Current Directions in Psychological Science
2017, Vol. 26(3) 282–287

DOI: 10.1177/0963721417694656
www.psychologicalscience.org/CDPS

From Faces to Prosocial Behavior: Cues, Tools, and Mechanisms

Ralph Adolphs and Anita Tusche
Division of the Humanities and Social Sciences, California Institute of Technology

Abstract
In this review, we ask how looking at people's faces can influence prosocial behaviors toward them. Components of this process have been studied in two disparate literatures: one focused on the perception of faces and judgments based on them, 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 mechanistic account of how processing of particular face attributes or features influences social judgments and behaviors. We review these two lines of research and suggest that combining some of their methodological tools may reveal the bridging mechanistic explanations.

Keywords
face perception, social judgments, eye tracking, drift-diffusion model, altruism

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 rapidly and with astonishingly little effort—for instance, trustworthiness judgments can be made reliably from faces shown for 100 milliseconds or less. Social judgments from faces are automatic and unrelated to intelligence, and they seem to satisfy all the criteria for an encapsulated "module" that delivers a judgment about perceived trustworthiness without any deliberative control (Bonnefon, Hopfensitz, & De Neys, 2013). Attractiveness is another popular example of how faces shape social inferences about others' personality. People tend to attribute more positive characteristics to physically attractive than to unattractive strangers (e.g., generosity, intelligence, trustworthiness), which affects a wide variety of social behaviors (reviewed in Maestripieri, Henry, & Nickels, 2017).

The factors that generate our social judgments based on faces are many. At a minimum, they include the detailed features of the face and their configuration (e.g., physiognomic features such as symmetry), how the face relates to other faces (e.g., how close it is to the "average" face), and how similar the face is to our own face (reflecting genetic relatedness; Todorov et al., 2015). Considerable work in developmental, evolutionary, and social psychology has provided 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, which 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 others' trust (Stirrat & Perrett, 2010), although this also depends on context (Stirrat & Perrett, 2012). Another example concerns facial cues of self-resemblance (signaling kinship), which can motivate prosocial behavior (A. Marsh, 2016): Faces that are more similar to one's own face 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

Corresponding Author:
Ralph Adolphs, HSS 228-77, California Institute of Technology, Pasadena, CA 91125
E-mail: radolphs@hss.caltech.edu

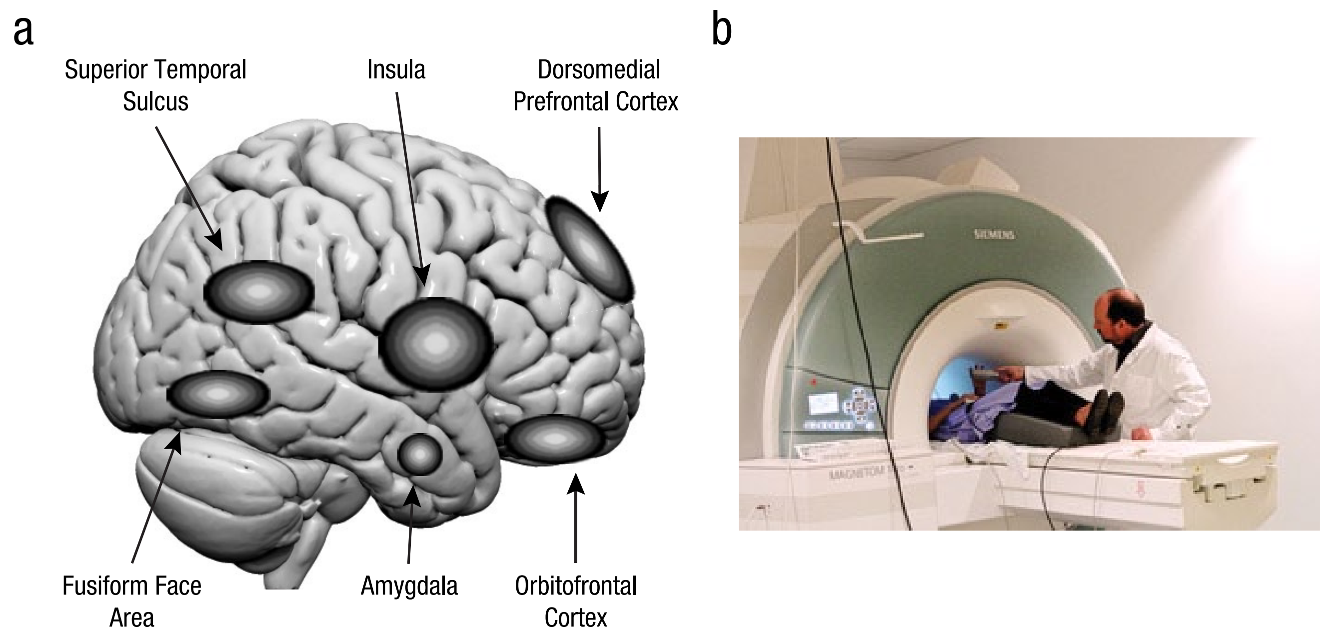

Fig. 1. Brain regions involved in face processing and face-based social judgments, and one tool for studying them. Panel (a) shows schematically some of the brain regions mentioned in the text. Note that several of these (e.g., the amygdala, the insula) are in fact interior to the brain; their location here represents where they would be if projected onto the lateral surface. Panel (b) illustrates the technical equipment used in fMRI studies, which measure and map brain activity. This technique is noninvasive and safe.

on prosociality have rarely spelled out the mediating mental or neural mechanisms, in good part because they have typically used tools from a single discipline and described 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

Our understanding of face perception, and the social judgments that build on it, has been substantially informed by recent neuroscience studies. It is clear from neuroscience data that a comprehensive representation of a face—of an object comprising many features 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 processing involves brain regions such as the fusiform face area (FFA) and the superior temporal sulcus (STS; see Fig. 1a). What exactly these brain regions each represent, and how they communicate with one another to synthesize a comprehensive visual representation of the entire face, has begun to be worked out in great detail by studies using a combination of neuroimaging (e.g., fMRI; Fig. 1b) and recordings from single brain cells (Freiwald, Tsao, & Livingstone, 2009). One approximate scheme is that some regions (e.g., the FFA) represent the static, physiognomic appearance of a face, whereas other regions (e.g., the STS) represent changeable features in faces, corresponding to the encoding of the identity and the emotional expression of a person, respectively (Haxby, Hoffman, & Gobbini, 2000). For example, the individual sets of muscles whose movements constitute emotional facial expressions can be decoded from neuroimaging patterns in the STS (Srinivasan, Golomb, & Martinez, 2016). Neuroscience data show 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 processes could thus implement the effect of a specific facial cue on social judgments and social behavior, examples of which 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 (Fig. 1a), 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 various social judgments that require some level of abstraction and causal inference, including attributions of mental states and personalities to people on the basis of 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, plausibly because we spontaneously attribute emotions to them upon seeing their expressions (Spunt, Ellsworth, & Adolphs, 2016). Representing reward value and inferring a person's mental state are thus at least two separate processes that may contribute to prosocial behaviors toward people with attractive faces.

Other brain structures relevant for social judgments based on faces include the amygdala and the insula (Fig. 1a). 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 illuminated some of the most dramatic deficits in social judgments from faces. For instance, such patients judge faces to look abnormally trustworthy and approachable (Adolphs, Tranel, & Damasio, 1998) and are unable to recognize fear from facial expressions (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 eye tracking (as described further in the following section; Adolphs et al., 2005). This last study tied together social judgments, a particular facial feature (the eyes), and a specific brain structure (the amygdala), and it is an example of the kind of mechanistic explanation we would ultimately like to have for all social judgments based on faces and their impact on social behavior.

Faces and Prosocial Behavior

Interestingly, neuroscience studies on the functional link between face perception and prosocial behavior have also indicated that the amygdala might play a key role. Compared with controls, 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 accuracy in recognizing fearful 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 people in distress. Besides real-world measures of altruistic behavior such as organ donation, increased sensitivity to fearful facial expressions also predicts increased prosocial behavior as 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.

Beyond facial expression, the mere physiognomy of the face (i.e., 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 decision making have also been observed in laboratory settings using economic game-theoretical paradigms: Players were offered more money if they were more attractive (Rosenblat, 2008; Solnick & Schweitzer, 1999), as signaled, for instance, by higher facial symmetry (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. Although these findings clearly show that facial cues are linked to individual differences in prosocial behavior, we know surprisingly little about the precise mental and neural mechanisms that link them. We propose that combining the advanced tools traditionally used in different research disciplines might help to bridge the 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 eye tracking, 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 increasing availability of easy-to-use, high-resolution eye trackers (with temporal sampling rates typically between 100 and 1,000 Hz) that do not

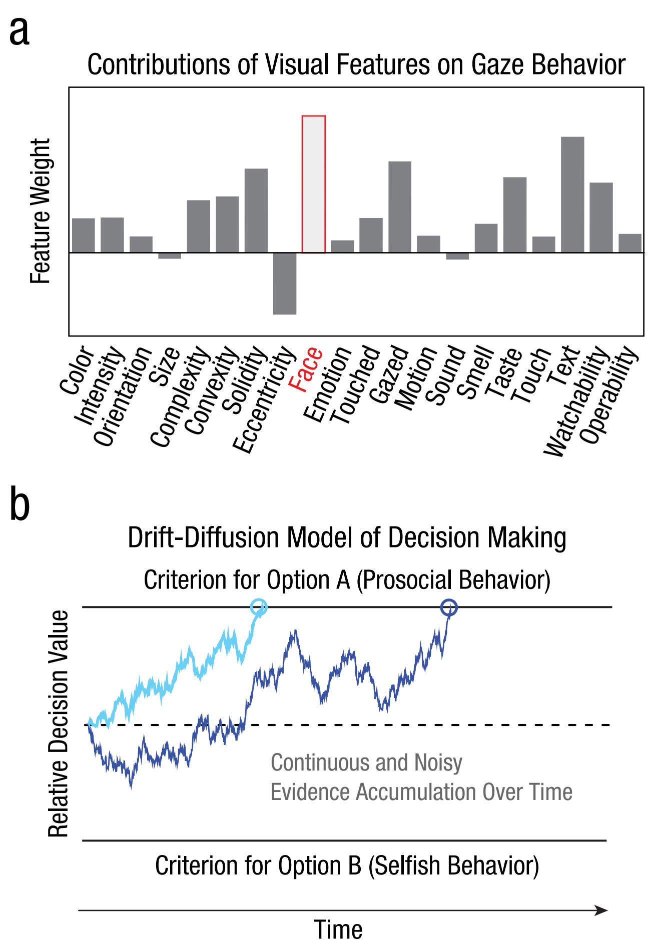

Fig. 2. Tools for extracting dimensions or features from visual stimuli that correlate with specific social judgments or social behavior. Panel (a) illustrates the results of a model-based analysis of eye-tracking data that yields saliency weights for specific features of visual stimuli. In the current 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. Panel (b) illustrates a drift-diffusion model as applied to altruistic choices. The curves (blue lines) plot the relative decision value in favor of one or the other behavioral option (in this example, to act either 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 enough evidence has accumulated and one of the decision thresholds (signified by the solid black lines at the top and bottom of the graph) is reached. Critically, how much attention is paid to choice options or their relevant features can bias the evolution of the curve. In this framework, attention (e.g., as measured via gaze behavior) can bias the decision in favor of prosocial behavior, as illustrated by the fact that the light blue line reaches the upper decision barrier earlier than the dark blue line. Panel (a) was adapted from "Atypical Visual Saliency in Autism Spectrum Disorder Quantified Through Model-Based Eye Tracking," by S. Wang, M. Jiang, X. M. Duchesne, D. P. Kennedy, R. Adolphs, and Q. Zhao, 2015, *Neuron, 88*, p. 611 (Fig. 5). Copyright 2015 by Cell Press.

require head mounting, collecting such data has become commonplace.

However, these data can also be analyzed with more sophisticated approaches. Thus, a second type of tool involves analyses that can map eye-tracking data onto psychologically meaningful attributes or dimensions. As one example, a linear classifier using a machine-learning algorithm was trained on fixations that people made to objects and faces in complex visual scenes, and the resultant model was then tested on a holdout data set (i.e., data that the algorithm hadn't seen before and that were not used for training). The net result produced a fairly detailed inventory of the relative weight that various attributes of visual scenes exert on visual attention—that is, their *visual saliency* (Xu, Jiang, Wang, Kankanhalli, & Zhao, 2014; see Fig. 2a). Interestingly, this analysis was conducted 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 using a big regression model estimating 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 a person's face), whereas others do not.

The third type of tool can take the results from the above two tools and use them to predict behavioral choices (see Fig. 2b for an example). This set of tools comprises models that convert where we look and what we attend to into decisions. One class of models accumulates sensory evidence over time; among 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 with 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 a choice-relevant feature particular to that option) contributes to noisy evidence accumulation over time. 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 eye-tracking 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 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 should bias our social judgments of and behavior toward the person.

A final set of tools probes the dimensions or features of faces more directly by manipulating them. Width-to-height ratio, skin color, or indeed any configuration that reliably correlates with a social judgment can be parametrically manipulated in computer-generated faces. Another approach uses a 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 a 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 (Adolphs, Nummenmaa, Todorov, & Haxby, 2016). These approaches complement the above set, and all of these tools taken together allow us to investigate how facial features relate to prosocial behaviors with both a broad, data-driven survey and more focused hypotheses.

Future Directions

The framework we have sketched suggests several directions for future research. First and most obviously, it motivates specific hypotheses about the mediating psychological processes (and their neural mechanisms) that link attention toward faces, on the one hand, to aspects of prosocial behavior, on the other hand. To test these hypotheses, one would need to go from focusing on the face (e.g., with eye-tracking studies) to focusing on the behavior (e.g., with behavioral economics studies) and incorporate the data generated into quantitative models (e.g., DDMs, machine-learning analyses of eye-tracking data). Second, it offers a sensitive and quantitative assessment that may not only reveal individual differences in these processes but also help in the diagnosis of psychiatric disorders. For instance, the Wang et al. (2015) study highlighted above used model-based eye tracking to investigate how people with autism view stimuli such as faces in an unusual way. Third, although 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. Fourth, future studies might explore the role of social processes such as empathy (which reliably predicts prosocial acts; Tusche, Böckler, Kanske, Trautwein, & Singer, 2016) as mediating factors that link attention to facial cues (e.g., related to distress) to subsequent helping. Finally, our framework suggests some speculative interventions for increasing prosocial behavior (Zaki & Cikara, 2015): If we can manipulate how people look at each other, we might be able to influence how they behave toward one another.

Recommended Reading

Adolphs, R. (2008). Fear, faces, and the human amygdala. *Current Opinion in Neurobiology*, *18*, 166–172. An article that reviews neuroscientific evidence on the role of the amygdala in face processing and fear perception, and discusses the automatic and nonconscious aspects of these processes.

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

Maestripieri, D., Henry, A., & Nickels, N. (2017). (See References). A clearly written, user-friendly review of prosocial biases (in both real-world and laboratory settings) in favor of attractive faces from various research disciplines.

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki. (2015). (See References). A comprehensible, highly accessible overview of the dimensions and features of faces that guide our social judgments about them.

Acknowledgments

We thank Juri Minxha and Shuo Wang for helpful comments on the manuscript, and Anna Skomorovsky for help with Figure 2.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

Research reported in this article was funded by a Conte Center grant to both authors from National Institute of Mental Health.

References

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, *433*, 68–72.

Adolphs, R., Nummenmaa, L., Todorov, A., & Haxby, J. V. (2016). Data-driven approaches in the investigation of social perception. *Philosophical transactions of the Royal Society of London, Series B: Biological Sciences*, *371*, Article 20150367. doi:10.1098/rstb.2015.0367

Adolphs, R., Tranel, D., & Damasio, A. R. (1998). The human amygdala in social judgment. *Nature*, *393*, 470–474.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, *372*, 669–672.

Bonnefon, J. F., Hopfensitz, A., & De Neys, W. (2013). The modular nature of trustworthiness detection. *Journal of Experimental Psychology: General*, *142*, 143–150.

Bzdok, D., Langner, R., Hoffstaedter, F., Turetsky, B. I., Zilles, K., & Eickhoff, S. B. (2012). The modular neuroarchitecture of social judgments on faces. *Cerebral Cortex*, *22*, 951–961.

DeBruine, L. (2011). Opposite-sex siblings decrease attraction, but not prosocial attributions, to self-resembling opposite-sex faces. *Proceedings of the National Academy of Sciences, USA*, *108*, 11710–11714.

Freeman, J. B., & Johnson, K. L. (2016). More than meets the eye: Split-second social perception. *Trends in Cognitive Sciences*, *20*, 362–374.

Freiwald, W. A., Tsao, D. Y., & Livingstone, M. S. (2009). A face feature space in the macaque temporal lobe. *Nature Neuroscience*, *12*, 1187–1196.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*, 223–233.

Hutcherson, C. A., Bushong, B., & Rangel, A. (2015). A neurocomputational model of altruistic choice and its implications. *Neuron*, *87*, 451–462. doi:10.1016/j.neuron.2015.06.031

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*, 1292–1298.

Krupp, D. B., DeBruine, L. M., & Barclay, P. (2008). A cue of kinship promotes cooperation for the public good. *Evolution & Human Behavior*, *29*, 49–55.

Landry, C., Lange, A., List, J. A., Price, M. K., & Rupp, N. (2006). Toward an understanding of the economics of charity: Evidence from a field experiment. *Quarterly Journal of Economics*, *121*, 747–782.

Maestripieri, D., Henry, A., & Nickels, N. (2017). Explaining financial and prosocial biases in favor of attractive people: Interdisciplinary perspectives from economics, social psychology, and evolutionary psychology. *Behavioral & Brain Sciences*, *40*, e19. doi:10.1017/S0140525X16000340

Marsh, A. (2016). Neural, cognitive and evolutionary foundations of human altruism. *Cognitive Science*, *7*, 59–71.

Marsh, A. A., Kozak, M. N., & Ambady, N. (2007). Accurate identification of fear facial expressions predicts prosocial behavior. *Emotion*, *7*, 239–251. doi:10.1037/1528-3542.7.2.239

Marsh, A. A., Stoycos, S. A., Brethel-Haurwitz, K. M., Robinson, P., VanMeter, J. W., & Cardinale, E. M. (2014). Neural and cognitive characteristics of extraordinary altruists. *Proceedings of the National Academy of Sciences, USA*, *111*, 15036–15041. doi:10.1073/pnas.1408440111

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

Raihani, N. J., & Smith, S. (2015). Competitive helping in online giving. *Current Biology*, *25*, 1–4.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873–922.

Rosenblat, T. S. (2008). The beauty premium: Physical attractiveness and gender in dictator games. *Negotiation Journal*, *24*, 465–481.

Solnick, S. J., & Schweitzer, M. E. (1999). The influence of physical attractiveness and gender on ultimatum game decisions. *Organizational Behavior and Human Decision processes*, *79*, 199–215.

Spunt, R., Ellsworth, E., & Adolphs, R. (2016). The neural basis of understanding the expression of the emotions in man and animals. *Social Cognitive and Affective Neuroscience*, *12*, 95–105.

Srinivasan, R., Golomb, J. D., & Martinez, A. M. (2016). A neural basis of facial action recognition in humans. *The Journal of Neuroscience*, *36*, 4434–4442.

Stirrat, M., & Perrett, D. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological Science*, *21*, 349–354.

Stirrat, M., & Perrett, D. I. (2012). Face structure predicts cooperation: Men with wider faces are more generous to their in-group when out-group competition is salient. *Psychological Science*, *23*, 718–722. doi:10.1177/0956797611435133

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, *66*, 519–545.

Tusche, A., Böckler, A., Kanske, P., Trautwein, F. M., & Singer, T. (2016). Decoding the charitable brain: Empathy, perspective taking, and attention shifts differentially predict altruistic giving. *The Journal of Neuroscience*, *36*, 4719–4732. doi:10.1523/JNEUROSCI.3392-15.2016

Wang, S., Jiang, M., Duchesne, X. M., Kennedy, D. P., Adolphs, R., & Zhao, Q. (2015). Atypical visual saliency in autism spectrum disorder quantified through model-based eye tracking. *Neuron*, *88*, 604–616.

Winston, J. S., Strange, B. A., O'Doherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience*, *5*, 277–283.

Xu, J., Jiang, M., Wang, S., Kankanhalli, M. S., & Zhao, Q. (2014). Predicting human gaze beyond pixels. *Journal of Vision*, *14*, Article 28. doi:10.1167/14.1.28

Zaatari, D., Palestis, B. G., & Trivers, R. (2009). Fluctuating asymmetry of responders affects offers in the ultimatum game oppositely according to attractiveness or need as perceived by proposers. *Ethology*, *115*, 627–632.

Zaki, J., & Cikara, M. (2015). Addressing empathic failures. *Current Directions in Psychological Science*, *24*, 471–476.