

1

Social Cognition, the Amygdala, and Autism

Ralph Adolphs

California Institute of Technology, USA

1.1 Three Broad Themes

At the outset, there are three broad themes that are important to consider that will guide the rest of this chapter. These are that (1) social cognition has enabling, or antecedent conditions; (2) the social world is complex; and (3) any specific method has fundamental limitations. We will discuss all these points with a focus on face processing, and through examples of findings in a psychiatric disease, autism, and in cognitive neuroscience, with a focus on the amygdala. Each of these three points suggests important ways forward, which we will discuss in further detail.

To help frame the discussion, we begin with a brief introduction to autism; we discuss the amygdala further below. Autism is a psychiatric disorder recognized since the 1940s, when Kanner and Asperger contemporaneously identified the disease in children (Kanner, 1943). It is a disease that arises early in life, and remains pervasive throughout life. Although it is currently diagnosed around age 3, there are precursors to it that already predict whether a child will develop autism or not. Autism is highly heritable, although no single gene accounts for a large percentage of autism; instead the disease arises from polymorphisms across many genes, each typically contributing only a very small effect size in isolation. These genes in turn code for protein products that influence many aspects of brain development and function, and in particular aspects of how neurons make and maintain synaptic connections with one another. Abnormal connectivity in the brain is currently one leading hypothesis for an intermediate phenotype that accounts for a substantial fraction of autism (Geschwind & Levitt, 2007). This

abnormal connectivity in turn causes abnormal brain function that manifests as a particular profile of abilities and disabilities – the ones used to diagnose the disease, which currently can be diagnosed only on behavioral criteria, not by a medical or genetic test of some kind.

In the psychiatric reference book used to diagnose disorders (the *Diagnostic and Statistical Manual*, DSM), autism was diagnosed as featuring impairments in three domains: social interaction, language, and stereotyped and repetitive behaviors. The first two are related, and have become fused in the transition from DSM-IV to DSM-V. The third is a somewhat heterogeneous category of impairments, including not only repetitive behaviors but also rigidity, and exceptional focus and attention to highly specific objects or topics. It has long been recognized that autism is a spectrum, and so it is often referred to as autism spectrum disorder (ASD), and it covers a very wide range from high-functioning individuals who have PhDs and whose primary complaint is skill in social interactions, to low-functioning individuals who are mentally retarded and mute. It remains an open question of considerable interest whether the processing deficits and behaviors seen in autism are truly continuous with the psychiatrically healthy population, and whether there might be subtypes of autism. It is hoped that research on the themes described below could help to answer these questions.

1.1.1 Antecedent Causes to Social Cognition

The first theme, that social cognition has antecedent conditions, is fairly obvious once we think about it. Social cognition does not emerge out of nowhere. It develops; it is caused by other processes; and it requires embedding in many other psychological processes in order to generate cognition and social behavior.

Perhaps the two most investigated antecedent conditions for social cognition are attention and motivation. Attention has long been noted to be critical for filtering sensory information, and could thus be thought of simply as a filter that determines sensory inputs, on which subsequent social cognition might be based. Thus, if we pay attention to somebody's face, we are able to make judgments about the emotion expressed on the face. If we do not pay attention to the face, we are unable to make such judgments (or make them much more poorly). But attention is much more than merely a filter, and should probably be thought of as an active seeking out of socially relevant information. We explore the visual world with our eyes, for instance, sampling relevant features as we make fixations onto them. Indeed, eyetracking has often been used to measure (overt) visual attention. This more active, instrumental view of attention of course raises a next question: so how do we decide where to attend in the first place? Presumably the value, salience, and interest of particular features of stimuli motivate us to pay attention to them. Thus, motivation is another key antecedent process that guides social cognition, together with attention.

Motivation can be thought of simply as that which causes instrumental behavior. Insofar as visual attention can be thought of as instrumental behavior, motivation can cause visual attention. An example would be top-down visual search, as when we are trying to find a person in a crowd. Conversely, it is also likely that attention influences motivation, since it is well known that our attention to stimuli influences both our preferences and choices. This particular association has been quantified with models such as drift-diffusion models, which model the accumulation of evidence that can cause motivation and choice. For instance, the more we look at a particular face, the more we are inclined to choose it as the preferred one, in two-alternative choice tasks with similar faces (Shimojo, Simion, Shimojo, & Scheier, 2003).

There are several specific factors that have been identified that contribute to motivation, and hence to attention. Perhaps the clearest one, and the one best studied in the laboratory, is reward value. If we find a particular feature rewarding, or predictive of reward, we will be motivated to attend there. This would be the simplest kind of explanation to account for why we like to thumb through magazines that have lots of pictures of people: images of people are intrinsically rewarding, and our attention is captured by them. But there are also other factors that can influence attention and motivation: attention can be captured by low-level saliency, such as the distinctiveness of a stimulus, and this in turn can drive motivation. We are also motivated to seek out information, even when it is not yet known whether that would lead to reward, and even when it is not distinctive. Reward value, saliency, and information are thus at least three factors that could in turn drive attention and motivation (Gottlieb, Hayhoe, Hikosaka, & Rangel, 2014), which in turn drive social behavior.

There is evidence to support the operation of all three factors with regard to face processing. Faces and other visual social stimuli are rewarding (Deaner, Khera, & Platt, 2005), and this rewarding property just of images of faces seems to be diminished (relative to other rewards, such as money) in people with autism (Lin, Rangel, & Adolphs, 2012). Their saliency is evident from the efficiency with which they can be detected in visual search, again an aspect that is impaired in people with autism (Wang et al., 2014), although the impairment in autism appears to be broader than just for faces (Wang et al., 2015). Finally, the information content of regions of the face drives how we attend to those regions. An interesting cross-cultural finding is that Asian observers tend to look more at the eyes in faces and less at the mouth than do Caucasian observers. A presumptive explanation for this is that the mouth carries less information in Asian people, because of cultural display rules that lead to reduced emotional expression around the mouth (Caldara, 2017).

Motivation and attention to social stimuli are thought to be dysfunctional in autism. One highly influential hypothesis about autism proposes that infants and children with autism do not find social stimuli (other people, faces) rewarding, and so are not motivated to attend to them (Chevallier, Kohls,

Troiani, Brodtkin, & Schultz, 2012). The developmental consequence of this deficit could then translate into social cognition difficulties later in life: if you do not attend to faces, you will not process faces as often, and consequently your brain will not develop expertise with faces, as it does in typically developing individuals. There is recent evidence that coarse mechanisms for attending to face-like configurations of visual stimuli may be present already in the womb: fetuses orient preferentially to lights in the configuration of eyes and mouth, when these are projected onto the abdomen of the mother (bright light can penetrate into the womb) (Reid et al., 2017).

There is a final important point to make. It is usually assumed that motivation and attention are domain-general processes that come into play at the front-end, so to speak, and that the apparent domain specificity of social cognition arises from subsequent mechanisms. But as we noted, motivation and attention can themselves exhibit selectivity for certain stimuli or features, and so can play a role both in the contemporaneous selective processing of social stimuli and in the development of domain-specific processes through experience (Spunt & Adolphs, 2017). It is even possible that attentional and motivational processes are sufficient to produce apparent category selectivity, if they amount to an intelligent enough filtering mechanism. For instance, if one combined attention to certain coarse features (the triangular configuration of eyes and mouth), and certain statistically specified locations in space (e.g., usually in the upper visual field, or foveal), and certain conjunctions of context (e.g., faces and voices), cells responding to such simple cues and their conjunctions might, in the aggregate, result in selective processing of faces.

1.1.2 The Social World is Complex

The second theme mentioned above is that realistic social stimuli are inherently complex. Other people, if we consider them as stimuli for a moment, are multimodal, moving objects with many features and attributes that all need to be processed together. They also occur in context, often involve substantial memory, and engage cognitive and behavioral processes that are typically bidirectionally interactive. Even just an image of an isolated face is complex, which is why it has been difficult to design computer vision algorithms to recognize faces. Many different features, and their relationships amongst one another, need to be represented in a flexible, viewpoint-invariant way, and need to be linked rapidly to the retrieval of often large amounts of semantic knowledge about the person whose face we are seeing.

The inherent complexity of social stimuli has typically been dealt with in the laboratory by using vastly impoverished stimuli, since these are easier to analyze and control. However, this is no longer necessary, since it is possible now to collect large amounts of data quickly, and to construct computational models that analyze such data. Some examples of this will be presented below, but

it is actually a rather common emerging theme in social neuroscience (see Adolphs, Nummenmaa, Todorov, & Haxby, 2016).

One way that the brain deals with the complexity of social stimuli is by representing them in a space with much lower dimensions. The identity of familiar individuals may be represented efficiently in a space with perhaps as few as 50 dimensions, and can be decoded from small ensembles of neurons, at least in experiments with monkeys (Chang & Tsao, 2017). More relevantly here, the social attributions that we make about people from their faces – their intentions, emotions, potential threat, and so forth – are likely represented in a space with only a few dimensions. Psychologists who study the impressions we glean from faces have identified three broad dimensions that account for much of variance in our attributions: attractiveness, dominance, and valence (or trustworthiness). There is considerable consensus, at least within a given culture, in the social attributions that we make from faces, and we are able to make them surprisingly rapidly, with less than 100 ms viewing time. Many of the core attributions are already seen in infants. It is an intriguing and very important general fact that we tend to be much more confident of our social attributions than we ought to be: we make the social judgments automatically and quickly, but they reflect more of our biases and stereotypes than providing accuracy. Alex Todorov's book, *Face Value*, provides a nice review of these effects (Todorov, 2017).

Two final sources of complexity are context and interaction. The social judgments that we make about other people depend critically on context, and in the real world involve interactions. Studying this dynamic and situated aspect of social cognition has been difficult and typically overlooked (Przyrembel, Smallwood, Pauen, & Singer, 2012), but there is now considerable interest in interactive experimental protocols, some with face-to-face encounters between people, others using virtual reality. These will be important directions for future development, conceptually, methodologically, and also in terms of the analysis tools.

1.1.3 Comparing Between Methods

The third and final broad theme of this chapter is that one must make comparisons across multiple approaches. No single approach will suffice, since each approach has limitations and shortcomings. Thus, the strongest eventual syntheses will come from studies that combine methods, or even species. Examples would be studies that use the same stimuli, and ask the same question, with electrophysiology and fMRI; or that ask parallel questions in monkeys and in humans; or that use correlational methods like fMRI as well as more causal methods like TMS or lesions. Of course, achieving this in a paper from a single laboratory is typically impossible. This highlights the need for collaborations as well. Ultimately, we want social neuroscience to be a cumulative science in

which multiple data points can build toward a convincing story, not isolated snippets that are difficult to compare.

It is worth briefly noting the major limitations with some of the most popular methods. It is well known that functional neuroimaging has clear advantages and disadvantages, for instance. Its strengths are its noninvasive nature, and whole-brain field-of-view. Limitations are the typically very small effect sizes and indirect nature of the primary measure (changes in magnetic susceptibility due to changes in blood oxygenation), artificial environment, modest spatio-temporal resolution, and correlational nature of the conclusions that are obtained (although there are methods that involve causal modeling as well). While electrophysiological measures such as EEG have similar limitations (but much better temporal resolution), invasive intracranial recordings in surgical patients provide the best spatio-temporal resolution – we give an example at the end of this chapter. Yet all these measures are primarily correlational in nature (although causal inferences can be derived from them with some effort), emphasizing the importance of perturbative approaches, such as TMS or lesion studies (which we also review below). The most compelling conclusions are ones that can be drawn from multiple approaches.

1.2 Impaired Attention to Eyes in Faces Following Human Amygdala Lesions

The example set of studies from our laboratory emphasize the first of the above three broad themes: the critical role of attention in social cognition. The story is particularly relevant, because it shows how an initially rather complex-seeming, and unexplained, specific deficit in one aspect of social perception (an inability to recognize fear in faces) could actually be explained, and even experimentally “cured,” through understanding attention.

The story begins with a famous patient, a woman named S.M., whom we have studied over several decades and who has provided the field of affective neuroscience with a wealth of insights about the necessary role of the amygdala in human social cognition and behavior (see Feinstein, Adolphs, & Tranel, 2016 for review) (Figure 1.1). S.M. has Urbach-Wiethe syndrome, an extremely rare genetic disease that results from deletions or mutations in the gene coding for extracellular matrix protein 1, a structural protein that is expressed not only in the brain but in many other organs. This disease, for reasons unknown, causes calcifications and lesions in the medial temporal lobe in a subset of patients (Hamada et al., 2002; Hofer, 1973). In S.M.’s case, it resulted in very focal and complete lesions of the amygdala, on both sides of the brain. The consequences of this in S.M.’s life have been profound: she does not seem to experience fear at all, and thus exhibits behaviors that have often put her at extreme risk (Feinstein, Adolphs, Damasio, & Tranel, 2011).

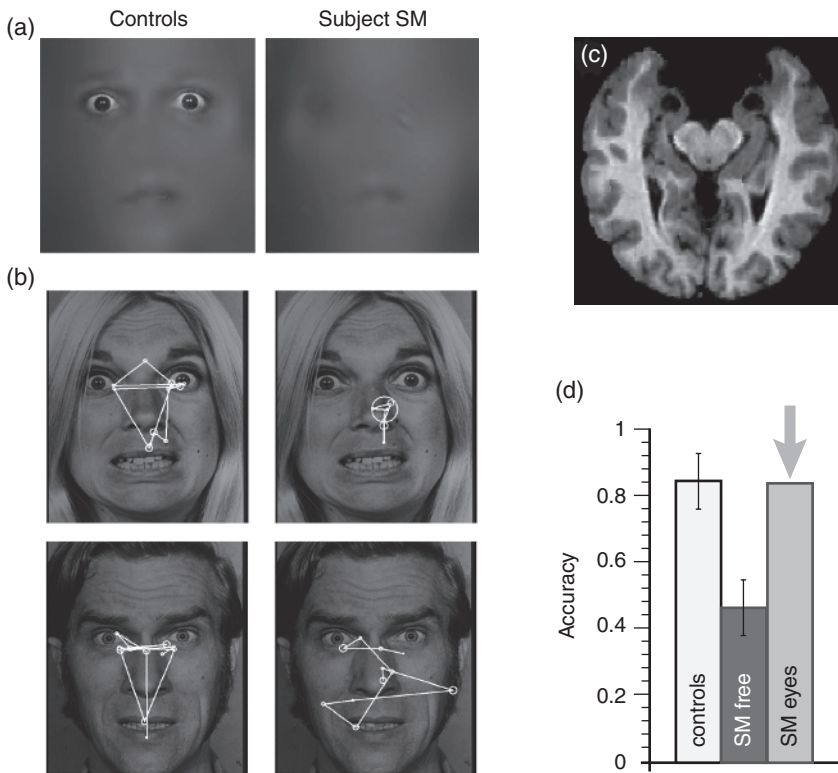


Figure 1.1 The brain and face processing in patient S.M. Bilateral amygdala lesions impair the use of the eyes and gaze to the eyes during emotion judgment. (A) A patient with bilateral damage to the amygdala made significantly less use of information from the eye region of faces when judging emotion. (B) While looking at whole faces, the patient (right column of images) exhibited abnormal face gaze, making far fewer fixations to the eyes than did controls (left column of images). This was observed across emotions (free viewing, emotion judgment, gender discrimination). (C) MRI scan of the patient's brain, whose lesion was relatively restricted to the entire amygdala, a very rare lesion in humans. The two round black regions near the top middle of the image are the lesioned amygdalae. (D) When the subject was instructed to look at the eyes ("SM eyes") in a whole face, she could do this, resulting in a remarkable recovery in ability to recognize the facial expression of fear. The findings show that an apparent role for the amygdala in processing fearful facial expressions is in fact more abstract, and involves the detection and attentional direction onto features that are socially informative. *Source:* © Ralph Adolphs.

It is important to say a few words about the amygdala and the human lesion cases here. The amygdala has long been implicated in fear, and there is substantial evidence across animal species including humans that it is necessary for many aspects of fear processing (Amaral & Adolphs, 2016), even though its role in the conscious experience of fear remains debated, especially in animals

(LeDoux, 2017). Two limitations of lesion studies of the amygdala are noteworthy. First, as with all lesion studies, the loss of function observed in a lesion case does not warrant the conclusion that the lesioned structure normally causes the function. So although amygdala lesions impair many aspects of fear processing, this doesn't mean that the amygdala normally implements those aspects of fear processing. Luckily, there is considerable evidence from other approaches that indeed does support that conclusion. Second, the amygdala is a complex structure consisting of a dozen different nuclei with further intermingled cell populations that subserve different functions. Lesions and fMRI thus have inadequate spatial resolution to resolve these populations, an issue that requires techniques like optogenetics, in which genetically targeted populations of cells can be manipulated. Plenty of those optogenetic studies have also been done now, and while they provide much more detail to the story, they largely support the conclusion that the amygdala participates in processing threat-related stimuli, although it also participates in processing rewarding stimuli. The modern-day conclusions are thus that the amygdala contains cell populations that implement functions that contribute to fear and anxiety. It also contains other cell populations that have different functions, and there are also other structures in the brain that participate in processing fear and anxiety. Whether a unitary function of some kind can be ascribed to the amygdala remains unclear, but when this has been attempted, functions related to social cognition have almost always emerged (Adolphs, 2010; Rutishauser, Mamelak, & Adolphs, 2015).

Across a large number of experiments, it was found that S.M. is selectively impaired in her ability to recognize fear from facial expressions. Although her basic vision is normal, and although she can discriminate all faces, even fear faces, normally, she fails to be able to recognize that a facial expression of fear signals the emotion fear (Adolphs, Tranel, Damasio, & Damasio, 1994). This deficit was subsequently discovered to be correlated with an inability to make use of the eye region of faces (Adolphs et al., 2005). To show this, we used a technique called "bubbles" in which participants were shown small, random pieces of a whole face and asked to recognize the emotion. Such a task, across many trials, can give us a "classification image" that shows which regions of the face carry discriminative information that allows viewers to classify them as fear or another emotion. In S.M.'s case, she had a very specific impairment on this task: she failed to make use of information from the eye region of the face. This made a lot of sense, since the eye region is normally the region of the face that is most informative about fear: wide eyes signal fear (Smith, Cottrell, Gosselin, & Schyns, 2005). So an inability to use this information from the eyes should result in impaired fear recognition, providing a mechanistic explanation for why S.M. was impaired in recognizing fear.

This finding still left two possible hypotheses. One hypothesis would state that S.M. looks at people's faces normally, and so has available at the level of the

retina exactly the same information that healthy people do when she looks at fear faces. Her impaired ability to utilize information from the eye region of faces in order to recognize fear, then, would be traced to a mechanism that depends on the amygdala. The amygdala would be necessary for some further processes that allow the brain to know that wide eyes signal fear.

A second hypothesis, however, would be that S.M. does not even look normally at the face stimuli in our experiment. That is, she might fixate faces in unusual ways, and thus might indeed not have available, at the level of the retina, the same information that healthy individuals do when they look at faces. To distinguish between these two possibilities, we used eyetracking to measure how S.M. looks at faces.

We found that S.M. indeed does not fixate faces normally. Often, she simply stares at the center of the image, not exploring it with her eyes. When she does move her eyes, she does not preferentially look at the eyes in faces, unlike healthy individuals. This finding thus provides a compelling mechanistic explanation of why S.M. is impaired in recognizing fear in faces. Normally, people look at the eyes in our face stimuli, and wide eyes signal fear. However, S.M. fails to look at the eyes in faces, and thus is unable to use information from the eye region of the face to tell her that the face expresses fear.

This story is particularly nice because it makes some testable further predictions. If true, it should be possible to help S.M. to recognize fear in faces. We could simply instruct her to look at faces the way that healthy people look at faces: fixate the eyes in faces. Would this improve her impaired fear recognition? When we did the experiment, we indeed found that it did. Unfortunately, the improvement only lasted the duration of the experiment. Without an explicit instruction to fixate the eyes in faces, S.M. would always revert back to not fixating the eyes, and to showing impaired recognition of fear.

This set of studies thus illustrates the important role of attention in social perception. It also raises the question whether we might find similar results in some other clinical populations that have difficulties in social cognition. One such population are people with autism, who also report difficulties figuring out how other people feel, and who are also often described as making poor eye contact. We turn to this clinical population next.

1.3 Atypical Visual Attention in People with Autism

The two antecedent processes that we mentioned as enabling social cognition have both been reported to be impaired in people with autism, and according to some hypotheses are thought to be responsible for the development of social difficulties in autism (Chevallier et al., 2012). It is known that people with autism fixate faces in unusual ways (Pelphrey et al., 2002) and it has also been reported that people with autism do not find pictures of faces normally

rewarding in guiding their instrumental behavior (Lin et al., 2012). While there is the belief that these deficits in social attention and social reward are specific, or at least disproportionate, for social stimuli, establishing this specificity is still an important and open question. It is possible that there are broader deficits in attention and reward processing, for all stimuli, and it is also possible that the deficits are specific to certain domains or features of stimuli, or computations performed on them, that happen to be disproportionately important when we process faces.

Be that as it may, the unusual fixation patterns of people with autism onto faces bear some intriguing resemblance to those seen in the patient with amygdala lesions, S.M. (Figure 1.2). Like S.M., people with autism tend to look less at the eyes in faces. While the patterns are far from identical, and while there are many other differences between patients with amygdala lesions and patients with autism, this superficial similarity is one piece of support for the hypothesis that amygdala dysfunction might contribute to autism (Baron-Cohen et al., 2000). Indeed, there is now overwhelming evidence that the amygdala is abnormal in autism (mostly from histological and structural studies), but it is also clear that (a) many other brain regions are also abnormal in autism, and (b) abnormalities in the amygdala contribute to all developmental disorders (and probably many adult-onset disorders), and not just to autism (Schumann, Bauman, & Amaral, 2011). Some of the most detailed ongoing studies that examine amygdala function in autism are using large data sets (such as those from the ABIDE network; Di Martino et al., 2014) to examine functional connectivity of the amygdala with other brain structures (often from resting-state fMRI data). It may be possible to diagnose autism just from the pattern of resting-state functional brain activation, although currently the number of false positives with such approaches is still too high.

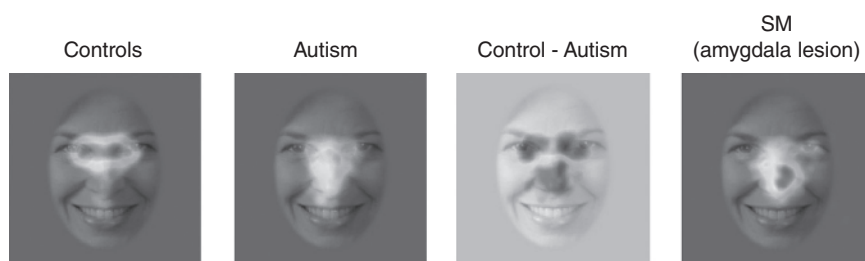


Figure 1.2 Fixations onto faces in S.M. and in people with autism show similarities. The images show data obtained from how participants fixate features from faces; hot colors denote higher density of fixations (except in the control–autism difference image, where red colors indicate that controls fixate more than autism at that location, and blue colors indicate that people with autism fixate there more than controls). Note that the images for autism and controls are obtained from groups of participants, whereas the image from S.M. is from a single individual. *Source:* Ralph Adolphs.

A further investigation by us of visual attention in autism illustrates the second of the broad themes we had outlined at the beginning of this chapter. That is the theme that social stimuli are complex, but with careful characterization can still be analyzed with sophisticated models and sufficient data. We asked the question: What features in visual stimuli capture people's visual attention, and how might this differ in people with autism? You could think of the answer to this question as producing something like a "fingerprint" that shows a profile of the weight that each visual feature has in attracting your visual attention.

In our study (Wang et al., 2015), we showed participants many different visual images. Importantly, all of these were natural scenes, and so were quite complex. They contained many different objects and features: people, animals, objects, trees, sky, background, and so forth. To capture all these different features in a computational model, we first used an automated algorithm to determine low-level visual saliency of specific regions on the image. This algorithm, developed by Christof Koch and Laurent Itti (Itti & Koch, 1998), essentially finds, in an automated way, regions that will attract visual attention because they are bright, or have high contrast, or have a particular color. So this aspect of visual saliency is relatively easy to quantify on our stimuli.

But we also wished to quantify semantic, object-based features in all our stimuli. You do not only fixate onto a region because it is bright, but also because of its meaning: whether it is showing a face, or an interesting animal, or something emotional. To characterize these semantic-level properties, we had a large number of students annotate the images (cf. Xu, Jiang, Wang, Kankanhalli, & Zhao, 2014). This produced a detailed model consisting of pixel-based (low-level saliency), object-based, and semantic-based features. We could then train this model on a subset of the eyetracking data, and ask how well it was able to predict new eyetracking data. The results of this produce a profile across all the different features, showing us how strong an effect they have on visual attention.

We then carried out exactly this same analysis in people with autism, and asked how their visual attention might be driven by different factors. Indeed, we found that there was a difference. Whereas normal controls show fixations that are driven less and less by pixel-based saliency over time, and more and more by semantic-based saliency, people with autism show much less of this effect and keep looking at low-level saliency regions in the image such as contrast and brightness. Thus, this analysis showed that visual attention in autism is characterized, at least in part, by an inability to be guided by the semantic meaning of objects in an image, and instead seems to stick to low-level cues.

There are many other examples of such a more data-driven, feature-based approach that uses relatively complex naturalistic stimuli. One type of stimulus that has become quite popular is videos or movies, which can be shown to subjects while obtaining fMRI data. Not only are these stimuli engaging and

thus capable of producing strong brain responses, but they offer a very efficient way of sampling a large range of different features within a context. Decomposing the complex stimulus of the movie into its constituent features is not trivial, but even without detailed decomposition it is possible to use such a rich stimulus to find abnormally activated brain networks in people with autism (Byrge, Dubois, Tyszka, Adolphs, & Kennedy, 2015), and then work backwards from this finding to ask what it is about the movie (e.g., which specific low-level or semantic-level features) might be most responsible for this. For instance, it was found that social awkwardness is one time-varying aspect of the movie that results in abnormal brain activation in autism (Pantelis, Byrge, Tyszka, Adolphs, & Kennedy, 2015).

1.4 Putting it All Together: Single-neuron Responses in the Amygdala

Finally, we turn to putting all three themes together, and in particular to highlighting the third of the themes, the need to use multiple methods. In this study, we used the “bubbles” method that was already introduced in Figure 1.1, we recorded from the amygdala, and we investigated amygdala responses in people with autism. The dependent measure this time, however, was not eye movements but single-neuron responses recorded from depth electrodes in the brains of neurosurgical patients.

These patients are all patients who have medically untreatable epilepsy, and whose seizures cannot be localized adequately with scalp EEG. The clinical goal is to find the region of the brain from which the seizures originate, so that this could be surgically removed. Often, the source is in the medial temporal lobe – in the amygdala or hippocampus – and resecting these tissues in a surgery called a temporal lobectomy can cure the epilepsy. But to decide exactly where the seizure originates, it is essential to be able to record the electrical activity of a seizure from electrodes in the brain, permitting a precise determination. For this clinical reason, neurosurgeons implant depth electrodes into the brains of such patients. They then spend 1–2 weeks in the hospital, with wires connected to the depth electrodes, so that one can record when a seizure occurs. During this time, the patients can also elect to participate in research studies, and one can show them stimuli and record single-neuron responses in the brain obtained through the depth electrodes. This is a very important and rare source of recordings from single neurons in the human brain, which has resulted in significant contributions in cognitive neuroscience (Fried, Rutishauser, Cerf, & Kreiman, 2014).

In our study (Rutishauser et al., 2013), we asked how single neurons in the amygdala would respond to the features of faces. Are their responses driven more by a specific part of the face, like the nose, or the eyes? To answer this

question, we used the “bubbles” technique in which small parts of faces, randomly chosen on each trial, were shown to the patients. Instead of obtaining a behavioral classification image, as we had done with patient S.M. (Figure 1.1), we now obtained a neuronal classification image, which told us the effect that each face feature had on eliciting neuronal responses from the amygdala neurons.

The result was quite striking. Whereas the control participants (a group of patients who also had epilepsy, but who did not have autism) had amygdala neurons that responded strongly to the eye region of faces, amygdala neurons in two rare patients who had autism (as well as epilepsy) showed an absence of such a response from the eyes. This finding, while limited by the very small sample size, and the unavailability of further control conditions, suggests a provocative hypothesis: neurons in the amygdala represent something like a saliency map. Normally they respond strongly to eyes in faces, but in people with autism they instead respond to the mouth. This pattern of response in amygdala neurons in patients with autism mirrors the pattern of fixations that they make onto faces, and thus suggests a mechanism that explains why people with autism do not fixate faces normally. Of course, to establish causality here, one would wish to carry out future experiments that might electrically stimulate the amygdala neurons, to see if this causes changes in fixations onto faces.

A second recent study highlights a similar convergence of approaches. In that study, we actually combined three different approaches in one paper: behavioral impairments in patients who have lesions of the amygdala, fMRI studies of the amygdala in healthy individuals, and single-unit recordings from the amygdala in neurosurgical patients. All three approaches used identical stimuli and tasks to investigate the question of which dimensions of emotional faces the amygdala might be responsible for processing. In particular, we asked whether the amygdala is involved in processing the ambiguity of the emotion, a hypothesis that Paul Whalen had suggested years earlier (Whalen, 1999) and for which there was some support (Herry et al., 2007); or whether the amygdala is involved in processing the intensity of fear in faces, which also had substantial support. We found evidence for both of these hypotheses, and could trace their origin to the presence of two largely nonoverlapping populations of cells with the single-neuron recordings: one population encoded ambiguity, the second encoded emotion intensity (Wang et al., 2017).

Taken together, the set of studies that we have reviewed here illustrate the power of approaching the study of social cognition with the three broad themes with which we began. To reiterate them briefly: we should attempt to deconstruct social cognition into its constituent, or antecedent stimuli; we should use realistic, ecologically valid social stimuli and try to quantify their full complexity in rich models; and we should strive to make comparisons across multiple methods. There is one large open domain that has not yet been well exploited: capturing this richness in features and processes in computational

models that aim to make explicit the processes. This approach has been hugely successful in learning and decision neuroscience, where sophisticated models are commonly used to estimate parameters such as the expected reward or the reward prediction error. While a few forays into the social domain have been undertaken, many of these are derivative to learning and decision-making more generically (Behrens, Hunt, Woolrich, & Rushworth, 2008). Important future topics for further development will be strategic deception (Hampton, Bossaerts, & O'Doherty, 2008) or social learning from the expertise of another person (Boorman, O'Doherty, Adolphs, & Rangel, 2013). Application of these models to the study of diseases like autism will be a major future topic in computational psychiatry.

References

- Adolphs, R. (2010). What does the amygdala contribute to social cognition? *Annals of the New York Academy of Sciences*, 1191, 42–61.
- Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P. G., & Damasio, A. (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 B: Biological Sciences*, 371. doi: 10.1098/rstb.2015.0367
- 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.
- Amaral, D. G., & Adolphs, R. (Eds.). (2016). *Living without an amygdala*. New York, NY: Guilford Press.
- Baron-Cohen, S., Ring, H. A., Bullmore, E. T., Wheelwright, S., Ashwin, C., & Williams, S. C. R. (2000). The amygdala theory of autism. *Neuroscience and Biobehavioral Reviews*, 24, 355–364.
- Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature*, 456, 245–249.
- Boorman, E., O'Doherty, J. P., Adolphs, R., & Rangel, A. (2013). The behavioral and neural mechanisms underlying the tracking of expertise. *Neuron*, 80, 1558–1570.
- Byrge, L., Dubois, J., Tyszka, J. M., Adolphs, R., & Kennedy, D. P. (2015). Idiosyncratic brain activation patterns are associated with poor social comprehension in autism. *Journal of Neuroscience*, 35, 5837–5850.
- Caldara, R. (2017). Culture reveals a flexible system for face processing. *Current Directions in Psychological Science*, 26, 249–255.
- Chang, L. J., & Tsao, D. Y. (2017). The code for facial identity in the primate brain. *Cell*, 169, 1013–1028.

- Chevallier, C., Kohls, G., Troiani, V., Brodtkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *TICS*, 16, 231–239.
- Deaner, R. O., Khera, A. V., & Platt, M. L. (2005). Monkeys pay per view: Adaptive valuation of social images by rhesus macaques. *Current Biology*, 15(6), 543–548.
- Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K.,... Milham, M. P. (2014). The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry*, 19, 659–667.
- Feinstein, J. S., Adolphs, R., Damasio, A., & Tranel, D. (2011). The human amygdala and the induction and experience of fear. *Current Biology*, 21, 34–38.
- Feinstein, J. S., Adolphs, R., & Tranel, D. (2016). A tale of survival from the world of patient S.M. In D. G. Amaral & R. Adolphs (Eds.), *Living without an amygdala*. New York, NY: Guilford Press.
- Fried, I., Rutishauser, U., Cerf, M., & Kreiman, G. (2014). *Single-neuron studies of the human brain: Probing cognition*. Cambridge, MA: MIT Press.
- Geschwind, D. H., & Levitt, P. (2007). Autism spectrum disorders: Developmental disconnection syndromes. *Current Opinion in Neurobiology*, 17, 103–111.
- Gottlieb, J., Hayhoe, M., Hikosaka, O., & Rangel, A. (2014). Attention, reward and information seeking. *Journal of Neuroscience*, 34, 15497–15504.
- Hamada, T., Irwin McLean, W. H., Ramsay, M., Ashton, G. H. S., Nanda, A., Jenkins, T.,...McGrath, J. A. (2002). Lipoid proteinosis maps to 1q21 and is caused by mutations in the extracellular matrix protein 1 gene (ECM1). *Human Molecular Genetics*, 11, 833–840.
- Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 6741–6746.
- Herry, C., Bach, D. R., Esposito, F., Di Salle, F., Perrig, W. J., Scheffler, K.,...Seifritz, E. (2007). Processing of temporal unpredictability in human and animal amygdala. *Journal of Neuroscience*, 27, 5958–5966.
- Hofer, P.-A. (1973). Urbach-Wiethe disease: A review. *Acta Dermato-Venereologica*, 53, 5–52.
- Itti, L., & Koch, C. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 1254–1259.
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2, 217–250.
- LeDoux, J. (2017). Semantics, surplus meaning, and the science of fear. *Trends in Cognitive Sciences*, 21, 303–306.
- Lin, A., Rangel, A., & Adolphs, R. (2012). Impaired learning of social compared to monetary rewards in autism. *Frontiers in Human Neuroscience*, 6, 143.
- Pantelis, P., Byrge, L., Tyszka, J. M., Adolphs, R., & Kennedy, D. (2015). A specific hypoactivation of right temporo-parietal junction/posterior superior temporal

- sulcus in response to socially awkward situations in autism. *Social Cognitive and Affective Neuroscience*, 10, 1348–1356.
- Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., & Piven, J. (2002). Visual scanning of faces in autism. *Journal of Autism and Developmental Disorders*, 32, 249–261.
- Przyrembel, M., Smallwood, J., Pauen, M., & Singer, T. (2012). Illuminating the dark matter of social neuroscience: Considering the problem of social interaction from philosophical, psychological, and neuroscientific perspectives. *Frontiers in Human Neuroscience*, 6, Article 190.
- Reid, V. M., Dunn, K., Young, R. J., Amu, J., Donovan, T., & Reissland, N. (2017). The human fetus preferentially engages with face-like visual stimuli. *Current Biology*, 27, 1825–1828.
- Rutishauser, U., Mamelak, A.N., & Adolphs, R. (2015). The primate amygdala in social perception: Insights from electrophysiological recordings and stimulation. *Trends in Neurosciences*, 38, 295–306.
- Rutishauser, U., Tudusciuc, O., Wang, S., Mamelak, A., Ross, I. B., & Adolphs, A. (2013). Single-neuron correlates of atypical face processing in autism. *Neuron*, 80, 887–899.
- Schumann, C. M., Bauman, M. D., & Amaral, D. G. (2011). Abnormal structure or function of the amygdala is a common component of neurodevelopmental disorders. *Neuropsychologia*, 49, 745–759.
- Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preferences. *Nature Neuroscience*, 6, 1317–1322.
- Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological Science*, 16, 184–189.
- Spunt, R., & Adolphs, R. (2017). A new look at domain specificity: Insights from social neuroscience. *Nature Reviews Neuroscience*, 18. doi: 10.1038/nrn.2017.76.
- Todorov, A. (2017). *Face value: The irresistible influence of first impressions*. Princeton, NJ: Princeton University Press.
- 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 eyetracking. *Neuron*, 88, 604–616.
- Wang, S., Xu, J., Jiang, M., Zhao, Q., Hurlemann, R., & Adolphs, R. (2014). Autism spectrum disorder but not amygdala lesions impairs social attention in visual search. *Neuropsychologia*, 63, 259–274.
- Wang, S., Yu, R., Tyszka, J. M., Zhen, S., Kovach, C. K., Sun, S.,...Rutishauser, U. (2017). The human amygdala parametrically encodes the intensity of specific facial emotions and their categorical ambiguity. *Nature Communications*, 8. doi: 10.1038/ncomms14821.
- Whalen, P. J. (1999). Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. *Current Directions in Psychological Science*, 7, 177–187.
- Xu, J., Jiang, M., Wang, S., Kankanhalli, M. S., & Zhao, Q. (2014). Predicting human gaze beyond pixels. *Journal of Vision*, 14, 28.