From faces to prosocial behavior: cues, tools, and mechanisms

Ralph Adolphs[1] and **Anita Tusche**[1]

[1]Division of the Humanities and Social Sciences, California Institute of Technology, HSS 228-77, Pasadena, CA 91125, U.S.A.

Abstract

In this review we ask how looking at people's faces can influence prosocial behaviors towards them. Components of this process have often been studied by disparate literatures: one focused on perception and judgment of faces, using both psychological and neuroscience approaches; and a second focused on actual social behaviors, as studied in behavioral economics and decision science. Bridging these disciplines requires a more mechanistic account of how processing of particular face attributes or features influences social judgments and behaviors. Here we review these two lines of research, and suggest that combining some of their methodological tools can provide the bridging mechanistic explanations.

Keywords

INTRODUCTION

Faces represent a potent and rich source of information, for instance about people's identity (e.g., are they kin?), emotional state (e.g., are they distressed?), or attractiveness, all of which can shape social behaviors such as helping or cooperation. We also routinely rely on facial cues to make inferences about people's *personality*, such as whether a person is trustworthy or not (Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). Moreover, we derive such social judgments from faces very rapidly and with astonishingly little effort: trustworthiness judgments can be made reliably within 100ms or less. Social judgments from faces are automatic, unrelated to intelligence, and seem to satisfy all the criteria for an encapsulated "module" that delivers a perceived trustworthiness judgment outside any deliberative control (Bonnefon, Hopfensitz, & De Neys, 2013). Attractiveness is another

RECOMMENDED READINGS

(Freeman & Johnson, 2016). (See References). A review of how rapid face perception can generate biases and stereotypes, and how these might be implemented in the brain.

(Todorov et al., 2015). (See References). A comprehensible, highly accessible overview of the dimensions and features of faces that guide our social judgments about them.

(Maestripieri et al., 2017) (See References). A clearly written, user-friendly review of prosocial biases (in the real-world and laboratory setting) in favor of attractive faces from various research disciplines.

Adolphs, R. (2008). Fear, faces, and the human amygdala. Current Opinion in Neurobiology. 18:1–7. This paper reviews neuroscientific evidence on the role of the amygdala in face processing and fear perception, and discusses the automatic and non-conscious aspects of these processes.

popular example of how faces shape social inferences about others' personality. People tend to attribute more positive characteristics to physically attractive compared to unattractive strangers (e.g. generosity, intelligence, or trustworthiness), which affects a wide variety of social behaviors (reviewed in (Maestripieri, Henry, & Nickels, 2017).

The ultimate factors that generate our social judgments based on faces are many. At a minimum, they include the detailed features and their configuration of the face (e.g. physiognomic features such as symmetry), how it relates to other faces (e.g., how close to the 'average' face) or how similar the face is to our own face (reflecting genetic relatedness) (Todorov et al., 2015). Considerable work in developmental, evolutionary, and social psychology provides initial clues about how specific face attributes are linked to social judgments and to prosocial (or antisocial) behaviors. For instance, physiognomic features of male faces such as the testosterone-related width-to-height ratio provide cues about whom to trust, and affect cooperative behavior. In particular, men with proportionally wider faces are perceived as less trustworthy, and indeed are more likely to act in their own self-interest and violate other's trust (M Stirrat & Perrett, 2010), although context might play a role (M. Stirrat & Perrett, 2012). Another example concerns facial cues of self-resemblance (signaling kinship) which in turn can motivate prosocial behavior (A. Marsh, 2016): faces that are more similar to oneself are perceived as more trustworthy (DeBruine, 2011) and facilitate cooperation (Krupp, DeBruine, & Barclay, 2008).

Yet the precise mechanisms underlying these findings remain largely unknown. Studies on the effects of faces on prosociality rarely spell out the mediating mental or neural mechanisms, in good part because they typically focus on tools from a single discipline, and describe only a piece of the entire process. Here we briefly introduce the relevant literatures, and suggest that putting together the pieces to provide a more comprehensive mechanistic account will require combining their approaches and tools. We begin with an overview of face perception, then turn to prosocial behavior, and conclude with a synthesis of tools from these disciplines.

FACE PERCEPTION AND THE BRAIN

Understanding face perception, and the social judgments that build on it, has been substantially informed by recent neuroscience studies. What is clear from neuroscience data is that a comprehensive representation of a face – of an object comprised of many features all configurally bound into a Gestalt percept – requires interactions within a network of brain structures that all implement somewhat distinct psychological processes. For instance, it requires brain structures that process the features of the face and their spatial relationships-- the eyes, the nose, the mouth, and how these are located with respect to one another. This includes brain regions such as the fusiform face area (FFA) and superior temporal sulcus (STS) (see Figure 1, left panel). What exactly each of these brain regions represents, and how they communicate with one another to synthesize a comprehensive visual representation of the entire face, is being worked out in great detail by studies that use a combination of neuroimaging (e.g. functional magnetic resonance imaging; Figure 1, right panel) and recordings from single brain cells (Freiwald, Tsao, & Livingstone, 2009). One approximate scheme is that some regions (like the FFA) represent the static, physiognomic

appearance of a face, whereas other regions (like the STS) represent changeable features in faces, corresponding to encoding of the identity vs. the emotional expression of a person, respectively (Haxby, Hoffman, & Gobbini, 2000). For example, the individual sets of muscles whose movements comprise emotional facial expressions can be decoded from neuroimaging patterns in the STS (Srinivasan, Golomb, & Martinez, 2016). What neuroscience data show is that there are various processes, occurring to some extent in distinct brain regions, that assemble a full perceptual representation of a face. A bias in any one of these components could thus implement the effect of a specific facial cue on social judgments and social behavior, examples we turn to next.

Let's look at the positive bias in favor of physically attractive people mentioned earlier. Reward-related regions of the brain, such as the orbitofrontal cortex (Figure 1), are activated by the perceived attractiveness of faces. These brain regions are thus likely candidates for neural processes that mediate the automatic biases for social judgments and generous behaviors based on attractiveness. Attractiveness judgments also activate the dorsomedial prefrontal cortex, a brain region that is also recruited when people make face-based inferences about other people's personality traits, such as their trustworthiness (Bzdok et al., 2012). This brain region is strongly implicated in all social judgments that require some level of abstraction and causal inference, so that we can attribute mental states and personalities to people from their observed behavior. In the case of faces, the dorsomedial prefrontal cortex is automatically activated whenever we see facial expressions (in people or in animals), since we spontaneously attribute emotions to them from seeing their expressions (Spunt, Ellsworth, & Adolphs, 2016). Representing reward value and inferring mental state are thus at least two separate processes that contribute to prosocial behaviors towards attractive faces.

Other brain structures relevant for social judgments based on faces include the amygdala and the insula (Figure 1). Trustworthiness judgments based on facial characteristics have been shown to involve brain responses in these regions (Winston, Strange, O'Doherty, & Dolan, 2002). Whereas the amygdala may provide a rapid and coarse evaluation of faces, and help direct attention to their features, the insula is thought to represent our own bodily reactions to the face – that is, how we feel about it. Focal damage to the amygdala in rare patients has provided some of the most dramatic deficits in social judgments from faces. For instance, such patients judge faces to look abnormally trustworthy and approachable (R Adolphs, Tranel, & Damasio, 1998) and also are unable to recognize fear from facial expressions (R Adolphs, Tranel, Damasio, & Damasio, 1994). This latter finding has been linked to a particular attentional impairment: patients with amygdala damage fail to judge faces as fearful because they do not look at the eye region of the face, a bias that can be revealed with eyetracking (see below) (R. Adolphs et al., 2005). This last study ties together social judgments, a particular facial feature (the eyes), and a specific brain structure (the amygdala), and is an example of the kind of mechanistic explanation we would ultimately like to have for all social judgments from faces, and their impact on social behavior.

FACES AND PROSOCIAL BEHAVIOR

Interestingly, neuroscience studies on the functional link of face perception and prosocial behavior also indicate that the amygdala might play a key role. Exceptionally altruistic people who volunteered to give up a kidney for the benefit of a total stranger, showed higher neural activity in the right amygdala when briefly exposed to fearful faces (A. A. Marsh et al., 2014). This difference in neural responses in the amygdala during face processing was also linked to superior recognition accuracy for these facial expressions. One possible explanation for these findings is that a heightened sensitivity to visual cues of personal distress might underlie increased motivations to respond altruistically to the person in distress. Besides real-world measure of altruistic behavior such as organ donation, increased sensitivity to fearful facial expressions also predicted individual differences in prosocial behavior assessed in the laboratory (A. A. Marsh, Kozak, & Ambady, 2007). This evidence clearly suggests that individual differences in face processing are linked to individual differences in altruistic behavior, mediated by variability in specific brain regions such as the amygdala.

In addition to facial expression, the mere physiognomy of the face (its neutral appearance in a person) also biases prosocial behavior. For example, prosocial biases in favor of physically attractive people have been observed in door-to-door fundraising (Landry, Lange, List, Price, & Rupp, 2006) and charitable donation behavior (Price, 2008; Raihani & Smith, 2015). Effects of facial attractiveness on prosocial decisionmaking have been observed in laboratory settings using economic game-theoretical paradigms: attractive people were offered more money, compared to unattractive players (Rosenblat, 2008; Solnick & Schweitzer, 1999) as signaled for instance in higher facial asymmetry (Zaatari, Palestis, & Trivers, 2009). Interestingly, physically attractive people themselves are actually *less* generous, less cooperative and less trustworthy (Maestripieri et al., 2017), suggesting that although we reliably infer traits about people from their faces, these judgments are often not valid. While these findings clearly show that facial cues are linked to individual differences in people's prosocial behavior, we know surprisingly little about the precise mental and neural mechanisms that link them. We propose that utilizing advanced tools traditionally used in different research disciplines might be utilized to bridge this gap.

TOOLS

There are a number of tools available for extracting dimensions or features from faces that correlate with specific social judgments. Some of these merely answer the question, "which particular regions of a face influence a social judgment the most?" Others go further than this and allow us also to ask, "what mechanism might be mediating that effect?"

The most commonly used tool is eyetracking, which quantifies eye movements via remote or head-mounted devices, allowing us to analyze where people are attending, and what specific facial information they are processing. With the availability of easy-to-use, high-resolution eyetrackers (with temporal sampling rates typically between 100–1000 Hz) that do not require headmounting, collecting such data is commonplace.

However, these data can also be analyzed with more sophisticated approaches. Thus, a second tool concerns analyses that can map eyetracking data onto psychologically meaningful attributes or dimensions. In one example, a linear classifier using a machine learning algorithm was trained on the fixations that people make to objects and faces in complex visual scenes, and the resultant model was then tested on a holdout dataset. The net result produced a fairly detailed inventory of the relative weight that various attributes of the visual scenes exert on visual attention, that is, their visual saliency (Xu, Jiang, Wang, Kankanhalli, & Zhao, 2014) (Figure 2, upper panel). Interestingly, this was done on an individual subject basis (with about 700 images) and used to study individual differences (Wang et al., 2015). You could think of this approach as analogous to a big regression model in which we have asked how strongly different features in a visual stimulus predict the location of where people will look—some features attract visual attention strongly (e.g., the eyes in face), whereas others do not.

The third type of tool comes from models that can take the results from the above two tools, and use these to predict behavioral choices (e.g., Figure 2, bottom panel). This set of tools comprises models that convert where we look and attend into decisions. One class of models accumulates sensory evidence over time, and one of the most influential of these models are so-called drift-diffusion models (DDMs) (Ratcliff & McKoon, 2008). These models have been successfully used to describe *perceptual* decision-making and are particularly powerful for various reasons: they are neurobiologically plausible, allow the estimation of parameters that correspond to specific psychological processes, can be fit from a range of different dependent measures (e.g., reaction times, visual fixations, firing rates of neurons in the brain), and can be extended to more than two behavioral options. Within the framework of these models, looking at an available choice-option or their choice-relevant feature contributes to the noisy evidence accumulation over time (Figure 1, bottom panel). As enough evidence is gathered and one of the two decision barriers is crossed, a decision in favor of this choice option is made. This means that eyetracking data can be directly incorporated into the DDM (Krajbich, Armel, & Rangel, 2010). One natural hypothesis, which has not yet been tested, is that a similar approach could be taken also for the features within faces: the more we look at somebody's eyes, nose, or mouth (or any relevant facial cue for that matter), the more information about this facial feature would bias our social judgment and behavior towards the person.

A final set of tools probes the dimensions or features of faces more directly by manipulating them: computer-generated faces can parametrically manipulate width-to-height ratio, skin color, or indeed any configuration that reliably correlates with a social judgment. Another approach uses random sampling of face regions, or adds random noise to faces, in order to extract, over many trials, those regions of a face where variability is most strongly associated with a social judgment (see (Todorov et al., 2015) for review). There are a number of such data-driven approaches being used in order to discover facial features or dimensions that one might not even have hypothesized to play a role in prosocial behaviors (R Adolphs, Nummenmaa, Todorov, & Haxby, 2016). These approaches complement the above set, and taken together allow us to investigate how facial features link to prosocial behaviors with both a broad, data-driven survey, and with more focused hypotheses.

FUTURE DIRECTIONS

The framework we have sketched suggests several future directions. First and most obviously, it would motivate specific hypotheses about the mediating psychological processes (and their neural mechanisms) that link attention to faces, on the one hand, to aspects of prosocial behavior, on the other hand. To test these hypotheses, one would need to go from the face (e.g., with eyetracking studies) to the behavior (e.g., with behavioral economics studies), and incorporate the data generated into quantitative models (e.g., DDMs, machine-learning analyses of eyetracking data). Second, it offers sensitive and quantitative assessment not only for individual differences in these processes, but also for help in diagnosis of psychiatric disorders. For instance the study we highlighted (Wang et al., 2015) used model-based eyetracking to investigate how people with autism view stimuli like faces differently. Third, while we have assumed throughout that attention to faces has a causal influence on prosocial behavior, the relation could of course go in the opposite direction as well (individual differences in people's prosociality may drive attention to faces) or both could be embedded in more complex networks of common causes, and these causal effects should be investigated as well. Fourth, future studies might explore the role of social processes such as empathy (that reliably predict prosocial acts, (Tusche, Bockler, Kanske, Trautwein, & Singer, 2016) as mediating factor that links attention to facial cues (e.g. related to distress) to subsequent helping. Finally, our framework suggests some speculative interventions for how people might become more prosocial in their behavior (Zaki & Cikara, 2015): if we can manipulate how people look at each other, we might be able to influence how they behave towards one another.

Acknowledgments

Bibliography

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. [PubMed: 15635411]

Adolphs R, Nummenmaa L, Todorov A, Haxby JV. Data-driven approaches in the investigation of social perception. Philos Trans R Soc Lond B Biol Sci. 2016; 371

Adolphs R, Tranel D, Damasio AR. The Human Amygdala in Social Judgment. Nature. 1998; 393:470–474. [PubMed: 9624002]

Adolphs R, Tranel D, Damasio H, Damasio A. Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. Nature. 1994; 372:669–672. [PubMed: 7990957]

Bonnefon JF, Hopfensitz A, De Neys W. The modular nature of trustworthiness detection. J Exp Psychol Gen. 2013; 142:143–150. [PubMed: 22686638]

Bzdok D, Langner R, Hoffstaedter F, Turetsky BI, Zilles K, Eickhoff SB. The modular neuroarchitecture of social judgments on faces. Cereb Cortex. 2012; 22:951–961. [PubMed: 21725038]

DeBruine L. Opposite-sex siblings decrease attraction, but not prosocial attributions, to self-resembling opposite-sex faces. PNAS. 2011; 108:11710–11714. [PubMed: 21709272]

Freeman JB, Johnson KL. More than meets the eye: split-second social perception. Trends in Cognitive Sciences. 2016; 20:362–374. [PubMed: 27050834]

Freiwald WA, Tsao DY, Livingstone MS. A face feature space in the macaque temporal lobe. Nature Neuroscience. 2009; 12:1187–1196. [PubMed: 19668199]

Haxby JV, Hoffman EA, Gobbini MI. The distributed human neural system for face perception. Trends in Cognitive Science. 2000; 4:223–233.

Hutcherson CA, Bushong B, Rangel A. A Neurocomputational Model of Altruistic Choice and Its Implications. Neuron. 2015; 87(2):451–462. [PubMed: 26182424]

Krajbich I, Armel C, Rangel A. Visual fixations and the computation and comparison of value in simple choice. Nature Neuroscience. 2010; 13:1292–1298. [PubMed: 20835253]

Krupp DB, DeBruine LM, Barclay P. A cue of kinship promotes cooperation for the public good. Evolution and Human Behavior. 2008; 29:49–55.

Landry C, Lange A, List JA, Price MK, Rupp N. Toward an understanding of the economics of charity: evidence from a field experiment. Quarterly Journal of Economics. 2006; 121:747–782.

Maestripieri D, Henry A, Nickels N. Explaining financial and prosocial biases in favor of attractive people: interdisciplinary perspectives from economics, social psychology, and evolutionary psychology. Behav Brain Sci, forthcoming. 2017

Marsh A. Neural, cognitive and evolutionary foundations of human altruism. Cognitive Science. 2016; 7:59–71. [PubMed: 26685796]

Marsh AA, Kozak MN, Ambady N. Accurate identification of fear facial expressions predicts prosocial behavior. Emotion. 2007; 7(2):239–251. [PubMed: 17516803]

Marsh AA, Stoycos SA, Brethel-Haurwitz KM, Robinson P, VanMeter JW, Cardinale EM. Neural and cognitive characteristics of extraordinary altruists. Proc Natl Acad Sci U S A. 2014; 111(42): 15036–15041. [PubMed: 25225374]

Price M. Fund-raising success and a solicitor's beauty capital: blondes raise more funds? Economics Letters. 2008; 100(351–354)

Raihani NJ, Smith S. Competitive helping in online giving. Current Biology. 2015; 25:1–4. [PubMed: 25532895]

Ratcliff R, McKoon G. The diffusion decision model: theory and data for two-choice decision tasks. Neural Computation. 2008; 20:873–922. [PubMed: 18085991]

Rosenblat TS. The beauty premium: Physical attractiveness and gender in dictator games. Negotiation Journal. 2008; 24:465–481.

Solnick SJ, Schweitzer ME. The influence of physical attractiveness and gender on ultimatum game decisions. Organizational Behavior and Human Decision processes. 1999; 79:199–215. [PubMed: 10471361]

Spunt R, Ellsworth E, Adolphs R. The neural basis of understanding the expression of the emotions in man and animals. Social Cognitive and Affective Neuroscience. 2016 in press.

Srinivasan R, Golomb JD, Martinez AM. A neural basis of facial action recognition in humans. The Journal of Neuroscience. 2016; 36:4434–4442. [PubMed: 27098688]

Stirrat M, Perrett D. Valid facial cues to cooperation and trust: male facial width and trustworthiness. Psychological Science. 2010 in press.

Stirrat M, Perrett DI. Face structure predicts cooperation: men with wider faces are more generous to their in-group when out-group competition is salient. Psychol Sci. 2012; 23(7):718–722. [PubMed: 22623509]

Todorov A, Olivola CY, Dotsch R, Mende-Siedlecki. Social attributions from faces: determinants, consequences, accuracy, and functional significance. Annual Review of Psychology. 2015; 66:519–545.

Tusche A, Bockler A, Kanske P, Trautwein FM, Singer T. Decoding the Charitable Brain: Empathy, Perspective Taking, and Attention Shifts Differentially Predict Altruistic Giving. J Neurosci. 2016; 36(17):4719–4732. [PubMed: 27122031]

Wang S, Jiang M, Duchesne XM, Kennedy DP, Adolphs R, Zhao Q. Atypical visual saliency in autism spectrum disorder quantified through model-based eyetracking. Neuron. 2015; 88:604–616. [PubMed: 26593094]

Winston JS, Strange BA, O'Doherty J, Dolan RJ. Automatic and intentional brain responses during evaluation of trustworthiness of faces. Nature Neuroscience. 2002; 5:277–283. [PubMed: 11850635]

Xu J, Jiang M, Wang S, Kankanhalli MS, Zhao Q. Predicting human gaze beyond pixels. Journal of Vision. 2014; 14:28.

Zaatari D, Palestis BG, Trivers R. Fluctuating asymmetry of responders affects offers in the ultimatum game oppositely according to attractiveness or need as perceived by proposers. Ethology. 2009; 115:627–632.

Zaki J, Cikara M. Addressing empathic failures. Current Directions in Psychological Science. 2015; 24:471–476.
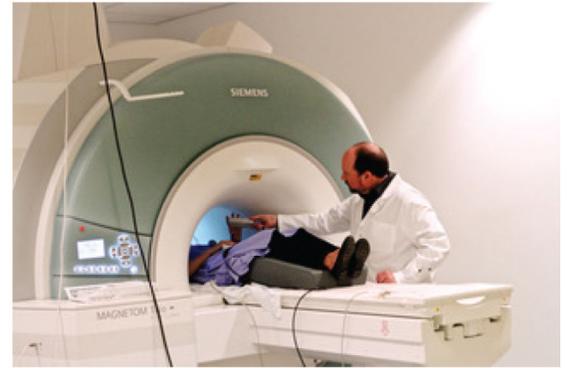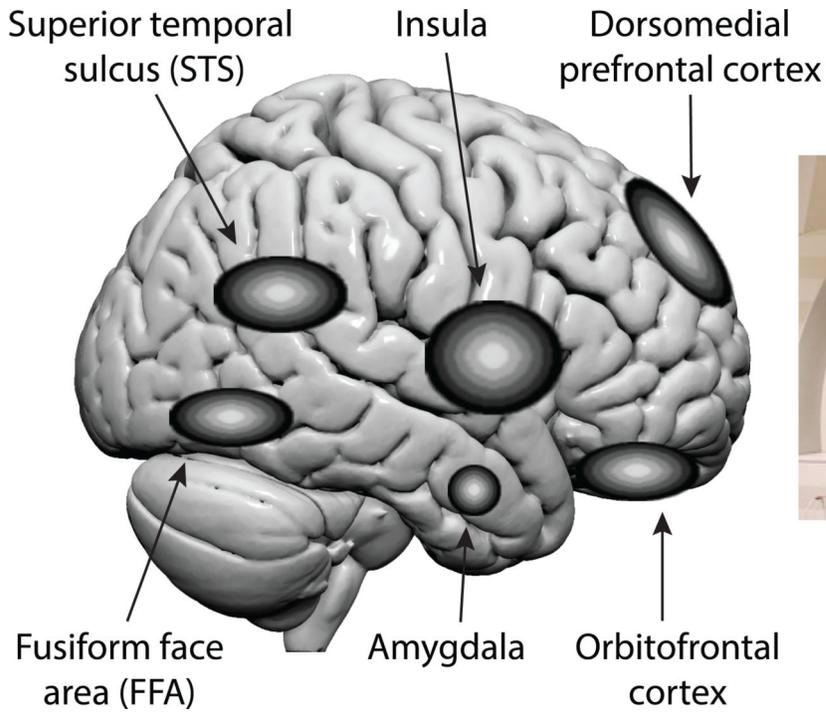
Figure 1.
Brain regions. The left panel of the figure shows schematically some of the brain regions mentioned in the text. Note that several of these (e.g. amygdala, insula) are in fact interior to the brain so that their location here represents where they are if projected onto the lateral surface. The right panel illustrates the technical equipment in functional magnetic resonance imaging (fMRI) studies that measure and map brain activity. This technique is noninvasive and safe.

Contributions of visual features on gaze behavior
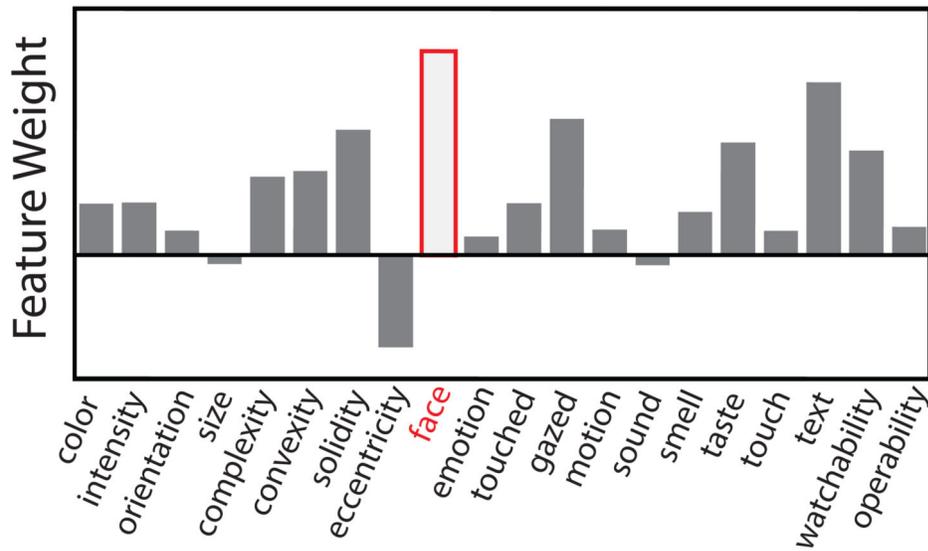
Drift-diffusion model (DDM) of decision-making
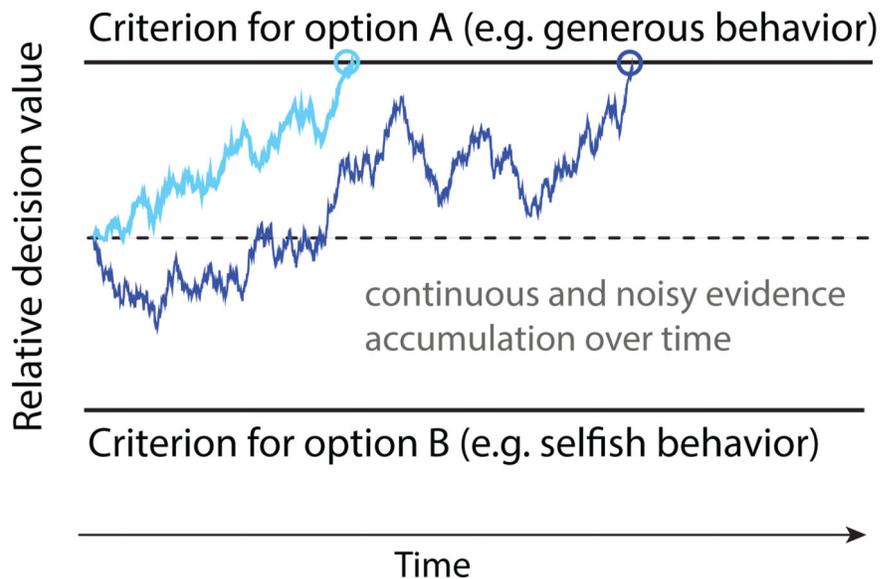
Figure 2.

Tools. The upper panel illustrates model-based analysis of eyetracking data that yield saliency weights onto specific features of visual stimuli. In the current example, these features were defined for complex, real-world visual scenes and their weights were computed using advanced machine learning algorithms. The high weight for faces reflects the fact that, when looking at a scene, viewers tend to fixate faces most frequently (modified from Wang et al., 2015). The bottom panel illustrates a driftdiffusion model as applied to altruistic choices. The curves (blue) plot the relative decision value in favor of one or the

other behavioral option (e.g. act prosocially or selfishly), as a function of time during the decision process. This accumulation of evidence over time is stochastic and noisy, as reflected in the moment-by-moment fluctuations of the plots. A decision is made once we have enough evidence accumulated and one of the thresholds is reached (upper and lower black lines). Critically, how much attention is paid to choice-options or relevant features can bias the evolution of the curve. In this framework, attention (e.g. as measured in gaze behavior) can bias the decision in favor of a generous decision, as illustrated by the light blue line that reaches the upper decision barrier earlier than in case of the dark blue line. Modified from (Hutcherson, Bushong, & Rangel, 2015).