inference

First, this approach offers a natural, systematic way to formalize highly-differentiated predictions of others'

Figure 1



A box and arrow simplification of part of the intuitive causal theory of other minds. Ovals denote unobservable, internal states; rectangles are externally observable. Constraints include appraisals of the costs of actions, what is possible, what is controllable, and other beliefs. Goals and values include both local goals and intentions, but also long term values like relationships and status, and therefore can directly influence expressions. At the core of the model is the inferred 'appraisal' process: interpreting external events through the lens of their relevance for one's goals, beliefs, costs, and so on. Inferred appraisals cause emotions (internal states) which cause expressions (observable behaviors). An observer can therefore predict emotions based on inferred appraisals (following the causal arrows) or from the observed expressions (inverse of the causal arrows). Compare similar models in Refs. [14,18••,19].

emotions, and the links between those predictions and the rest of our sophisticated reasoning about other minds. Although no existing model has yet fulfilled this promise, parts of the intuitive theory of mind have already been well-described in Bayesian generative causal models [16•,17]. Capitalizing on this progress, the same formalizations can be used to model (some) human emotion predictions. For example, in a simple lottery context, two parameters of the target's appraisal could be inferred directly from a description of the event—his overall reward, and his prediction error—and combined to capture in quantitative detail the emotions that observers predicted [18••]. Relatedly, Wu *et al.* showed participants simple moral scenarios, in which Grace puts white powder in another girl's coffee [19]. The powder turns out to be poison, and the girl dies. Participants use Grace's smiling facial expression to infer both that Grace knew the powder was poison, and that she wanted the girl to die. These inferences could be precisely described as inverse inferences in the participants' intuitive theory of mind. (In the real world, observers make similarly momentous inverse inferences based on emotional reactions [20,21].)

Even children's earliest understanding of others' emotions implies (simple) inferred appraisals. Based on an agent's observed motion path (and a principle of rational action), preverbal infants can infer the agent's goal (*e.g.*, to get over the wall); then, relative to that goal, infants can distinguish between outcomes that the agent would appraise as goal-consistent or not [22]. Critically, by 10 months old, infants also appear to predict a relevant emotion (or affective state) that causes subsequent expressions (laughing or crying) and are surprised if the agent whose goal was fulfilled then shows a negative-valence behavior, crying [23••] (see also Ref. [24]). During development, children's intuitive theory of mind becomes more sophisticated, and their third-person emotion attributions follow suit [15,25–30]. (Note that while some developmental psychologists reserve the term 'theory of mind' for a meta-representational understanding of beliefs, *e.g.*, [30], here the Bayesian model of theory of mind is a generative causal theory, encompassing goals and actions as well as beliefs, costs, and values [10,16•].)

The long-term goal, however, is not just to capture one or two components of observers' emotion knowledge; rather, it is to develop a formal model that captures all of the same inferred appraisals as human observers do. Promising for this line of work, when given human labels for the target's appraisals, computational models can already capture a relatively wide and differentiated range of human emotion predictions. Two recent studies provide converging evidence. Using human ratings for 25 appraisal features, a model correctly chose an emotion label (out of 14) for 51% of 6000 real-life events; only 10% of the model's choices were judged 'wrong' by human observers [31]. Similarly, using human ratings for 38 inferred appraisals, a simple model correctly chose the emotion label (out of 20) for 57% of 200 short stories; human accuracy on the same test was 63% [32••]. These models do not yet capture the link from the event to the target's values, goals, beliefs, and costs, and thus to inferred appraisals. Still, the models' success suggests that once these links are included, the intuitive theory of mind will capture a substantial portion of shared human knowledge about emotions.

Ambiguous perception and precise predictions

Second, our proposal offers novel insight into predictions based on combinations of inferred appraisals (forward inference) and perceived emotional expressions (inverse inference). People intuit that faces contain the most revealing information about others' emotions [33,34]. Perhaps surprisingly, mounting scientific evidence shows that human emotion attribution from faces is actually uncertain, noisy, and low-dimensional [35–37]. Many different emotions can be attributed to the same facial configuration [38•,39–41]; and the space of emotions perceived in faces can be captured in just a handful of dimensions [42,43•]. Even the valence of the event (goal-congruent or not) is not reliably perceived in high-intensity faces: the exact same facial configuration can be attributed to extreme joy (the unexpected return of a child from military service), extreme distress (witnessing a terrorist attack), extreme pleasure (orgasm), or extreme pain (unanesthetized nipple piercing) with equal plausibility [33]. To disambiguate these emotions, observers rely on body posture (open arms, lifted chest [44]) or inferred appraisals of the event ('he won the race' [45•]).

Although both body posture and event information are known to disambiguate emotion recognition [33,37,45•,46], our model makes a novel distinction between inverse inferences (from bodies) and forward inferences (from event-appraisals). On one hand, observers intuitively infer a common cause (an underlying emotion) of observable face, body and vocal cues. Thus, integrating facial and body configuration, as well as vocal tone, can improve the reliability and specificity of inverse inferences [44,47,48]. Postural information is less ambiguous than facial configuration when perceived at high intensity, from a distance, and so on [44]; similarly, vocal bursts are more informative for distinguishing among positive emotions [49]. As a result, depending on the context, the modality with the most reliable information will appear to dominate emotion attributions [18••,46]; when one cue is ambiguous, cues from other modalities can 'sharpen' the inferred cause by shifting attributions among similar, or nearby, emotions [37]. On the other hand, event information is intuitively relevant to the *cause* of the emotion, rather than its consequences. Additional event information can make emotion attributions more reliable not only by continuously shifting among similar emotions, but also by selecting among separated possibilities [50], because partial event knowledge can generate predictions of distinct (dissimilar, non-overlapping) alternative emotions (*e.g.*, how will he feel after he asks his crush on a first date?). This difference between forward and inverse inferences has been obscured in prior research that confused postural and event-context cues: for example, a photograph of nipple piercing [33] contains mainly event information supporting inferred appraisals, not an emotional posture.

Relatedly, we can distinguish between two ways that 'dynamic' facial expressions contain more information than static ones [51]. On the one hand, dynamic change can more precisely differentiate expressive from structural facial features (*e.g.*, a person with dark brows from a person making an angry expression) [52–55]. Dynamic change can also provide more clarity on mixed expressions, by separating the mixture in time [56]. In these ways, dynamic expressions may lead to more specific or more confident inverse inferences (though observers can also be surprisingly insensitive to dynamic information *per se* [36,57]). On the other hand, when temporal change in the face coincides with temporal change in the external event structure, dynamics support forward inference by highlighting the emotionally-relevant aspect of an event [58]. For example, observers are generally quite insensitive to elements of surprise ('wide-eyed') in mixed expressions [19,42]. When a change of expression is temporally coincident with an event outcome, though, observers accurately infer that the information was unexpected and change their inferred appraisals accordingly [19]. The temporal sequence of emotions can further constrain inferred appraisals; if people intuit that cognitive processes occur at different speeds then the order of expressions can indicate which hidden mental variable is associated with which emotion.

Third, we propose that there is a key asymmetry between forward and inverse inferences of emotion. The forward inference depends on inferred appraisals which are highly differentiated and granular. However, people's intuitive theory of mind is also biased and based on simplifying heuristics, inducing systematic errors [59]. We assume people share our desires, values, norms [60]. We underestimate people's ability to cope, recover, and rebound from significant events [61,62]. These biases in the intuitive theory of mind translate into systematic errors in predictions of emotions. By contrast, inverse inference from emotional expressions is uncertain and low-dimensional, but also relatively accurate and unbiased. Combining both sources is therefore uniquely powerful: forward inferences from inferred appraisals can suggest highly specific, granular, differentiated predictions of another person's emotions; perception of that person's expressions can confirm or contradict these predictions, allowing for rapid correction within a reduced possibility space.

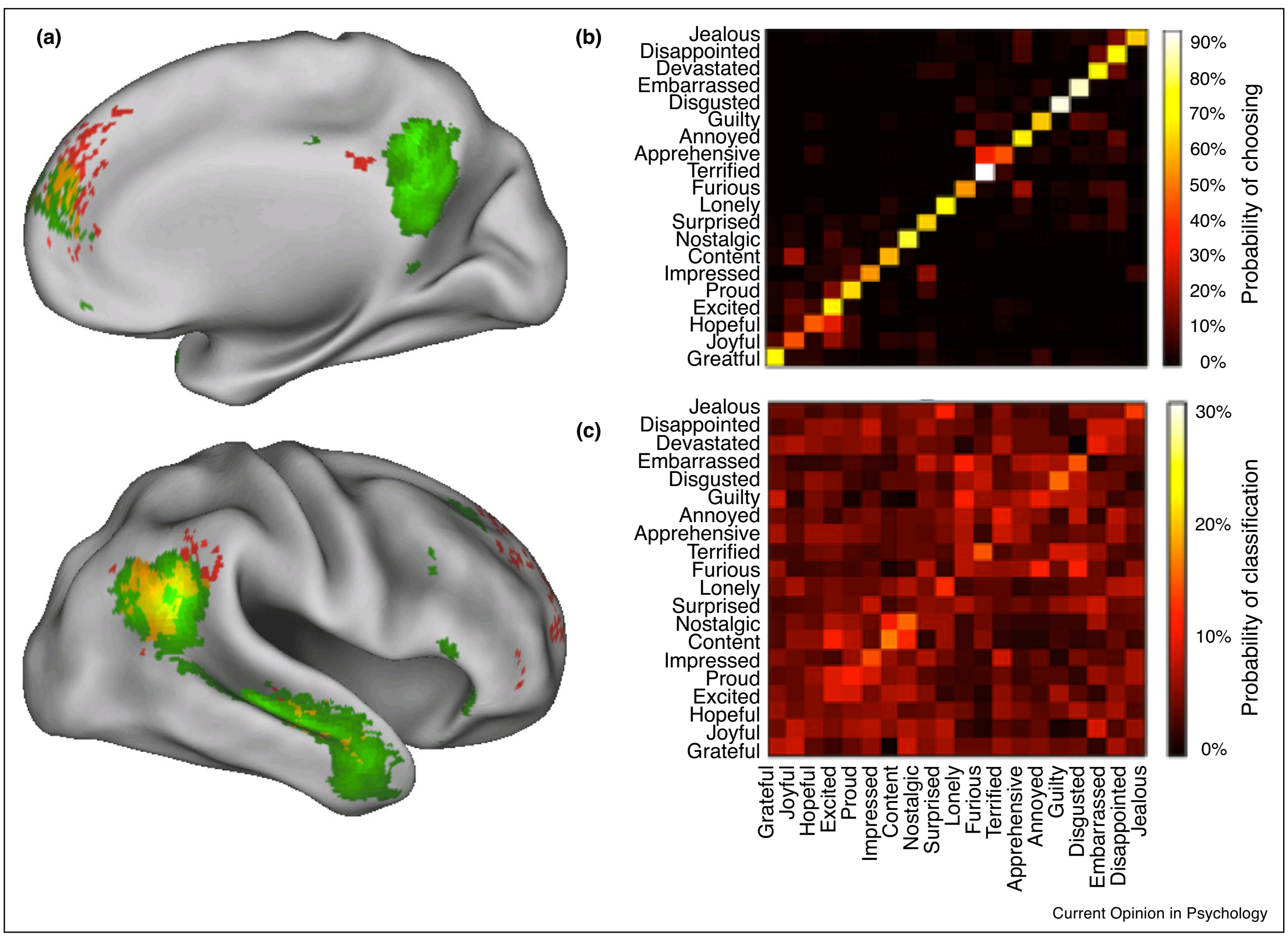
Neural representations of fine-grained emotion concepts

Finally, situating emotion concepts within the intuitive theory of mind fits well with recent neuroscientific evidence. Highly-differentiated representations of others' emotions are almost exclusively found in brain regions associated with theory of mind, especially in temporo-parietal and medial frontal cortex [32••,63,64] (Figure 2). These representations are abstract and amodal, generalizing across emotions inferred from stories, events, and expressions [32••,65]. By contrast, perception of emotional expressions, and even integration of those expressions across modalities, depends on distinct brain regions, especially the superior temporal sulcus [66–70]. These two processes are dissociable in individual differences [71–73], and in neurodegenerative disorders [74]. Taken together, these lines of evidence strongly support the link between emotion concepts and the rest of an observer's intuitive theory of mind.

Conclusion

Two lines of scientific research have made substantial progress in parallel, and now stand to make even more progress in concert. On the one hand, formal computational models have begun to capture the core of people's intuitive theory of mind. These models can accurately model inferences over continuous quantitative variables, within abstract hierarchical structures. As of yet, however, these models have made limited progress in the domain of emotion understanding. On the other hand, the conceptual act theory of emotion attribution identifies the

Figure 2



(a) Fine-grained discrimination of others' emotions (20-way classification) from short verbal narratives (in red) depends on the same brain regions that are involved in intuitive reasoning about other minds (green), especially temporo-parietal junction and medial prefrontal cortex. **(b)** The behavioral confusion matrix for human participants: bright colors show the probability of a human choosing each label, averaged over 10 stories in each of 20 categories. **(c)** The confusion matrix based on patterns of response in brain regions associated with theory of mind. Adapted with permission from Ref. [32••].

powerful influence of emotion concepts on emotion attribution (though emotion concepts are usually operationalized as words, or labels) [75,76]. Appraisal theory describes some of the content of shared knowledge about emotional events (though as a hand-picked and manually-coded list, rather than a generative causal model) [31]. Using the intuitive theory of mind as a framework to formalize observers' inferences about a target's appraisals offers a powerful tool to capture, and even recreate in a computer [77], our detailed knowledge of how others feel.

Conflict of interest statement

Nothing declared.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Lake BM, Ullman TD, Tenenbaum JB, Gershman SJ: **Building** • **machines that learn and think like people**. *Behav. Brain Sci.* 2016:1-101 http://dx.doi.org/10.1017/S0140525X16001837.
Comprehensive argument for using computational generative models to build human-like artificial intelligence to address the current major shortcoming of machine learning. Stresses the use of intuitive theories of mind to infer human cognitive processes and proposes machine learning challenges that target specific methodical leaps in service of this aim.

2. Barrett LF: **The conceptual act theory: a précis**. *Emot. Rev.* 2014, **6**:292-297 http://dx.doi.org/10.1177/1754073914534479.

3. Moors A: **Flavors of appraisal theories of emotion**. *Emot. Rev.* 2014, **6**:303-307 http://dx.doi.org/10.1177/1754073914534477.

4. Tracy JL: **An evolutionary approach to understanding distinct emotions**. *Emot. Rev.* 2014, **6**:308-312 http://dx.doi.org/10.1177/1754073914534478.

5. Adolphs R: **How should neuroscience study emotions? By distinguishing emotion states, concepts, and experiences**. *Soc. Cogn. Affect. Neurosci.* 2016 http://dx.doi.org/10.1093/scan/nsw153. nsw153–8.

6. Zaki J, Williams WC: **Interpersonal emotion regulation**. *Emotion* 2013, **13**:803-810 http://dx.doi.org/10.1037/a0033839.

7. Baker CL, Jara-Ettinger J, Saxe R, Tenenbaum JB: **Rational quantitative attribution of beliefs, desires, and percepts in human mentalizing**. *Nat. Hum. Behav.* 2017 . (in press) *https://www.nature.com/articles/s41562-017-0064*.

8. Murphy GL, Medin DL: **The role of theories in conceptual coherence**. *Psychol. Rev.* 1985, **92**:289-316.

9. Wellman HM: *Making Minds: How Theory of Mind Develops*. Oxford University Press; 2014.

10. Baker CL, Saxe R, Tenenbaum JB: **Action understanding as inverse planning**. *Cognition* 2009, **113**:329-349 http://dx.doi.org/10.1016/j.cognition.2009.07.005.

11. Jara-Ettinger J, Gweon H, Tenenbaum JB, Schulz LE: **Children's understanding of the costs and rewards underlying rational action**. *Cognition* 2015, **140**:14-23 http://dx.doi.org/10.1016/j.cognition.2015.03.006.

12. Gershman SJ, Gerstenberg T, Baker CL, Cushman FA: **Plans, habits, and theory of mind**. *PLoS One* 2016, **11**:e0162246-24 http://dx.doi.org/10.1371/journal.pone.0162246.

13. Ong DC, Zaki J, Goodman ND: **Emotions in lay explanations of behavior**. *CogSci*. Austin, TX: Cognitive Science Society; 2016, 360-365.

14. Böhm G, Pfister H-R: **How people explain their own and others' behavior: a theory of lay causal explanations**. *Front. Psychol.* 2015, **6**:139 http://dx.doi.org/10.3389/fpsyg.2015.00139.

15. Harris PL, de Rosnay M, Pons F: **Understanding emotion**. In *Handbook of Emotions,* fourth ed.. Edited by Barrett LF, Lewis M, Haviland-Jones JM.New York: Guildford Press; 2016.

16. Jara-Ettinger J, Gweon H, Schulz LE, Tenenbaum JB: **The naïve** • **utility calculus: computational principles underlying commonsense psychology**. *Trends Cogn. Sci.* 2016, **20**:589-604 http://dx.doi.org/10.1016/j.tics.2016.05.011.
Formal generative Bayesian model of how both children and adults use intuitive theories of others' subjective utility functions to infer the contents of their minds, and employ these inferences as bases for action. Such inferences can be sophisticated and abstract, spanning others' beliefs, desires, preferences, prior knowledge, and character traits.

17. Moutoussis M, Fearon P, El-Deredy W, Dolan RJ, Friston KJ: **Bayesian inferences about the self (and others): a review**. *Conscious. Cogn.* 2014, **25**:67-76 http://dx.doi.org/10.1016/j.concog.2014.01.009.

18. Ong DC, Zaki J, Goodman ND: **Affective cognition: exploring lay** •• **theories of emotion**. *Cognition* 2015, **143**:141-162 http://dx.doi.org/10.1016/j.cognition.2015.06.010.
Uses a formal intuitive theory of mind to model observer's quantitative predictions of others' emotions (in a lottery context) based on inferred event appraisals, based on emotional expressions, and based on combinations of both sources of information.

19. Wu Y, Baker CL, Tenenbaum JB: **Joint inferences of belief and desire from facial expressions**. *CogSci*. Austin, TX: Cognitive Science Society; 2014, 1796-1801.

20. The weeping Oscar Pistorius and a final question: Has it all been an act? https://www.washingtonpost.com/news/morning-mix/wp/2014/10/15/the-weeping-oscar-pistorius-and-a-final-question-has-it-all-been-an-act; 2014.

21. Armstrong K, Semien R, Miller TC, Glass I, Anatomy of Doubt, This American Life. https://www.thisamericanlife.org/radio-archives/episode/581/anatomy-of-doubt; 2016.

22. Gergely G, Csibra G: **Teleological reasoning in infancy: the naïve theory of rational action**. *Trends Cogn. Sci.* 2003, **7**:287-292 http://dx.doi.org/10.1016/S1364-6613(03)00128-1.

23. Skerry AE, Spelke ES: **Preverbal infants identify emotional** •• **reactions that are incongruent with goal outcomes**. *Cognition* 2014, **130**:204-216 http://dx.doi.org/10.1016/j.cognition.2013.11.002.
Three elegant experiments provide evidence that 10-month-old infants predict that an agent's emotional reactions will depend on whether the agent's inferred goals are fulfilled or not.

24. Repacholi BM, Meltzoff AN, Toub TS, Ruba AL: **Infants' generalizations about other people's emotions: foundations for trait-like attributions**. *Dev. Psychol.* 2016, **52**:364-378 http://dx.doi.org/10.1037/dev0000097.

25. Ong DC, Asaba M, Gweon H: **Young children and adults integrate past expectations and current outcomes to reason about others' emotions**. *CogSci*. Austin, TX: Cognitive Science Society; 2016, 135-140.

26. Ronfard S, Harris PL: **When will Little Red Riding Hood become scared? Children's attribution of mental states to a story character**. *Dev. Psychol.* 2014, **50**:283-292 http://dx.doi.org/10.1037/a0032970.

27. Ornaghi V, Grazzani I: **The relationship between emotional-state language and emotion understanding: a study with school-age children**. *Cogn. Emot.* 2013, **27**:356-366 http://dx.doi.org/10.1080/02699931.2012.711745.

28. Nelson JA, de Lucca Freitas LB, O'Brien M, Calkins SD, Leerkes EM, Marcovitch S: **Preschool-aged children's understanding of gratitude: relations with emotion and mental state knowledge**. *Br. J. Dev. Psychol.* 2012, **31**:42-56 http://dx.doi.org/10.1111/j.2044-835X.2012.02077.x.

29. Weimer AA, Sallquist J, Bolnick RR: **Young children's emotion comprehension and theory of mind understanding**. *Early Educ. Dev.* 2012, **23**:280-301 http://dx.doi.org/10.1080/10409289.2010.517694.

30. O'Brien M, Miner Weaver J, Nelson JA, Calkins SD, Leerkes EM, Marcovitch S: **Longitudinal associations between children's understanding of emotions and theory of mind**. *Cogn. Emot.* 2011, **25**:1074-1086 http://dx.doi.org/10.1080/02699931.2010.518417.

31. Scherer KR, Meuleman B: **Human emotion experiences can be predicted on theoretical grounds: evidence from verbal labeling**. *PLoS One* 2013, **8**:e58166-8 http://dx.doi.org/10.1371/journal.pone.0058166.

32. Skerry AE, Saxe R: **Neural representations of emotion are**
•• **organized around abstract event features**. *Curr. Biol.* 2015, **25**:1945-1954 http://dx.doi.org/10.1016/j.cub.2015.06.009.
Fine-grained emotion distinctions (20-way classification) were measured behaviorally and neurally in human observers. Information relevant to these distinctions was present in brain regions involved in intuitive theory of mind. Both behavioral and neural representations were well captured by a high-dimensional space of abstract event features.

33. Aviezer H, Trope Y, Todorov A: **Body cues, not facial expressions, discriminate between intense positive and negative emotions**. *Science* 2012, **338**:1225-1229 http://dx.doi.org/10.1126/science.1224313.

34. Bonnefon J-F, Hopfensitz A, De Neys W: **The modular nature of trustworthiness detection**. *J. Exp. Psychol. Gen.* 2013, **142**:143-150 http://dx.doi.org/10.1037/a0028930.

35. Russell JA: **A sceptical look at faces as emotion signals**. In *The Expression of Emotion: Philosophical, Psychological and Legal Perspectives.* Edited by Abell C, Smith J. Cambridge: Cambridge University Press; 2016:157-172.

36. Bartlett MS, Littlewort GC, Frank MG, Lee K: **Automatic decoding of facial movements reveals deceptive pain expressions**. *Curr. Biol.* 2014, **24**:738-743 http://dx.doi.org/10.1016/j.cub.2014.02.009.

37. Hassin RR, Aviezer H, Bentin S: **Inherently ambiguous: facial expressions of emotions, in context**. *Emot. Rev.* 2013, **5**:60-65 http://dx.doi.org/10.1177/1754073912451331.

38. Wenzler S, Levine S, van Dick R, Oertel-Knöchel V, Aviezer H:
• **Beyond pleasure and pain: facial expression ambiguity in adults and children during intense situations**. *Emotion* 2016, **16**:807-814 http://dx.doi.org/10.1037/emo0000185.
A replication and extention of Aviezer et al. [33], finding that observers cannot discriminate high intensity positive versus negative emotions, based on naturally occurring facial expressions.

39. Aviezer H, Messinger DS, Zangvil S, Mattson WI, Gangi DN, Todorov A: **Thrill of victory or agony of defeat? Perceivers fail to utilize information in facial movements**. *Emotion* 2015, **15**:791-797 http://dx.doi.org/10.1037/emo0000073.

40. Widen SC, Russell JA: **Differentiation in preschooler's categories of emotion**. *Emotion* 2010, **10**:651-661 http://dx.doi.org/10.1037/a0019005.

41. Russell JA, Suzuki N, Ishida N: **Canadian, Greek, and Japanese freely produced emotion labels for facial expressions**. *Motiv. Emot.* 1993, **17**:337-351 http://dx.doi.org/10.1007/BF00992324.

42. Mehu M, Scherer KR: **Emotion categories and dimensions in the facial communication of affect: an integrated approach**. *Emotion* 2015, **15**:798-811 http://dx.doi.org/10.1037/a0039416.

43. Dobs K, Bülthoff I, Breidt M, Vuong QC, Curio C, Schultz J:
• **Quantifying human sensitivity to spatio-temporal information in dynamic faces**. *Vis. Res.* 2014, **100**:78-87 http://dx.doi.org/10.1016/j.visres.2014.04.009.
Systematic manipulation of spatio-temporal information in dynamic faces found that people used sparse cues to process facial motion and gave noisy similarity ratings, but were highly sensitive to deviations from natural motion dynamics.

44. Martinez L, Falvello VB, Aviezer H, Todorov A: **Contributions of facial expressions and body language to the rapid perception of dynamic emotions**. *Cogn. Emot.* 2015, **30**:939-952 http://dx.doi.org/10.1080/02699931.2015.1035229.

45. Kayyal M, Widen S, Russell JA: **Context is more powerful than**
• **we think: contextual cues override facial cues even for valence**. *Emotion* 2015, **15**:287-291 http://dx.doi.org/10.1037/emo0000032.
Observers viewed real photographs of Olympic athletes at the moment of winning or losing their event, and were given either correct or misleading information about the context. Verbal descriptions of the event dominated the facial expression in determining the attributed emotion.

46. Zaki J: **Cue integration: a common framework for social cognition and physical perception**. *Perspect. Psychol. Sci.* 2013, **8**:296-312 http://dx.doi.org/10.1177/1745691613475454.

47. de Gelder B, de Borst AW, Watson R: **The perception of emotion in body expressions**. *WIREs Cogn. Sci.* 2015, **6**:149-158 http://dx.doi.org/10.1002/wcs.1335.

48. Schlegel K, Grandjean D, Scherer KR: **Emotion recognition: unidimensional ability or a set of modality- and emotion-specific skills?** *Pers. Individ. Differ.* 2012, **53**:16-21 http://dx.doi.org/10.1016/j.paid.2012.01.026.

49. Simon-Thomas ER, Keltner DJ, Sauter D, Sinicropi-Yao L, Abramson A: **The voice conveys specific emotions: evidence from vocal burst displays**. *Emotion* 2009, **9**:838-846 http://dx.doi.org/10.1037/a0017810.

50. Ma WJ, Zhou X, Ross LA, Foxe JJ, Parra LC: **Lip-reading aids word recognition most in moderate noise: a Bayesian explanation using high-dimensional feature space**. *PLoS One* 2009, **4**:e4638-14 http://dx.doi.org/10.1371/journal.pone.0004638.

51. Krumhuber EG, Scherer KR: **The look of fear from the eyes varies with the dynamic sequence of facial actions**. *Swiss J. Psychol.* 2016, **75**:5-14 http://dx.doi.org/10.1024/1421-0185/a000166.

52. Hehman E, Flake JK, Freeman JB: **Static and dynamic facial cues differentially affect the consistency of social evaluations**. *Pers. Soc. Psychol. Bull.* 2015, **41**:1123-1134 http://dx.doi.org/10.1177/0146167215591495.

53. Todorov A, Olivola CY, Dotsch R, Mende-Siedlecki P: **Social attributions from faces: determinants, consequences, accuracy, and functional significance**. *Annu. Rev. Psychol.* 2015, **66**:519-545 http://dx.doi.org/10.1146/annurev-psych-113011-143831.

54. Todorov A, Porter JM: **Misleading first impressions**. *Psychol. Sci.* 2014, **25**:1404-1417 http://dx.doi.org/10.1177/0956797614532474.

55. Goren A, Todorov A: **Two faces are better than one: eliminating false trait associations with faces**. *Soc. Cogn.* 2009, **27**:222-248 http://dx.doi.org/10.1521/soco.2009.27.2.222.

56. Jack RE, Garrod OGB, Schyns PG: **Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time**. *Curr. Biol.* 2014, **24**:187-192 http://dx.doi.org/10.1016/j.cub.2013.11.064.

57. Widen SC, Russell JA: **Do dynamic facial expressions convey emotions to children better than do static ones?** *J. Cogn. Dev.* 2015, **16**:802-811 http://dx.doi.org/10.1080/15248372.2014.916295.

58. Mumenthaler C, Sander D: **Automatic integration of social information in emotion recognition**. *J. Exp. Psychol. Gen.* 2015, **144**:392-399 http://dx.doi.org/10.1037/xge0000059.

59. Saxe R: **Against simulation: the argument from error**. *Trends Cogn. Sci.* 2005, **9**:174-179 http://dx.doi.org/10.1016/j.tics.2005.01.012.

60. Coleman MD: **Emotion and the false consensus effect**. *Curr. Psychol.* 2016:1-7 http://dx.doi.org/10.1007/s12144-016-9489-0.

61. Miloyan B, Suddendorf T: **Feelings of the future**. *Trends Cogn. Sci.* 2015, **19**:196-200 http://dx.doi.org/10.1016/j.tics.2015.01.008.

62. Cooney G, Gilbert DT, Wilson TD: **The unforeseen costs of extraordinary experience**. *Psychol. Sci.* 2014, **25**:2259-2265 http://dx.doi.org/10.1177/0956797614551372.

63. Sebastian CL, Fontaine NMG, Bird G, Blakemore S-J, De Brito SA, McCrory EJP *et al.*: **Neural processing associated with cognitive and affective theory of mind in adolescents and adults**. *Soc. Cogn. Affect. Neurosci.* 2012, **7**:53-63 http://dx.doi.org/10.1093/scan/nsr023.

64. Ferrari C, Lega C, Vernice M, Tamietto M, Mende-Siedlecki P, Vecchi T *et al.*: **The dorsomedial prefrontal cortex plays a causal role in integrating social impressions from faces and verbal descriptions**. *Cereb. Cortex* 2016, **26**:156-165 http://dx.doi.org/10.1093/cercor/bhu186.

65. Skerry AE, Saxe R: **A common neural code for perceived and inferred emotion**. *J. Neurosci.* 2014, **34**:15997-16008 http://dx.doi.org/10.1523/JNEUROSCI.1676-14.2014.

66. Lee KH, Siegle GJ: **Different brain activity in response to emotional faces alone and augmented by contextual information**. *Psychophysiology* 2014, **51**:1147-1157 http://dx.doi.org/10.1111/psyp.12254.

67. Watson R, Latinus M, Noguchi T, Garrod O, Crabbe F, Belin P: **Crossmodal adaptation in right posterior superior temporal sulcus during face-voice emotional integration**. *J. Neurosci.* 2014, **34**:6813-6821 http://dx.doi.org/10.1523/JNEUROSCI.4478-13.2014.

68. Srinivasan R, Golomb JD, Martinez AM: **A neural basis of facial action recognition in humans**. *J. Neurosci.* 2016, **36**:4434-4442 http://dx.doi.org/10.1523/JNEUROSCI.1704-15.2016.

69. Wegrzyn M, Riehle M, Labudda K, Woermann F, Baumgartner F, Pollmann S *et al.*: **Investigating the brain basis of facial expression perception using multi-voxel pattern analysis**. *Cortex* 2015, **69**:131-140 http://dx.doi.org/10.1016/j.cortex.2015.05.003.

70. Kotz SA, Kalberlah C, Bahlmann J, Friederici AD, Haynes J-D: **Predicting vocal emotion expressions from the human brain**. *Hum. Brain Mapp.* 2012, **34**:1971-1981 http://dx.doi.org/10.1002/hbm.22041.

71. Anderson LC, Rice K, Chrabaszcz J, Redcay E: **Tracking the neurodevelopmental correlates of mental state inference in early childhood**. *Dev. Neuropsychol.* 2015, **40**:379-394 http://dx.doi.org/10.1080/87565641.2015.1119836.

72. Rice K, Redcay E: **Spontaneous mentalizing captures variability in the cortical thickness of social brain regions**. *Soc. Cogn. Affect. Neurosci.* 2015, **10**:327-334 http://dx.doi.org/10.1093/scan/nsu081.

73. Rice K, Viscomi B, Riggins T, Redcay E: **Amygdala volume linked to individual differences in mental state inference in early childhood and adulthood**. *Dev. Cogn. Neurosci.* 2014, **8**:153-163 http://dx.doi.org/10.1016/j.dcn.2013.09.003.

74. Lindquist KA, Gendron M, Barrett LF, Dickerson BC: **Emotion perception, but not affect perception, is impaired with semantic memory loss**. *Emotion* 2014, **14**:375-387 http://dx.doi.org/10.1037/a0035293.

75. Lindquist KA, Satpute AB, Gendron M: **Does language do more than communicate emotion?** *Curr. Dir. Psychol. Sci.* 2015, **24**:99-108 http://dx.doi.org/10.1177/0963721414553440.

76. Lindquist KA, Gendron M: **What's in a word? Language constructs emotion perception**. *Emot. Rev.* 2013, **5**:66-71 http://dx.doi.org/10.1177/1754073912451351.

77. Gratch J, Marsella S: **Appraisal models**. In *The Oxford Handbook of Affective Computing.* Edited by Calvo RA, D'Mello S, Gratch J, Kappas A. USA: Oxford University Press; 2014:54-67.