



Delusions and Prediction Error

Philip Corlett

Abstract Different empirical and theoretical traditions approach delusions differently. This chapter is about how cognitive neuroscience – the practice of studying the brain to draw conclusions about the mind – has been applied to the problem of belief and delusion. In particular, the focus is on a particular bridging theory, that of predictive coding. This theory holds that the brain contains a model of the world (and the self as an agent in that world). It uses that model to make predictions in order to adapt to the environment. Errors in those predictions can garner belief updating or be ignored, depending on how each prediction error response sustains adaptive fitness. The discussion will cover how delusions might arise and be maintained under the influence of aberrant prediction errors and what psychological and neural mechanisms of prediction error processing pertain to delusions, comparing and contrasting the theory with other prominent theories of delusions. The conclusion is that the single factor, prediction error account gives a parsimonious account of delusions that generates novel predictions about how best to treat delusions and incorporates numerous biological, clinical and phenomenological data regarding delusions.

P. Corlett (✉)
Connecticut Mental Health Center, New Haven, CT, USA
e-mail: philip.corlett@yale.edu

Keywords Aberrant prediction errors • Belief updating • Cognitive neuroscience • Delusions • Predictive coding • Prediction error processing

2.1 A MILLENNIAL CULT AND THE PSYCHOLOGY OF DELUSION

Defining, explaining and ultimately understanding delusions has proven challenging. There are many instances of people adopting and acting upon beliefs that appear delusional, despite an apparent lack of serious mental illness.

One real-world historical example may be particularly instructive.

The Chicago Tribune reported, in December 1954, that Dr. Charles Laughead (a Christian with a fascination with UFOs) foresaw the end of the world. He was speaking on behalf of Dorothy Martin, who was supposedly relaying a prophecy from extra-terrestrials from the planet Clarion. The prophecy of course did not manifest. Martin was placed in psychiatric care and charged with contributing to the delinquency of minors – the children that she and Laughead warned of the forthcoming apocalypse were so scared they had trouble sleeping.

Martin ultimately settled in Sedona, Arizona where she lived until she was 92, continuing to proselytize about aliens, but ultimately evading interaction with psychiatric services. Did Martin have delusions? What about her acolytes? Their beliefs were certainly bizarre and firm and occasionally held with some distress. There is growing appreciation that strong beliefs and delusions exist on a continuum and may be difficult to distinguish (DSM-V). This is a challenge. However, there are also opportunities. The psychology and neurobiology of belief may inform our understanding and treatment of delusions – which is a clear unmet clinical need.

Unbeknownst to Martin and Laughead, some of their followers were social psychologists from the University of Minnesota, led by Leon Festinger. The academics studied the group as the end-times loomed, resulting in a book; *When Prophecy Fails: A social psychological study of a modern group that predicted the destruction of the world* (Festinger, Riecken, & Schachter, 1956). The authors focused on cognitive dissonance; the internal discord felt from holding conflicting beliefs simultaneously [in this case was between the prophecy and real-world events (Festinger,

1962)]. People in the cult acted to reduce their dissonance. Many disavowed the apocalypse and left the group. However, some *increased* their conviction in the face of contradictory data. Martin's failed predictions were re-contextualized as actually having come to fruition (a minor earthquake did occur in California). Confounded expectations were explained away ("the aliens did come for us, but they were scared off by the crowds of press"). These sleights of mind (McKay, Langdon, & Coltheart, 2005) will be familiar to those who have spoken to patients with delusions (Garety, 1991, 1992), who can respond to challenges to their beliefs by incorporating the challenging data, and sometimes the challenger, into their delusional narrative.

Cognitive dissonance contains the kernel of the prediction error account of delusions. In brief, when beliefs about reality, prediction errors result – which are mismatches between expectation and experience. One may update one's beliefs or ignore conflicting data – minimizing the conflict. When conflict is detected inappropriately, delusions result (Adams, Stephan, Brown, Frith, & Friston, 2013; Corlett, 2015; Corlett & Fletcher, 2014; Corlett, Frith, & Fletcher, 2009a; Corlett, Honey, & Fletcher, 2007; Corlett, Honey, Krystal, & Fletcher, 2010; Corlett, Taylor, Wang, Fletcher, & Krystal, 2010; Fletcher & Frith, 2009; Gray, Feldon, Rawlins, Hemsley, & Smith, 1991)

2.2 REASONING ABOUT BELIEFS FROM BIOLOGY, PSYCHOLOGY, AND COGNITIVE NEUROSCIENCE

Delusions are challenging to study in the laboratory – the sufferer often denies any problem (Gibbs & David, 2003) and does not present to clinical attention until delusions are fully formed (Corlett et al., 2007). The neural correlates of hallucinations can be captured when people experiencing them report their experiences in a functional imaging scanner (Zmigrod, Garrison, Carr, & Simons, 2016). Delusions on the other hand, do not typically wax and wane on a timescale that lends itself to such capture. Experimental models can provide a unique window onto an otherwise inaccessible disease process (Corlett et al., 2007). Prior work has capitalized on one such drug model of delusions: ketamine; the NMDA glutamate receptor antagonist drug that transiently and reversibly engenders delusion-like ideas in healthy people (Pomarol-Clotet et al., 2006) and other animals (Honsberger, Taylor, & Corlett, 2015).

These delusions might be manifestations of aberrant prediction errors (Corlett, Taylor, et al., 2010), the mismatch between what we expect and what we experience (Rescorla & Wagner, 1972). Derived from formal learning theory to explain mechanisms of animal conditioning, prediction error (Rescorla & Wagner, 1972) is signaled by dopamine and glutamate activity in the brain (Lavin et al., 2005). It has also become a key process in theoretical models of human causal learning and belief formation (Dickinson, 2001). By minimizing prediction error we model the causal structure of our environment (Dickinson, 2001). If prediction errors occur when they ought not to, aberrant associations are formed and strengthened, culminating in delusional beliefs.

Beliefs and the Brain

The cognitive neuroscience of belief has been slow to develop. The absence of a consilient psychological theory of belief formation led the late Jerry Fodor – both a philosopher and a cognitive scientist, to assert that, whilst beliefs are among the most interesting cognitive phenomena, they are not ready to be explained in the same cognitive and neural terms as more accessible processes, such as vision (Fodor, 1975, 2000). However, there are now cognitive and neural frameworks of belief (Dickinson, 2001) amenable to quantitative analysis and applicable to studies on healthy subjects (Corlett et al., 2004) in clinical settings (Corlett, Frith, & Fletcher, 2009b; Corlett, Taylor, et al., 2010), and across species (Dickinson, 2001).

A Bridging Hypothesis: From Mind to Brain?

Associationists believe that the mind is a network of associations between ideas (Warren, 1921). It began with Plato (Plato, 350 B.C./1999). Aristotle outlined the first laws of association (Aristotle, 350 B.C./1930). John Locke described the role of improper association of ideas in mental illness (Locke, 1690/1976). David Hume added cause and effect (contiguity in time) as a law of association (Hume, 1739/2007). Pavlov explored the mechanisms of association empirically (Pavlov, 1927). His conditioning paradigms highlighted that mere contiguity is not sufficient for learning. For example, Leon Kamin discovered blocking, which involves the retardation of learning about a novel cue-outcome association when that

cue is paired with a stimulus that already predicts the outcome – the pre-trained cue blocks learning about the novel cue (Kamin, 1969). Blocking demands that the association of ideas is sensitive to surprise (McLaren & Dickinson, 1990).

Widrow and Hoff created a simple connectionist neural network of nodes, representing inputs and outputs as links between nodes (Widrow & Hoff, 1960). Those links were strengthened by reducing an error signal, the mismatch between the desired output from a given input and the output that actually occurred. A similar algorithm was proposed for animal conditioning by Rescorla and Wagner (Rescorla & Wagner, 1972); environmental stimuli induce expectations about subsequent states of the world, exciting representations of those states. Any mismatch between the expectancies and actual experience is a PE. PEs are used as teaching signals to update future expectancies about stimuli and states. Under this scheme, blocking occurs because the outcome of the compound of pre-trained and novel cues is completely predicted, by the pre-trained cue, which precludes the generation of prediction error signal and, subsequently, learning about the association between the novel cue and the outcome. Consequently, a greater magnitude PE should weaken blocking. This has been demonstrated with amphetamine administration in experimental animals (O’Tuathaigh et al., 2003), chemogenetic manipulations of cingulate cortex in rats (Yau & McNally, 2015) and optogenetic manipulation of dopamine neurons in mice (Steinberg et al., 2013). In humans, weaker blocking has been observed in patients with schizophrenia (Moran, Al-Uzri, Watson, & Reveley, 2003) and the extent to which the neural PE signal is inappropriately engaged correlates with delusion-like beliefs (Corlett & Fletcher, 2012).

Attention is also critical for associative learning. Cues that are predictably associated with important outcomes are allocated most attention, and thus more readily enter associative relationships (Mackintosh, 1975). However, stimuli with an uncertain predictive history also garner attention (Pearce & Hall, 1980). Clearly attention is important to association formation in different ways under different circumstances. One crucial circumstance involves reward prediction; stimuli garner incentive salience to the extent that they drive goal-directed action (Robinson & Berridge, 2001). We must recognize the important impact of Kapur’s perspicuous incentive salience theory of psychosis (Kapur, 2003), that delusions form as a consequence of aberrant incentive salience driven by an excess of dopamine in the ventral striatum. We note though that it was presaged

by more mechanistic theories grounded in associative learning theory (Gray et al., 1991; Miller, 1976), it did not readily explain the role of other neurotransmitters like glutamate and that the data on dopamine release capacity (Howes et al., 2009) have implicated the associative striatum (not the ventral striatum) in the genesis of psychosis. Nevertheless, there do seem to be phenomenological and empirical data linking the broad category of salient events to delusions.

How do we reconcile salience and associative learning accounts with the phenomenology and neurobiology of psychosis? Bayesian models have been invoked to explain both associative learning and psychosis (Corlett, Frith & Fletcher, 2009; Corlett et al., 2010).

2.3 BAYESIAN MINDS AND BRAINS

Thomas Bayes was a British clergyman and mathematician whose theorem was published posthumously. His is a theorem of conditional probabilities, of event A given event B, expressed as follows:

$$P(A_i | B) = \frac{P(A_i)P(B | A_i)}{\sum_{j=1}^k P(A_j)P(B | A_j)}$$

Bayes may offer a way of bridging levels of explanation – from single neurons, to groups of cells, to systems, and ultimately associative learning and belief (Clark, 2013).

Bayesian Brains Are Predictive Brains

Under this account of brain function, organisms have a brain to anticipate future situations, thus enabling survival by maximizing rewards and minimizing punishments. This is achieved computationally by making predictions and minimizing prediction errors through the hierarchical anatomy of the brain – wherein predictions are communicated in a top-down fashion, from higher to lower layers. When predictions encounter bottom-up sensory information that does not match – prediction errors are generated which are either accommodated (ignored) or assimilated (incorporated into future predictions).

Predictions originate in areas columns with less laminar differentiation (e.g. agranular cortex) and are propagated to areas with greater laminar differentiation (such as granular cortex). In the prototypical case, prediction signals originate in the deep layers (primarily layer V) and terminate in the supragranular division of dysgranular and granular regions — principally on dendrites in layer I, as well as on neurons in layers II and III.

Predictions then change the firing rates of neurons in layers I–III in anticipation of thalamic input. If the pattern of firing in a cortical column sufficiently anticipates the afferent thalamic input, there will be little or no prediction error. However, a mismatch will entail a prediction error. Some pyramidal neurons within a cortical column function as precision units that dynamically modify the gain on neurons that compute prediction error. Precision units modulate the weight of prediction errors on the basis of the relative confidence in the descending predictions compared to incoming sensory signals.

Chanes and Feldman Barrett applied this analysis more broadly to agranular cortices, notably to the limbic regions that regulate visceral control of the body’s internal milieu. Regions including the ACC, insula and thalamus may compute predictions and prediction errors and then other higher and lower cortical regions represent the specific domains being computed. We believe these sorts of models will guide prediction, inference and interpretation of neural data gathered during the formation and operation of beliefs. This arrangement may allow for the encapsulation of beliefs, without having to postulate a modular mental organization (see below).

The specific path the information takes is governed by the relative precision of the priors, as well as prediction errors (Adams et al., 2013). As Körding and Wolpert (2004) showed, the relative precision that governs how strongly we will rely on incoming data can be expressed as a linear function of priors and likelihood (probability of observing the data we see if the prior was true)¹:

$$E(\textit{Posterior}) \propto (1 - r_{\textit{reliance}}) * \textit{Prior} + r_{\textit{reliance}} * \textit{Likelihood}$$

¹We are assuming that both the distribution of priors and likelihood is Gaussian, with $\varepsilon_{\textit{prior}} \sim N(\mu, \sigma_{\textit{prior}}^2)$ and $\varepsilon_{\textit{likelihood}} \sim N(\mu, \sigma_{\textit{likelihood}}^2)$.

If the pool of our priors is both large and heterogeneous, the incoming data will play an important role in influencing our prediction. But if our priors are precise it will have a negligible role in updating.

Dopamine, serotonin and acetylcholine may code the precision of priors and prediction errors in separate hierarchies (Marshall et al., 2016). For example, acetylcholine is involved in specifying the precision of perceptual priors. However, stimulating dopamine neurons in the VTA, drives acetylcholine release in the nucleus basalis, which expands the cortical representation of sensory stimuli that coincide with the stimulation (Bao, Chan, & Merzenich, 2001). This could be a mechanism through which salient events garner greater cortical representation.

The importance of the element of surprise in the learning process has long been appreciated. C. S. Peirce coined the term abduction as a key aspect of his explanation of inference. He dissociated abduction from other mechanisms of explanation like deduction and induction (Peirce, 1931–58). Abductive inference has been used to help describe the generation of explanations for distorted perception culminating in delusions (Coltheart, Menzies, & Sutton, 2010).

Capgras syndrome is one of the most rare neurological delusions: (Capgras & Reboul-Lachaux, 1923). Here, an individual, sees his loved ones as imposters.

The confusion that accompanies living with this feeling of ongoing strangeness could become exhausting – a clear explanation, like “that’s actually not my wife” – may be protective, although far from comforting.

Kihlstrom and Hoyt (1988) have discussed the explanation process as it might pertain to misconstrued experiences. They appealed to a number of heuristics and biases to which healthy people are susceptible discussed at length by Kahneman, Slovic, and Tversky (1982).

Kihlstrom and Hoyt (1988) describe a man, walking down the street minding his own business, who suddenly and unexpectedly has an anomalous experience – he hears his name perhaps or perhaps a strange or unpleasant thought crosses his mind. All he knows is that something unusual just happened to him. The person then will initiate a search for the cause of an event; people seem to have a general propensity towards causal explanation (Michotte, 1963), and anomalous schema and incongruent events demand such explanation.

Bayesian Biases?

The Bayesian approach can be used to formalize several well-studied belief biases. For example, we know that providing people with counterarguments that undermine their beliefs is not only insufficient, but it can also ironically enhance their confidence in these beliefs – just like the Seekers in the millennial cult.

The cognitive psychology of explanation involves conscious deliberative processes; our models of delusions, perception, and learning are not committed to a requirement for conscious processing. While some associative learning effects require subjects to be aware of contingencies (Shanks & Channon, 2002), there are examples of prediction error-driven learning about stimuli that were presented subliminally (Pessiglione et al., 2008). Helmholtz considered perception to be a process of unconscious inference over alternate hypotheses about the causes of sensory stimulation (von Helmholtz, 1878/1971). Fleminger applied this reasoning to misidentification delusions, arguing that misidentification of familiar perceptual objects and scenes was due to a dysfunction in the pre-conscious specification of perceptual predictions (Fleminger, 1992) that would engender a prediction error demanding explanation.

Psychotic illnesses like schizophrenia are associated with resistance to perceptual illusions (Dima et al., 2009). It seems that in patients with delusions, perceptual priors are more flexible and prone to change, and therefore less likely to affect perception. However, extra-perceptual priors, may be stronger. A team lead by Paul Fletcher (Teufel et al., 2015) recently showed that it is this extra perceptual knowledge sphere, where recent prior experience can change subsequent processing, which is hyper-engaged in individuals prone to schizophrenia and correlates with their symptom severity.

Perhaps most relevant to the present discussion is confirmation bias (Lord et al., 1979; Nickerson, 1998), through which prior beliefs bias current decision-making. More specifically, contradictory data are ignored if they violate a cherished hypothesis. Prediction error-driven learning models have been generated that instantiate a confirmation bias. According to theoretical (Grossberg, 2000) and quantitative computational models (Doll, Jacobs, Sanfey, & Frank, 2009), confirmation biases favor learning that conforms to beliefs through the top-down influence of the frontal cortex on striatal prediction error learning. DARPP-32 and DRD2 are two striatally enriched proteins. DARPP-32 – an intracellular signaling

nexus, DRD2 a key component of dopamine D2 receptors. Both proteins are involved in prediction error signaling (Frank, Moustafa, Haughey, Curran, & Hutchison, 2007; Heyser, Fienberg, Greengard, & Gold, 2000) and involved in the top-down cancellation of striatal positive and negative prediction error signals that conflict with prior beliefs. Using a behavioral neurogenetic approach, Doll and colleagues (2009) found that genes for DARPP-32 and DRD2.

Of special interest to this discussion, confirmation bias is increased in individuals with delusions (Balzan, Delfabbro, Galletly, & Woodward, 2013). Also, DARPP-32 has been implicated in the genetic risk for schizophrenia, the effects of psychotomimetic drugs (Svenningsson et al., 2003), learning changes in instrumental contingencies (Heyser et al., 2000), as well as the functional and structural coupling between frontal cortex and striatum (Meyer-Lindenberg et al., 2007). On the other hand, Doll and colleagues (2014) found that patients with chronic schizophrenia did not show an enhanced fronto-striatal confirmation bias. Furthermore, it is possible that confirmation biases are specific to delusion contents (encapsulated) rather than a general deficit (Balzan et al., 2013).

People attribute causal significance to the most salient perceptual elements co-occurring with the event to be explained (Taylor & Fiske, 1978). In the terms of associative theories, aberrant prediction error signals might randomly increase the attentional salience of aspects of the perceptual field, leading subjects to attribute inappropriate importance to irrelevant features of the environment (Beninger & Miller, 1998; Gray, 1993, 1998a, 1998b; Gray, Feldon, Rawlins, Hemsley, & Smith, 1991; Hemsley, 1993, 2005; Kapur, 2003, 2004; Kapur, Mizrahi, & Li, 2005; Miller, 1993).

People tend to jump to conclusions, employing short cuts and heuristics. For example, people assume that the features of a causal event should resemble the features of its outcome. Unpleasant effects should have unpleasant causes. Furthermore, peoples' causal judgments tend to be greatly influenced by their *a priori* theories about causation: If someone has the idea that many unpleasant events in the outside world reflect the activities of an international terrorist conspiracy, those same terrorists may be held responsible for unpleasant internal events as well. It seems possible to appeal to an associative mechanism to explain this heuristic, a particular personal bias may be mediated by associations; the increased salience of a particular out-group may increase the propensity to form associations between that group and events in the environment.

The availability heuristic posits that the basis for judgment is the ease with which a plausible scenario can be constructed mentally. Judgments of causality are affected by the ease with which the person can imagine a path from a presumed cause to a known effect. When unpredicted events occur, the simulation process traces causal links back to prior causes. Consider a psychotic patient searching the environment for a likely cause of their anomalous experiences (Kihlstrom & Hoyt, 1988). Salient objects and events – a honk or a wave from a passing driver, perhaps a member of a minority group standing on a street corner – will inevitably draw attention and be given special weight as a likely cause of their troublesome internal events. If there is nothing perceptually salient, events may be retrieved from memory – a curse uttered in anger by a co-worker (Kihlstrom & Hoyt, 1988). If no suitable cause is generated through perception or memory, the simulation process may be invoked (Kihlstrom & Hoyt, 1988). The person may imagine possible causes and grasp the first one that comes to mind as the most likely explanation (Kihlstrom & Hoyt, 1988; Maher, 1974, 1988a, 1988b).

It is plausible that the simulation heuristic may be mediated by associative mechanisms, namely the retrieval of associative chains such that the individual can mentally trace the associations from outcome to cause. A probability tree-search mechanism mediated by prefrontal cortex may underpin this heuristic (Daw, Niv, & Dayan, 2005). Under the influence of aberrant subcortical prediction error signals, this mechanism may be invoked to account for the apparent relatedness of stimuli and events or the aberrant attentional salience of previously irrelevant background stimuli (Kihlstrom & Hoyt, 1988).

While the heuristics described so far are involved in the initial generation of a causal explanation, anchoring and adjustment might be involved in the maintenance of delusional beliefs. Many judgments begin as hypotheses – tentative conclusions that can be revised on the basis of newly acquired evidence. However, it has long been appreciated that final judgments are inordinately influenced by first impressions: The initial judgment serves as an anchor for the final one, and there is very little subsequent adjustment. The anchoring and adjustment heuristic reflects a general tendency to rely on initial or partial judgments, giving too little weight to newly acquired information. By virtue of its use, judgments of causality tend not to accommodate new information that should instigate revision. Instead, knowledge gained subsequent to the initial judgment may be distorted so as to fit the original causal theory. Subjects thus adopt suboptimal

verificationist strategies, seeking and paying special attention to information that is consistent with their hypothesis (Snyder & Swann, 1978). As many researchers will attest, when confronted with evidence that counters a cherished belief, individuals often react by challenging the evidence (Bentall, Corcoran, Howard, Blackwood, & Kinderman, 2001). Once an explanation for odd perceptual and attentional phenomena is arrived at, the patient experiences relief from anxiety. The experience of insight relief diminishes the person's subsequent motivation to question his or her original conclusions and increases resistance to contrary information. This theme is represented in Miller's (1993) associative learning based account of psychosis. He argues that arriving at a causal explanation that accounts for aberrant experiences is so rewarding/relieving that it is accompanied by a surge of dopamine (Miller, 1993). Dopamine also has impacts on the consolidation of memories (Dalley et al., 2005), and as such, an incorrect conclusion may be "stamped-in" to long-term memory by dopamine, rendering it relatively impervious to disconfirmatory evidence.

The anchoring and adjustment heuristic may relate to another prominent cognitive theory of delusional belief formation, the "jumping to conclusions bias" (Garety, Hemsley, & Wessely, 1991; Hemsley & Garety, 1986; Huq, Garety, & Hemsley, 1988). This bias was well-documented in healthy subjects (Asch, 1946; Kahneman, 2011), where individuals tend to make decisions hastily, and on the basis of little evidence. But the bulk of empirical evidence for this account comes from investigations of clinical patients' performance on probabilistic reasoning tasks; typically, participants are presented with two jars holding colored beads in different proportions. The jars are removed from view and subjects are presented with beads, drawn one at a time from a jar, and patients are then asked to predict which jar the beads are coming. Individuals with delusions tend to make a decision after only one bead (Fear & Healy, 1997; Garety et al., 1991; Huq et al., 1988; Moritz & Woodward, 2005). It is important to note that the bias is not specific to individuals with delusions (Menon, Pomarol-Clotet, McKenna, & McCarthy, 2006) and may represent a desire to end cognitive testing more rapidly or to avoid uncertain experiences (Moutoussis, Bentall, El-Deredy, & Dayan, 2011). Hence, this bias may also pertain to the defensive functions of beliefs (protecting against low self-esteem resulting from poor cognitive performance and the toxic effects of uncertainty).

The jumping to conclusions bias may represent a need for closure (McKay, Langdon, & Coltheart, 2006) in the face of aberrant prediction error signals that engender a stressful state of uncertainty about the world.

Recent behavioral and neuroimaging data suggest that as uncertainty increases, so do learning rates (Behrens, Hunt, Woolrich, & Rushworth, 2008; Pearce & Hall, 1980). When non-delusional healthy subjects jump to conclusions (updating their beliefs extensively after one trial in conditions of high uncertainty), there is hyper-connectivity between the ventrolateral prefrontal cortex and hippocampus functional magnetic resonance signals (Lee, O'Doherty, & Shimojo, 2015).

Moritz and Woodward suggest that a liberal acceptance bias might account for apparent jumping to conclusions. When only two mutually exclusive options are available (as in the beads task), individuals rapidly *accept* that the beads are coming from a particular jar, but they do not *decide* that they are to the exclusion of other possibilities (Moritz & Woodward, 2005). This account allows for over-adjustment following contradictory evidence, since although they have strongly accepted one conclusion (the beads are from one jar), they do not exclude the alternative conclusion (that the beads are coming from the other jar).

When given more than two alternatives (for example in a thematic apperception task, where participants are shown pictures and asked to rate the plausibility of particular interpretations), psychotic patients entertain a broader range of possible interpretations (rating multiple alternatives as excellent or good interpretations of a particular scenario), whereas healthy participants are more cautious and effectively narrow down the set of possible alternatives. The broadening of plausible explanations may be a manifestation of Miller's inappropriate relatedness of entities (Miller, 1976, 1993). And while it can undoubtedly minimize the rigidity with which one may hold on to an explanation, when new information arrives, at a higher, representational level, it may lead to the entertainment of implausible or absurd accounts for a particular set of circumstances.

Since anomalous perceptual and attentional experiences may be unpleasant (Maher, 1974, 1988b), it is important to consider the biases that distort causal judgments about negatively valenced events. For example, when humans make causal attributions, they tend to fall for beneficance bias, such that they internalize the cause of positive events and externally attribute negatively valenced events (Greenwald, 1980; Kaney & Bentall, 1992). Such *Lake Woebegone Effects* – where everyone is smarter and more beautiful than average – are exaggerated in patients with paranoia (Kaney & Bentall, 1992). Hence a psychotic individual seeking an explanation for their unpleasant anomalous experiences will most often look to the environment outside them, rather than say, to a dysfunction in

their own brain or body. These biases were the only types of belief afforded the status of adaptive misbeliefs by McKay and Dennett (2009). If these biases may be related to delusions, perhaps then, certain delusions could be adaptive misbeliefs.

In an fMRI study of the self-serving hindsight bias in healthy individuals, subjects silently read sentences describing positively and negatively valenced social events, then imagined the event happening to them, and finally decided the cause of the event, whether internal (was it something about you?) or external (was it something about your friend? was it something about the situation or circumstances?). Self-serving biased attributions (internal attribution of positive and external attribution of negative events) were associated with striatal activation (Blackwood et al., 2003), previously implicated in the motivational control of behavior (Robbins & Everitt, 1996), as well as in the mediation of delusions (Laruelle, Abi-Dargham, Gil, Kegeles, & Innis, 1999). Menon and colleagues (2011) showed that delusions of reference were associated with inappropriate striatal engagement during reading of sentences that were inappropriately judged to be self-related.

2.4 DELUSIONS, SELF AND OTHERS

Thus far, we have discussed beliefs in the context of individuals. However, they are constructed in a social context that involves interacting with others and engaging with their perspectives. In our theory, the brain models incoming data and minimizes prediction error (Friston & Kiebel, 2009). However, it also actively samples those data, by performing actions on the world (e.g. moving through it) (Friston, Daunizeau, Kilner, & Kiebel, 2010). By predicting (and ignoring) the sensory consequences of our actions we also model ourselves as agents that exist. And, by identifying with the top layers of the hierarchy, the conscious experience of being that self emerges (Blanke & Metzinger, 2009).

Passivity experiences – the sense that ones' actions are under external control – may arise when the predictive modeling of one's actions fails and the active sampling of sensory data becomes noisy (Stephan, Friston, & Frith, 2009). In such circumstances thoughts and actions that were self-generated are not attributed to self and, at the extremes, one no longer identifies with ones' hierarchical model of the world. On the other hand, paranoia and referential delusions may be associated with excessive responsibility, our sense of self extends to areas it should not.

Ketamine augments experience of the rubber hand illusion, the spurious sense of ownership of a prop-hand if the hand is stroked at the same time as one's own hand (Morgan et al., 2011). People on ketamine get the illusion more strongly and they experience it even in a control condition when the real and rubber hands are stroked asynchronously (Morgan et al., 2011). Patients with schizophrenia (Peled, Pressman, Geva, & Modai, 2003) and chronic ketamine abusers show the same excessive experience of the illusion, in the synchronous and asynchronous conditions (Tang et al., 2015). Activity in the right anterior insula cortex increases to the extent that individuals experience the illusion. Anil Seth and others have argued that the anterior insula is a key nexus for the PE driven inferences that guide perceptions of bodily ownership and agency (Palmer, Seth, & Hohwy, 2015; Seth, 2013; Seth, Suzuki, & Critchley, 2011). Others highlight the parietal cortex as a key locus for the illusion (Ehrsson, Holmes, & Passingham, 2005).

2.5 BLANKETS, BRAINS, BELIEFS

The “*Markov blanket*” might be one means a Bayesian brain distinguishes self and others (Friston & Frith, 2015). A Markov blanket it like a cell membrane. It shields the interior of the cell from direct exposure to the conditions outside it, but it contains sufficient information (in the form of actual and potential structures) for the cell to be influenced by and to influence those external conditions. The Markov blanket of “an animal” encloses the Markov blankets of “the organs,” which enclose the Markov blanket of “their cells,” which enclose the Markov blankets of “their nuclei,” etc. To distinguish such levels of hierarchies, Pearl used the terms “parents” and “children.” The Markov blanket for a node in a Bayesian network is the set of nodes composed of its parents, its children, and its children's other parents. The Markov blanket of a node contains all the variables that shield the node from the rest of the network. This means that the Markov blanket of a node is the only knowledge required to predict the behavior of the node.

A Markov blanket separates states into *internal* and *external*. External states are hidden (insulated) from the internal states. In other words, from the node's, or individual's perspective, the external states can be seen only indirectly by the internal states, via the Markov blanket. The internal state models (learns and make inferences about) the external, lying on the other side of the blanket.

Despite serving as a boundary, the Markov blanket may also have a role in synchronizing self with others. This occurs, for example, when we speak to another agent. In our predictive coding scheme, we adapt language comprehension to the demands of any given communicative situation, estimating the precision of our prior beliefs at a given representational level and the reliability of new inputs to that level (Friston & Frith, 2015).

In a hermeneutic setting, though, Bayesian brains do not predict each other; they predict themselves provided those predictions are enacted. The enactment of sensory (proprioceptive) predictions is a tenet of active inference, as we can minimize prediction errors by actively sampling data that conform to our predictions (Friston & Frith, 2015). This framework for communication is inherently embodied and enactive in nature.

The internal states (of each agent) and external states (their partner) – and the Markov blanket that separates them – possess something called a random dynamical attractor that mediates the synchrony (Friston, Sengupta, & Auletta, 2014). Through this attractor, the external and internal states track each other, or the states one agent occupies impose constraints on states the other can occupy. However, if the Markov blanket or attractor become dysfunctional, first rank psychotic symptoms (Schneider, 1957) may result. That is, you may hear voices from recognizable social agents that communicate with you, or believe that your thoughts, actions, and emotions have been inserted into your mind by others.

Of particular relevance is the implication of temporoparietal junction in hearing voices. According to Saxe, this is a central role in representing others' mental states through predictive coding (Koster-Hale & Saxe, 2013). Stimulating temporoparietal junction induces a "sensed presence." Taken together with theories that suggest that hallucinations and delusions arise when reality monitoring (or more accurately reality filtering) fails such that inner speech is confused with external speech (Johnson & Raye, 1981), one can see how perturbations of these inferential mechanisms could render inner speech experienced as the communicative intent of an external agent. Similarly, Fernyhough, after Vygotsky, argues that children learn language through interaction with others; this begins out loud and later when we internalize speech as thought aberrations of this process subtend an inner voice that does not belong and is rather another

agent (Jones & Fernyhough, 2007). Predictive social models may also be set awry by poor attachment (Fineberg, Steinfeld, Brewer, & Corlett, 2014).

2.6 THERAPEUTIC IMPLICATIONS

Predictive coding seems to entail learning about different contingencies: low-level contingencies, detected within a perceptual module (e.g. V1 or A1) and higher-level contingencies that involve integrating across time, space, and sensory modalities. When low-level contingency detection fails, higher-level, top-down knowledge-based contingency detection compensates – hence a stronger reliance on high-level priors as Teufel et al. (2015) observed in people at risk for psychosis. To reinforce this idea, we point to a phenomenon from social psychology – lack of personal control. Remembering a time in one’s life when one lacked control, such as preparing to skydive from an airplane, triggers a compensatory increase in illusory pattern perception like superstitious behavior and belief in conspiracy theories; there needn’t necessarily be a direct connection between uncertainty and the way in which it is compensated (Proulx, Inzlicht, & Harmon-Jones, 2012), as any belief will do. Ultimately, this conception of belief underlines our aversion to uncertainty and our preference for reasons and explanations.

Why beliefs backfire in response to challenges is not yet fully understood, however, there are models of the political polarization of beliefs in response to the same evidence that suggest the strength of priors are important. If priors are strong, polarizing effects are more likely. Personally relevant priors that contribute to self-identity are likely to be the strongest. This is not encouraging with regard to efforts to change strong beliefs.

However, there are some encouraging new data. One promising line of inquiry, with respect to vaccine beliefs, is the involvement of individuals who used to object to vaccines and have now changed their minds in engaging with others who are against vaccinations (Brendan Nyhan, *personal communication*). Many researchers agree delusions and beliefs are often grounded in personal experiences. To the credulous, personal experiences are a reliable source. Relinquishing those beliefs on the basis of others’ testimony is strongly related to the credibility of the source (Nyhan & Reifler, 2013); for example, do the individuals trying to change another’s mind have a vested reason to disagree, like professional status, roles, or affiliations? Perhaps large-scale anti-stigma educational activities

in mental health have failed because they did not employ individuals with lived experience to spread the word about mental illness (Corrigan, 2012). With regard to fixed and distressing delusional beliefs, perhaps peer-support might supplement our standard approaches to mollifying delusions. People with lived experience who have recovered from delusions or learned how to manage them might be better at helping their peers experiencing ongoing delusions. More direct methods might involve hearing the story and imagining the position of someone directly affected by the belief. This technique was tested for beliefs about transgendered individuals with success (Broockman & Kalla, 2016). With regards to the putative circuit, perhaps engaging empathy in this manner permits assimilation and belief updating rather than the discarding of prediction error.

2.7 WHAT SHAPE IS THE MIND?

There are of course other theories of beliefs and delusions. Extant cognitive neuropsychiatric (Halligan & David, 2001) explanations of delusions range from single factor (Maher, 1974), to two-factor to interactionist. The single factor account appeals to a deficit in perception; the delusion formation process being a logical consequence of such an unsettling experience (Maher, 1974, 1988a). Two-factor theorists appeal to a deficit in familiarity processing with an additional dysfunction in belief evaluation such that the unlikely explanation (“My loved one has been replaced”) is favored (Coltheart, 2010; Coltheart, Langdon, & McKay, 2007; Mendoza, 1978).

Two-factor theory is attractive in its simplicity. It derives from cognitive neuropsychology; the consideration of patients who develop delusions following brain damage (Coltheart, 2010; Coltheart & Davies, 2000; Coltheart, Langdon, & McKay, 2010). It holds that two factors are necessary for delusions; a perceptual dysfunction and a belief evaluation dysfunction. Each is attributable to separate locations of brain damage – for Capgras delusion the perceptual dysfunction may involve ventromedial prefrontal cortex damage that renders familiar faces unfamiliar. However, people with this damage do not always have Capgras (Tranel & Damasio, 1985). Coltheart and others posit that a further deficit in belief evaluation is necessary for the delusion. They suggest right dorsolateral prefrontal cortex may be the locus of the second factor (Coltheart, 2010; Coltheart & Davies, 2000; Coltheart, Langdon, et al., 2010). The logic

here is flawed; a second factor is only suggested. It would be necessitated by a double dissociation of functions (Coltheart, 2002). The data here are still consistent with a single factor: The ventromedial perceptual dysfunction could occur to a greater or lesser degree, delusions could arise in those patients who have more extensive damage and they may be absent in people with less extensive damage. Nevertheless, two factor theories emphasized the role of perception and belief in the genesis of psychotic symptoms. Updated versions of the theory implicated Bayesian mechanisms of belief evaluation in delusion formation (Coltheart, Menzies, et al., 2010) and interactive models suggest that perception and belief intersect in a Bayesian manner that may become deranged when delusions form (Young, 2008). This update moves two-factor theory nearer to PE theory.

However, PE theory challenges the strict distinction between perception and belief, and therefore the necessity for two factors to explain delusions (Powers, Kelley, & Corlett 2016). The disagreement is not about delusions per se, but rather cognitive neuropsychology more broadly, the shape of the mind, the allowable relationships between processes and how one ought to relate the mind with the brain. These may seem arcane. However, we try to explain delusions to better treat them. Understanding their component cognitive and neural mechanisms is essential.

Modularity Versus Penetrability

In *The Modularity of Mind* (1983), Fodor sketched a mental architecture comprised of modules—systems that process a single specific kind of information (Fodor, 1983). 2-factor theory demands this encapsulated modularity. Belief and perception are separate and can be damaged independently. Information flows from perception to belief and never in the opposite direction (Fotopoulou, 2014). An encapsulated perceptual system, kept separate from the influence of beliefs, could keep our beliefs grounded in the truth offered by our senses (Quine & Quine, 1951). However, a cognitively penetrable perceptual apparatus, per PE theory, may be equally adaptive, despite misperceiving and misbelieving (Johnson & Fowler, 2011; McKay & Dennett, 2009). We perceive what would need to be present in order for our sensations to make sense, not necessarily what is actually there (von Helmholtz, 1867; Hume, 1900). Predictive perception is penetrated by beliefs to the extent this minimizes overall long-term PE (Lupyan & Clark, 2015).

Ultimately, the two explanations (two-factor and predictive processing) are cast at different explanatory levels. Two-factor theory is concerned with describing cognitive architectures. Predictive processing aims to unite brain, behavioral and phenomenological data for all delusions (neurological and those that occur in schizophrenia) as well as other psychotic symptoms like hallucinations and a-motivation.

2.8 CONCLUSION

A better understanding of delusions may be achieved by taking a reductionist approach to beliefs, conceiving of them as learned associations between representations that govern perception (both internal and external) and action. Central to the process of associative belief formation is PE; the mismatch between prior expectation and current circumstances. Depending on the precision (or inverse variance) of the PE (relative to the prior), it may drive new learning – updating of the belief, or it may be disregarded. We have argued that this process of PE signaling and accommodation/assimilation may be awry in people with psychotic illnesses. In particular, we believe delusions form when PE is signaled inappropriately with high precision, such that it garners new and aberrant learning. We have described animal research that has furnished a mechanistic understanding of PE signaling in terms of underlying neurobiology; glutamatergic mechanisms underlie the specification of PE (NMDA receptors signal top-down expectancies, AMPA the feedforward error signal), and, depending on the specific hierarchy, slower neuromodulators (like dopamine, acetylcholine, serotonin, noradrenaline and oxytocin) signal precision of priors and PE. There are thus many routes through which PE can be aberrantly signaled and many heterogeneous consequences of aberrant PE. The inferences that are perturbed give rise to the specific contents of delusions (they are about other people and one's relationships to them, because these are the hardest inferences to make). We have described how such error correcting inferential mechanisms might give rise to the sense of bodily agency (the sense of being a self) and to a sense of reality more broadly. Disrupting these senses is profoundly distressing and results in psychosis. Armed with an understanding of exactly how people with delusions fare on these tasks and exactly which neural mechanisms underpin them, we will be much better placed to determine the pathophysiology underpinning delusions and to tailor treatment approaches aimed at that pathophysiology.

REFERENCES

- Adams, R. A., Stephan, K. E., Brown, H. R., Frith, C. D., & Friston, K. J. (2013). The computational anatomy of psychosis. *Frontiers in Psychiatry*, *4*, 47. <https://doi.org/10.3389/fpsy.2013.00047>
- Arbib, M., & Bota, M. (2003). Language evolution: neural homologies and neuroinformatics. *Neural Networks*, *16*(9), 1237–1260. <https://doi.org/10.1016/j.neunet.2003.08.002>
- Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *The Behavioral and Brain Sciences*, *28*(2), 105–124 discussion 125–167.
- Aristotle. (Ed.). (350 B.C./1930). *On Memory and Reminiscence* (Vol. 3). Oxford, UK: Clarendon Press.
- Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, *41*(3), 258–290.
- Balzan, R., Delfabbro, P., Galletly, C., & Woodward, T. (2013). Confirmation biases across the psychosis continuum: The contribution of hypersalient evidence-hypothesis matches. *The British Journal of Clinical Psychology/The British Psychological Society*, *52*(1), 53–69. <https://doi.org/10.1111/bjc.12000>
- Bao, S., Chan, V. T., & Merzenich, M. M. (2001). Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature*, *412*(6842), 79–83. <https://doi.org/10.1038/35083586>
- Behrens, T. E., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. (2008). Associative learning of social value. *Nature*, *456*(7219), 245–249. <https://doi.org/10.1038/nature07538>
- Beninger, R. J., & Miller, R. (1998). Dopamine D1-like receptors and reward-related incentive learning. *Neuroscience and Biobehavioral Reviews*, *22*(2), 335–345.
- Bentall, R. P., Corcoran, R., Howard, R., Blackwood, N., & Kinderman, P. (2001). Persecutory delusions: A review and theoretical integration. *Clinical Psychology Review*, *21*(8), 1143–1192.
- Blackwood, N. J., Bentall, R. P., Ffytche, D. H., Simmons, A., Murray, R. M., & Howard, R. J. (2003). Self-responsibility and the self-serving bias: An fMRI investigation of causal attributions. *NeuroImage*, *20*(2), 1076–1085. [https://doi.org/10.1016/S1053-8119\(03\)00331-8](https://doi.org/10.1016/S1053-8119(03)00331-8)
- Blanke, O., & Metzinger, T. (2009). Full-body illusions and minimal phenomenal selfhood. *Trends in Cognitive Sciences*, *13*(1), 7–13. <https://doi.org/10.1016/j.tics.2008.10.003>
- Broockman, D., & Kalla, J. (2016). Durably reducing transphobia: A field experiment on door-to-door canvassing. *Science*, *352*(6282), 220–224. <https://doi.org/10.1126/science.aad9713>

- Brown, M., & Kuperberg, G. R. (2015). A hierarchical generative framework of language processing: Linking language perception, interpretation, and production abnormalities in schizophrenia. *Frontiers in Human Neuroscience*, 9, 643. <https://doi.org/10.3389/fnhum.2015.00643>
- Capgras, J., & Reboul-Lachaux, J. (1923). L'illusion des 'sosies' dans un delirium systematisé chronique. *Bulletin de la Société clinique de médecine mentale*, 2, 6–16.
- Chawke, C., & Kanai, R. (2016). Alteration of political belief by non-invasive brain stimulation. *Frontiers in Human Neuroscience*, 9, 621.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *The Behavioral and Brain Sciences*, 36(3), 181–204. <https://doi.org/10.1017/S0140525X12000477>
- Coltheart, M. (2002). Cognitive neuropsychology. In H. Pashler & J. Wixted (Eds.), *Steven's handbook of experimental psychology* (Vol. 4, 3rd ed.). Hoboken, NJ: Wiley.
- Coltheart, M. (2010). The neuropsychology of delusions. *Annals of the New York Academy of Sciences*, 1191(1), 16–26. doi:NYAS5496 [pii] <https://doi.org/10.1111/j.1749-6632.2010.05496.x>.
- Coltheart, M., Cox, R., Sowman, P., Morgan, H., Barnier, A., Langdon, R., ... Polito, V. (2018). Belief, delusion, hypnosis, and the right dorsolateral prefrontal cortex: A transcranial magnetic stimulation study. *Cortex*. <https://doi.org/10.1016/j.cortex.2018.01.001>
- Coltheart, M., & Davies, M. (2000). *Pathologies of belief*. Oxford, UK: Blackwell.
- Coltheart, M., Langdon, R., & McKay, R. (2007). Schizophrenia and monotheistic delusions. *Schizophrenia Bulletin*, 33(3), 642–647.
- Coltheart, M., Langdon, R., & McKay, R. (2010). Delusional Belief. *Annual Review of Psychology*. <https://doi.org/10.1146/annurev.psych.121208.131622>
- Coltheart, M., Menzies, P., & Sutton, J. (2010). Abductive inference and delusional belief. *Cognitive Neuropsychiatry*, 15(1), 261–287. <https://doi.org/10.1080/13546800903439120>
- Corlett, P. R. (2015). Answering some phenomenal challenges to the prediction error model of delusions. *World Psychiatry*, 14(2), 181–183. <https://doi.org/10.1002/wps.20211>
- Corlett, P. R., Aitken, M. R. F., Dickinson, A., Shanks, D. R., Honey, G. D., Honey, R. A. E., ... Fletcher, P. C. (2004). Prediction error during retrospective revaluation of causal associations in humans: fMRI evidence in favor of an associative model of learning. *Neuron*, 44(5), 877. <https://doi.org/10.1016/j.neuron.2004.11.022>
- Corlett, P. R., & Fletcher, P. C. (2012). The neurobiology of schizotypy: Frontostriatal prediction error signal correlates with delusion-like beliefs in healthy people. *Neuropsychologia*, 50(14), 3612–3620. <https://doi.org/10.1016/j.neuropsychologia.2012.09.045>

- Corlett, P. R., & Fletcher, P. C. (2014). Computational psychiatry: A Rosetta Stone linking the brain to mental illness. *Lancet Psychiatry*, *1*, 399.
- Corlett, P. R., Frith, C. D., & Fletcher, P. C. (2009a). From drugs to deprivation: A Bayesian framework for understanding models of psychosis. *Psychopharmacology*, *206*(4), 515–530.
- Corlett, P. R., Frith, C. D., & Fletcher, P. C. (2009b). *From drugs to deprivation: A Bayesian framework for understanding models of psychosis* (Vol. 206, pp. 515–530). Berlin/Heidelberg, Germany: Springer Science & Business Media.
- Corlett, P. R., Honey, G. D., & Fletcher, P. C. (2007). From prediction error to psychosis: Ketamine as a pharmacological model of delusions. *Journal of Psychopharmacology*, *21*(3), 238–252. <https://doi.org/10.1177/0269881107077716>
- Corlett, P. R., Honey, G. D., Krystal, J. H., & Fletcher, P. C. (2010). Glutamatergic model psychoses: Prediction error, learning, and inference. *Neuropsychopharmacology*. doi:npp2010163 [pii] <https://doi.org/10.1038/npp.2010.163>
- Corlett, P. R., Taylor, J. R., Wang, X. J., Fletcher, P. C., & Krystal, J. H. (2010). Toward a neurobiology of delusions. *Progress in Neurobiology*, *92*(3), 345–369. <https://doi.org/10.1016/j.pneurobio.2010.06.007>
- Corrigan, P. W. (2012). Research and the elimination of the stigma of mental illness. *The British Journal of Psychiatry*, *201*(1), 7–8. <https://doi.org/10.1192/bjp.bp.111.103382>
- Dalley, J. W., Laane, K., Theobald, D. E., Armstrong, H. C., Corlett, P. R., Chudasama, Y., & Robbins, T. W. (2005). Time-limited modulation of appetitive Pavlovian memory by D1 and NMDA receptors in the nucleus accumbens. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(17), 6189–6194. <https://doi.org/10.1073/pnas.0502080102>
- Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704–1711.
- Dayan, P., Kakade, S., & Montague, P. R. (2000). Learning and selective attention. *Nature Neuroscience*, *3*(Suppl), 1218–1223. <https://doi.org/10.1038/81504>
- Dickinson, A. (2001). The 28th Bartlett memorial lecture causal learning: An associative analysis. *The Quarterly Journal of Experimental Psychology. B*, *54*(1), 3–25.
- Dima, D., Roiser, J. P., Dietrich, D. E., Bonnemann, C., Lanfermann, H., Emrich, H. M., & Dillo, W. (2009). Understanding why patients with schizophrenia do not perceive the hollow-mask illusion using dynamic causal modelling. *NeuroImage*, *46*(4), 1180–1186. <https://doi.org/10.1016/j.neuroimage.2009.03.033>

- Doll, B. B., Jacobs, W. J., Sanfey, A. G., & Frank, M. J. (2009). Instructional control of reinforcement learning: A behavioral and neurocomputational investigation. *Brain Research*, 1299, 74–94. <https://doi.org/10.1016/j.brainres.2009.07.007>
- Doll, B. B., Waltz, J. A., Cockburn, J., Brown, J. K., Frank, M. J., & Gold, J. M. (2014, June). Reduced susceptibility to confirmation bias in schizophrenia. *Cognitive, Affective & Behavioral Neuroscience CABN*, 14(2), 715–728.
- Ehrsson, H. H., Holmes, N. P., & Passingham, R. E. (2005). Touching a rubber hand: Feeling of body ownership is associated with activity in multisensory brain areas. *The Journal of Neuroscience*, 25(45), 10564–10573. <https://doi.org/10.1523/JNEUROSCI.0800-05.2005>
- Fear, C. F., & Healy, D. (1997). Probabilistic reasoning in obsessive-compulsive and delusional disorders. *Psychological Medicine*, 27(1), 199–208.
- Festinger, L. (1962). Cognitive dissonance. *Scientific American*, 207, 93–102.
- Festinger, L., Riecken, H. W., & Schachter, S. (1956). *When prophecy fails*. Minneapolis: University of Minnesota.
- Fineberg, S. K., Steinfeld, M., Brewer, J. A., & Corlett, P. R. (2014). A computational account of borderline personality disorder: Impaired predictive learning about self and others through bodily simulation. *Frontiers in Psychiatry*, 5, 111. <https://doi.org/10.3389/fpsy.2014.00111>
- Fleminger, S. (1992). Seeing is believing: The role of ‘preconscious’ perceptual processing in delusional misidentification. *The British Journal of Psychiatry*, 160, 293–303.
- Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: A Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience*, 10(1), 48–58.
- Fodor, J. A. (1975). *The language of thought*. New York: Crowell.
- Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.
- Fodor, J. A. (2000). *The mind Doesn't work that way*. Cambridge, MA: MIT.
- Fotopoulou, A. (2014). Time to get rid of the ‘Modular’ in neuropsychology: A unified theory of anosognosia as aberrant predictive coding. *Journal of Neuropsychology*, 8(1), 1–19. <https://doi.org/10.1111/jnp.12010>
- Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences of the United States of America*, 104(41), 16311–16316. <https://doi.org/10.1073/pnas.0706111104>
- Friston, K., & Frith, C. (2015). A duet for one. *Consciousness and Cognition*, 36, 390–405. <https://doi.org/10.1016/j.concog.2014.12.003>
- Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364(1521), 1211–1221. <https://doi.org/10.1098/rstb.2008.0300>

- Friston, K., Sengupta, B., & Auletta, G. (2014). Cognitive dynamics: From attractors to active inference. *Proceedings of the Institute of Electrical and Electronics Engineers*, 102(4), 427–445. <https://doi.org/10.1109/JPROC.2014.2306251>
- Friston, K. J., Daunizeau, J., Kilner, J., & Kiebel, S. J. (2010). Action and behavior: A free-energy formulation. *Biological Cybernetics*, 102(3), 227–260. <https://doi.org/10.1007/s00422-010-0364-z>
- Garety, P. (1991). Reasoning and delusions. *The British Journal of Psychiatry*, 14(Supplement), 14–18.
- Garety, P. A. (1992). Making sense of delusions. *Psychiatry*, 55(3), 282–291; discussion 292–286.
- Garety, P. A., Hemsley, D. R., & Wessely, S. (1991). Reasoning in deluded schizophrenic and paranoid patients. Biases in performance on a probabilistic inference task. *Journal of Nervous and Mental Disease*, 179(4), 194–201.
- Gibbs, A. A., & David, A. S. (2003). Delusion formation and insight in the context of affective disturbance. *Epidemiologia e Psichiatria Sociale*, 12(3), 167–174.
- Gray, J. A. (1993). Consciousness, schizophrenia and scientific theory. *Ciba Foundation Symposium*, 174, 263–273 discussion 273–281.
- Gray, J. A. (1998a). Abnormal contents of consciousness: The transition from automatic to controlled processing. *Advances in Neurology*, 77, 195–208; discussion 208–111.
- Gray, J. A. (1998b). Integrating schizophrenia. *Schizophrenia Bulletin*, 24(2), 249–266.
- Gray, J. A., Feldon, J., Rawlins, J. N. P., Hemsley, D., & Smith, A. D. (1991). The neuropsychology of schizophrenia. *The Behavioral and Brain Sciences*, 14, 1–84.
- Greenwald, A. G. (1980). The totalitarian ego. *American Psychologist*, 35(7), 603–618.
- Grossberg, S. (2000, July). How hallucinations may arise from brain mechanisms of learning, attention, and volition. *Journal of the International Neuropsychological Society: JINS*, 6(5), 583–592.
- Halligan, P. W., & David, A. S. (2001). Cognitive neuropsychiatry: Towards a scientific psychopathology. *Nature Reviews. Neuroscience*, 2(3), 209–215.
- Helmholtz, H. von. (1867). *Handbuch der physiologischen Optik*. Leipzig.: Voss.
- Helmholtz, H. von. (1878/1971). The facts of perception. In R. Kahl (Ed.), *Selected Writings of Herman von Helmholtz*. Middletown, CT: Wesleyan University Press.
- Hemsley, D. R. (1993). A simple (or simplistic?) cognitive model for schizophrenia. *Behaviour Research and Therapy*, 31(7), 633–645.
- Hemsley, D. R. (2005). The schizophrenic experience: Taken out of context? *Schizophrenia Bulletin*, 31(1), 43–53. <https://doi.org/10.1093/schbul/sbi003>

- Hemsley, D. R., & Garety, P. A. (1986). The formation and maintenance of delusions: A Bayesian analysis. *The British Journal of Psychiatry*, *149*, 51–56.
- Heyes, C. (2010). Where do mirror neurons come from? *Neuroscience and Biobehavioral Reviews*, *34*(4), 575–583. <https://doi.org/10.1016/j.neubio-rev.2009.11.007>
- Heyser, C. J., Fienberg, A. A., Greengard, P., & Gold, L. H. (2000). DARPP-32 knockout mice exhibit impaired reversal learning in a discriminated operant task. *Brain Research*, *867*(1–2), 122–130.
- Hickok, G. (2013). Do mirror neurons subserve action understanding? *Neuroscience Letters*, *540*, 56–58. <https://doi.org/10.1016/j.neulet.2012.11.001>
- Hirstein, W., & Ramachandran, V. S. (1997). Capgras syndrome: A novel probe for understanding the neural representation of the identity and familiarity of persons. *Proceedings of the Royal Society B: Biological Sciences*, *264*(1380), 437–444.
- Honsberger, M. J., Taylor, J. R., & Corlett, P. R. (2015). Memories reactivated under ketamine are subsequently stronger: A potential pre-clinical behavioral model of psychosis. *Schizophrenia Research*. <https://doi.org/10.1016/j.schres.2015.02.009>
- Howes, O. D., Montgomery, A. J., Asselin, M. C., Murray, R. M., Valli, I., Tabraham, P., ... Grasby, P. M. (2009). Elevated striatal dopamine function linked to prodromal signs of schizophrenia. *Archives of General Psychiatry*, *66*(1), 13–20.
- Hume, D. (1739/2007). *A treatise of human nature*. Oxford, UK: Oxford University Press.
- Hume, D. (1900). *An enquiry concerning human understanding*. Chicago: The Open Court Publishing Co.; etc.
- Huq, S. F., Garety, P. A., & Hemsley, D. R. (1988). Probabilistic judgements in deluded and non-deluded subjects. *The Quarterly Journal of Experimental Psychology*, *40*(4), 801–812.
- Johnson, D. D., & Fowler, J. H. (2011). The evolution of overconfidence. *Nature*, *477*(7364), 317–320. <https://doi.org/10.1038/nature10384>
- Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, *88*(1), 67–85. <https://doi.org/10.1037//0033-295x.88.1.67>
- Jones, S. R., & Fernyhough, C. (2007). Thought as action: Inner speech, self-monitoring, and auditory verbal hallucinations. *Consciousness and Cognition*, *16*(2), 391–399. <https://doi.org/10.1016/j.concog.2005.12.003>
- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty*. Cambridge, UK: Cambridge University Press.
- Kamin, L. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior*. New York: Appleton-Century-Crofts.

- Kaney, S., & Bentall, R. P. (1992). Persecutory delusions and the self-serving bias. Evidence from a contingency judgment task. *The Journal of Nervous and Mental Disease*, 180(12), 773–780.
- Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. *The American Journal of Psychiatry*, 160(1), 13–23.
- Kapur, S. (2004). How antipsychotics become anti-“psychotic”—From dopamine to salience to psychosis. *Trends in Pharmacological Sciences*, 25(8), 402–406.
- Kapur, S., Mizrahi, R., & Li, M. (2005). From dopamine to salience to psychosis-linking biology, pharmacology and phenomenology of psychosis. *Schizophrenia Research*, 79(1), 59–68.
- Kihlstrom, J. F., & Hoyt, I. P. (1988). Hypnosis and the psychology of delusions. In T. F. Oltmanns & B. A. Maher (Eds.), *Delusional beliefs*. New York: Wiley.
- Kilner, J., Friston, K., & Frith, C. (2007). Predictive coding: An account of the mirror neuron system. *Cognitive Processing*, 8(3), 159–166. <https://doi.org/10.1007/s10339-007-0170-2>
- Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, 427(6971), 244–247. <https://doi.org/10.1038/nature02169>
- Koster-Hale, J., & Saxe, R. (2013). Theory of mind: A neural prediction problem. *Neuron*, 79, 836–848.
- Laruelle, M., Abi-Dargham, A., Gil, R., Kegeles, L., & Innis, R. (1999). Increased dopamine transmission in schizophrenia: Relationship to illness phases. *Biological Psychiatry*, 46(1), 56–72.
- Lavin, A., Nogueira, L., Lapish, C. C., Wightman, R. M., Phillips, P. E., & Seamans, J. K. (2005). Mesocortical dopamine neurons operate in distinct temporal domains using multimodal signaling. *The Journal of Neuroscience*, 25(20), 5013–5023.
- Lawson, R. P., Friston, K. J., & Rees, G. (2015). A more precise look at context in autism. *Proceedings of the National Academy of Sciences of the United States of America*, 112(38), E5226. <https://doi.org/10.1073/pnas.1514212112>
- Lawson, R. P., Rees, G., & Friston, K. J. (2014). An aberrant precision account of autism. *Frontiers in Human Neuroscience*, 8, 302. <https://doi.org/10.3389/fnhum.2014.00302>
- Lee, S. W., O’Doherty, J. P., & Shimojo, S. (2015). Neural computations mediating one-shot learning in the human brain. *PLoS Biology*, 13(4), e1002137. <https://doi.org/10.1371/journal.pbio.1002137>
- Leff, J., Williams, G., Huckvale, M. A., Arbutnot, M., & Leff, A. P. (2013). Computer-assisted therapy for medication-resistant auditory hallucinations: Proof-of-concept study. *The British Journal of Psychiatry: The Journal of Mental Science*, 202, 428–433. <https://doi.org/10.1192/bjp.bp.112.124883>
- Locke, J. (1690/1976). *An essay concerning human understanding*. London: Dent.
- Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37(11), 2098–2109.

- Lupyan, G., & Clark, A. (2015). Words and the world: Predictive coding and the language-perception-cognition Interface. *Current Directions in Psychological Science*, 24(4), 279–284. <https://doi.org/10.1177/0963721415570732>
- Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, 82, 276–298.
- Maher, B. A. (1974). Delusional thinking and perceptual disorder. *Journal of Individual Psychology*, 30(1), 98–113.
- Maher, B. A. (1988a). Anomalous experience and delusional thinking: The logic of explanations. In T. F. Oltmanns & B. A. Maher (Eds.), *Delusional Beliefs* (pp. 15–33). New York: Wiley.
- Maher, B. A. (1988b). *Delusions as normal theories*. New York: Wiley.
- Marshall, L., Mathys, C., Ruge, D., de Berker, A. O., Dayan, P., Stephan, K. E., & Bestmann, S. (2016). Pharmacological fingerprints of contextual uncertainty. *PLoS Biology*, 14(11), e1002575. <https://doi.org/10.1371/journal.pbio.1002575>
- McKay, R., Langdon, R., & Coltheart, M. (2005). “Sleights of mind”: Delusions, defences, and self-deception. *Cognitive Neuropsychiatry*, 10(4), 305–326.
- McKay, R., Langdon, R., & Coltheart, M. (2006). Need for closure, jumping to conclusions, and decisiveness in delusion-prone individual. *Journal of Nervous and Mental Disease*, 194(6), 422–426.
- McKay, R. T., & Dennett, D. C. (2009). The evolution of misbelief. *The Behavioral and Brain Sciences*, 32(6), 493–510; discussion 510–461. <https://doi.org/10.1017/S0140525X09990975>.
- McLaren, I. P., & Dickinson, A. (1990). The conditioning connection. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 329(1253), 179–186. <https://doi.org/10.1098/rstb.1990.0163>
- Mendoza, P. (1978). [In memoriam of Gerardo Varela]. *Gaceta medica de Mexico*, 114(5), 250.
- Menon, M., Pomarol-Clotet, E., McKenna, P. J., & McCarthy, R. A. (2006). Probabilistic reasoning in schizophrenia: A comparison of the performance of deluded and nondeluded schizophrenic patients and exploration of possible cognitive underpinnings. *Cognitive Neuropsychiatry*, 11(6), 521–536. <https://doi.org/10.1080/13546800544000046>
- Menon, M., Anderson, A., Schmitz, T., Graff, A., Korostil, M., Mamo, D., et al. (2011). Exploring the neural correlates of delusions of reference: An fMRI study. *Biological Psychiatry*, 70(12), 1127–1133.
- Meyer-Lindenberg, A., Straub, R. E., Lipska, B. K., Verchinski, B. A., Goldberg, T., Callicott, J. H., ... Weinberger, D. R. (2007). Genetic evidence implicating DARPP-32 in human frontostriatal structure, function, and cognition. *The Journal of Clinical Investigation*, 117(3), 672–682.
- Michotte, A. (1963). *The perception of causality*. Oxford, England: Basic Books.
- Miller, R. (1976). Schizophrenic psychology, associative learning and the role of forebrain dopamine. *Medical Hypotheses*, 2(5), 203–211.

- Miller, R. (1993). Striatal dopamine in reward and attention: A system for understanding the symptomatology of acute schizophrenia and mania. *International Review of Neurobiology*, 35, 161–278.
- Moran, P. M., Al-Uzri, M. M., Watson, J., & Reveley, M. A. (2003). Reduced Kamin blocking in non paranoid schizophrenia: Associations with schizotypy. *Journal of Psychiatric Research*, 37(2), 155–163.
- Morgan, H. L., Turner, D. C., Corlett, P. R., Absalom, A. R., Adapa, R., Arana, F. S., ... Fletcher, P. C. (2011). Exploring the impact of ketamine on the experience of illusory body ownership. *Biological Psychiatry*, 69(1), 35–41. <https://doi.org/10.1016/j.biopsych.2010.07.032>
- Moritz, S., & Woodward, T. S. (2005). Jumping to conclusions in delusional and non-delusional schizophrenic patients. *The British Journal of Clinical Psychology/The British Psychological Society*, 44.(Pt 2), 193–207. <https://doi.org/10.1348/014466505X35678>
- Moutoussis, M., Bentall, R. P., El-Deredy, W., & Dayan, P. (2011). Bayesian modelling of Jumping-to-Conclusions bias in delusional patients. *Cognitive Neuropsychiatry*, 16(5), 422–447. <https://doi.org/10.1080/13546805.2010.548678>
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175–220.
- Nyhan, B., & Reifler, J. (2013). *Which corrections work?* Washington, DC: New America Foundation.
- O’Tuathaigh, C. M., Salum, C., Young, A. M., Pickering, A. D., Joseph, M. H., & Moran, P. M. (2003). The effect of amphetamine on Kamin blocking and overshadowing. *Behavioural Pharmacology*, 14(4), 315–322. <https://doi.org/10.1097/01.fbp.0000080416.18561.3c>
- Palmer, C. J., Seth, A. K., & Hohwy, J. (2015). The felt presence of other minds: Predictive processing, counterfactual predictions, and mentalising in autism. *Consciousness and Cognition*, 36, 376–389. <https://doi.org/10.1016/j.concog.2015.04.007>
- Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex* (G. V. Anrep, Trans.). New York: Dover Publications.
- Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, 87(6), 532–552. <https://doi.org/10.1037/0033-295x.87.6.532>
- Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems* (Revised second printing ed.). San Mateo, CA: Morgan Kaufmann Publishers Inc.
- Pearl, J., & Russel, S. (2001). Bayesian networks. In M. Arbib (Ed.), *Handbook of brain theory and neural networks*. Cambridge, MA: MIT Press.
- Peirce. (1931–58). *Collected papers of Charles Sanders Peirce* (Vol. 1–6). Cambridge, MA: Harvard University Press.
- Peled, A., Pressman, A., Geva, A. B., & Modai, I. (2003). Somatosensory evoked potentials during a rubber-hand illusion in schizophrenia. *Schizophrenia Research*, 64(2–3), 157–163.

- Pessiglione, M., Petrovic, P., Daunizeau, J., Palminteri, S., Dolan, R. J., & Frith, C. D. (2008). Subliminal instrumental conditioning demonstrated in the human brain. *Neuron*, 59(4), 561–567. <https://doi.org/10.1016/j.neuron.2008.07.005>
- Plato. (350 B.C./1999). *Phaedo* (D. Gallop, Trans.). Oxford, UK: Oxford University Press.
- Pomarol-Clotet, E., Honey, G. D., Murray, G. K., Corlett, P. R., Absalom, A. R., Lee, M., ... Fletcher, P. C. (2006). Psychological effects of ketamine in healthy volunteers. Phenomenological study. *The British Journal of Psychiatry*, 189, 173–179.
- Powers, A. R., Mathys, C., & Corlett, P. R. (2017). Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. *Science*, 357(6351), 596–600. <https://doi.org/10.1126/science.aan3458>
- Powers, A. R., III, Kelley, M., & Corlett, P. R. (2016). Hallucinations as top-down effects on perception. *Biological Psychiatry: CNNI*, 1, 393–400.
- Proulx, T., Inzlicht, M., & Harmon-Jones, E. (2012). Understanding all inconsistency compensation as a palliative response to violated expectations. *Trends in Cognitive Science*, 16(5), 285–291. <https://doi.org/10.1016/j.tics.2012.04.002>
- Quine, W. V., & Quine, W. V. (1951). Two dogmas of empiricism. *Philosophical Review*, 60, 20–43.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.
- Robbins, T. W., & Everitt, B. J. (1996). Neurobehavioural mechanisms of reward and motivation. *Current Opinion in Neurobiology*, 6(2), 228–236.
- Robinson, T. E., & Berridge, K. C. (2001). Incentive-sensitization and addiction. *Addiction*, 96(1), 103–114.
- Schneider, K. (1957, September). Primary & secondary symptoms in schizophrenia. *Fortschritte der Neurologie, Psychiatrie, und ihrer Grenzgebiete*, 25(9), 487–490.
- Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Sciences*, 17(11), 565–573. <https://doi.org/10.1016/j.tics.2013.09.007>
- Seth, A. K., Suzuki, K., & Critchley, H. D. (2011). An interoceptive predictive coding model of conscious presence. *Frontiers in Psychology*, 2, 395. <https://doi.org/10.3389/fpsyg.2011.00395>
- Shanks, D. R., & Channon, S. (2002). Effects of a secondary task on “implicit” sequence learning: Learning or performance? *Psychological Research*, 66(2), 99–109.
- Snyder, M., & Swann, W. B., Jr. (1978). Behavioural confirmation in social interaction; from social perception to social reality. *Journal of Experimental Social Psychology*, 14, 148–162.

- Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience*, *16*(7), 966–973. <https://doi.org/10.1038/nn.3413>
- Stephan, K. E., Friston, K. J., & Frith, C. D. (2009). Dysconnection in schizophrenia: From abnormal synaptic plasticity to failures of self-monitoring. *Schizophrenia Bulletin*, *35*(3), 509–527. <https://doi.org/10.1093/schbul/sbn176>
- Svenningsson, P., Tzavara, E. T., Carruthers, R., Rachleff, I., Wattler, S., Nehls, M., ... Greengard, P. (2003). Diverse psychotomimetics act through a common signaling pathway. *Science*, *302*(5649), 1412–1415.
- Tang, J., Morgan, H. L., Liao, Y., Corlett, P. R., Wang, D., Li, H., ... Chen, X. (2015). Chronic administration of ketamine mimics the perturbed sense of body ownership associated with schizophrenia. *Psychopharmacology*, *232*(9), 1515–1526. <https://doi.org/10.1007/s00213-014-3782-0>
- Taylor, S. E., & Fiske, S. T. (1978). *Salience, attention and attribution: Top of the head phenomena* (Vol. 11). San Diego, CA: Academic.
- Teufel, C., Naresh, S., Veronika, D., Jesus, P., Johanna, F., Puja, R. M., ... Paul, C. F. (2015). Shift toward prior knowledge confers a perceptual advantage in early psychosis and psychosis-prone healthy individuals. *Proceedings of the National Academy of Sciences*, *112*(43), 13401–13406. <https://doi.org/10.1073/pnas.1503916112>
- Tranel, D., & Damasio, A. R. (1985). Knowledge without awareness: An automatic index of facial recognition by prosopagnosics. *Science*, *228*(4706), 1453–1454.
- Varela, F. G. (1971). Self-consciousness: Adaptation or epiphenomenon? *Studium generale; Zeitschrift für die Einheit der Wissenschaften im Zusammenhang ihrer Begriffsbildungen und Forschungsmethoden*, *24*(4), 426–439.
- Wada, M., Takano, K., Ora, H., Ide, M., & Kansaku, K. (2016). The rubber tail illusion as evidence of body ownership in mice. *The Journal of Neuroscience*, *36*(43), 11133–11137. <https://doi.org/10.1523/JNEUROSCI.3006-15.2016>
- Warren, H. C. (1921). *A history of the association psychology*. New York: Charles Scribner's Sons.
- Widrow, B., & Hoff, M. E., Jr. (1960). *Adaptive switching circuits*. Paper presented at the IRE WESCON convention rec.
- Yau, J. O., & McNally, G. P. (2015). Pharmacogenetic excitation of dorsomedial prefrontal cortex restores fear prediction error. *The Journal of Neuroscience*, *35*(1), 74–83. <https://doi.org/10.1523/JNEUROSCI.3777-14.2015>
- Young, G. (2008). Capgras delusion: An interactionist model. *Consciousness and Cognition*, *17*(3), 863–876.
- Zmigrod, L., Garrison, J. R., Carr, J., & Simons, J. S. (2016). The neural mechanisms of hallucinations: A quantitative meta-analysis of neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, *69*, 113–123. <https://doi.org/10.1016/j.neubiorev.2016.05.037>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

