Prediction and Expectation

RI Schubotz, University of Münster, Münster, Germany

Glossary

Active inference A term introduced in Friston's version of predictive coding to refer to the generalization of the predictive coding principle from perceptual systems to include the motor system. Movement understood as active inference is a top-down ('feedback') signal that suppresses the proprioceptive prediction error at the spinal level. This account abandons the idea of 'motor commands' and 'efference copies' that become proprioceptive and somatosensory predictions.

Entropy Entropy is a term that is used differently in different scientific disciplines. In the context of information theory, it quantifies the average surprise of outcomes in a probability distribution. In the context of prediction, entropy captures the average prediction error an animal experiences, that is, an animal's informational cost for representing the sensory input by its (not optimal) internal model.

Mutual information An information theoretical term that quantifies how much a known variable reduces uncertainty about another: If the mutual information is high between two variables (or state distributions), and one variable is known, this reduces the number of yes/no questions required to guess the unknown variable. Applied to the brain, mutual information should be high between our intern models (memories) and our sensory experiences.

Predictive coding The brain is organized hierarchically, and many theories on brain function assume that there is an interaction between 'higher' and 'lower' areas. However, the question is: in which way? In contrast to other models, the predictive coding account holds that feedback projections from higher areas are predictive in that they suppress the forward flow of information from lower areas, except a remaining residual prediction error that propagates upward.

Introducing Prediction

Prediction is omnipresent in our life. We are well aware of predictive requirements when crossing a street, singing in chorus, or handling a chain saw, but expectations do also emerge when we listen to familiar music or watch people going by. Actually, changes in the seconds range are of vital meaning to all animals that move. For instance, if something is moving close by, it can be prey, predator, a conspecific of lower or higher social rank, or simply a leaf in the wind. Obviously, it is highly relevant not only to timely react in an appropriate fashion but also to foresee how things will change and what demands they entail in terms of one's own behavior.

The *predictive brain* has become a rediscovered buzzword in cognitive neuroscience (Bubic, von Cramon, & Schubotz, 2010). The novel view on this well-established notion is that prediction is not an optimizing mechanism, but an intrinsic built-in principle operative in all cortical and subcortical processing (Clark, 2013). The roots of this view trace back to Helmholtz' nineteenth-century proposal of efference copies canceling self-induced sensations during eye movements in order to remain a stable head-centered representation of object locations (Helmholtz, 1860/1962). Since then, theoretical implementations of predictive signals have undergone a conceptual evolution whose most recent offspring are motor control theory (Wolpert, 2007) and predictive coding (Friston, 2005a; Mumford, 1992; Rao & Ballard, 1999). Launched as a simplification and generalization of the principles adopted by motor control theory, Friston's (2011) predictive coding account postulates that the brain generally strives to mitigate its surprise (i.e., prediction error). Thus, the brain notoriously estimates (not measures) the world's and body's states and it constantly learns from faulty guesses and estimates. It does so by *active inference*, that is, predictive codes traveling top-down to probabilistically weigh alternative excitation patterns in downstream areas. In turn, the latter sends the so-called prediction errors upstream to modify initial and future predictions. Thus, prediction is not a one-way propagation of information, but an ongoing exchange between cortical layer-specific *representation neurons* and *error neurons* that keep the nervous system prepared for and updated about states of environment and body. As an upshot, Helmholtz' problem of anticipatorily optimized motor control is reduced to perceptual inference in the sense that movement simply amounts to the suppression of proprioceptive prediction error (Adams, Shipp, & Friston, 2013; Friston, 2011).

While these models were originally developed to explain optimal motor control, the field has meanwhile drifted towards predictions applied in perception and beyond (Clark, 2013; Grush, 2004). One only starts to spell out exactly how a multitude of predictive requirements and phenomena are implemented in the brain. Generally, prediction surfaces as increased activity in cortical and subcortical areas that modulates and biases activity in downstream sensory areas. Consequently, brain research has mostly addressed the difference between brain responses to predicted versus unpredicted stimulation. While the principle of top-down prediction and bottom-up prediction error is suggested operative on all levels of the cortical hierarchy (cf. Rauss & Pourtois, 2013), most studies have addressed short-term prediction in perception and motor control. This article will focus on perceptual prediction (for motor prediction, see, e.g., Adams, Shipp, et al., 2013; Bastian, 2006; Shadmehr, Smith, & Krakauer, 2010; Wolpert & Flanagan, 2001; Wolpert, 2007).

Prediction at Different Timescales

Predicting the Present

A major application of perceptual prediction is to decide what states the world and the body are in. Thus, prediction enhances and speeds up disambiguation of either noisy or competing sensory information. For instance, detailed object information traveling up the ventral stream is suggested to be accompanied by two distinct sources of top-down prediction signals. On the one hand, object-sensitive visual areas in the posterior temporal cortex are suggested to provide the orbitofrontal cortex (OFC) with low-spatial-frequency information of the current image and receive from there an 'initial guess' about the most likely stimulus interpretations (Bar, 2003). On the other hand, low-spatial-frequency representations are also passed to parahippocampal areas sensitive to scene gist and contextual frames and receive from there experience-based associations that facilitate recognition of the presented object in that particular situation (Bar, 2004). For instance, the context 'beach' renders the occurrence of some objects ('sunshades') more probable than others ('umbrellas') (Fenske, Aminoff, Gronau, & Bar, 2006; Kveraga, Ghuman, & Bar, 2007). Finally, bottom-up analysis and top-down suggestions are integrated to result in object perception (Figure 1(a)). Top-down modulations on perception in support of a hierarchical predictive account are abundant in the literature (e.g., Alink, Schwiedrzik, Kohler, Singer, & Muckli, 2010; Kumar & Schönwiesner, 2012; Murray, Kersten, Olshausen, Schrater, & Woods, 2002; Vetter, Grosbras, & Muckli, 2013).

In line with the notion of early 'initial guesses,' it shows that stimulus response latencies of higher cortical, including frontal, areas are too short to adhere to a strictly serial bottom-up communication but instead dovetail with a reciprocating flow of signals underlying early predictive shaping (Bullier, 2001). Indeed, evidence has accumulated that even earliest processes in the primary visual cortex (V1) are sculptured by top-down modulation (Gilbert & Sigman, 2007; Rauss, Schwartz, & Pourtois, 2011; Spratling, 2010). These predictive signals have several effects on sensory cortices.

The earliest effect is a stimulus repetition-induced attenuation of activity in the respective sensory cortex (repetition suppression; e.g., Summerfield, Trittschuh, Monti, Mesulam, & Egner, 2008). It is slightly later followed by an additional attenuation called expectation suppression, which also interacts with repetition suppression (Todorovic & de Lange, 2012). While repetition suppression depends on local transition probabilities, expectation suppression rather depends on stochastic regularities that have been learned over the course of longer time periods. As will be discussed later, event-related brain potential (ERP) recordings seem to parallel this finding: violation of local transition probabilities trigger a so-called mismatch negativity (MMN), which is directly followed by a sustained positive-going deflection (P300) that is sensitive to deviance from rules or probabilities stored in long-term memory and currently active in working memory. Different explanations have been put forward as to what mechanism gives rise to repetition suppression, including neuronal fatigue, sharpening (i.e., sparser representation of stimuli), and facilitation (for review, see Grill-Spector, Henson, & Martin, 2006). Recent findings indicate that repetition suppression is not due to a passive, bottom-up neural adaptation to repeated stimuli, but due to violation of local prediction (Wacogne et al., 2011).

Strikingly, repetition enhancement has also been reported (Segaert, Weber, de Lange, Petersson, & Hagoort, 2013). This
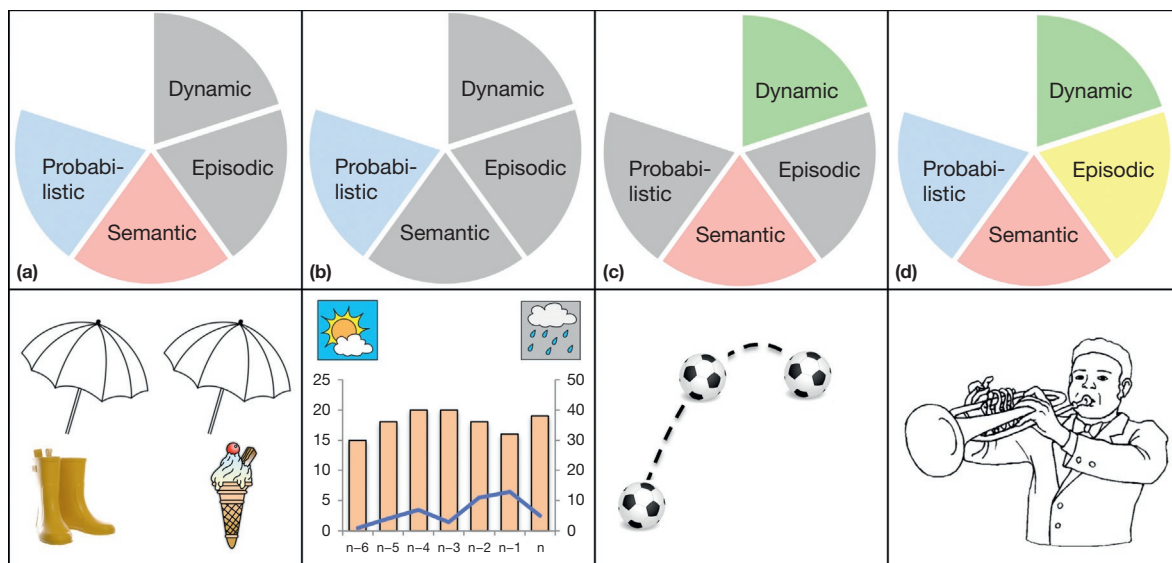
Figure 1 Predictions differ. (a) Low-frequency information of a retinal image is passed to higher visual and orbitofrontal areas so as to constrain object recognition. Here, other objects in the scene help to identify the shape as a sunshade, not an umbrella. (b) We come to learn contingencies on the basis of repeated experiences that follow a stochastic distribution. For instance, given sampling of the preceding days, we expect a sunny day tomorrow with a certain probability. (c) Dynamic prediction is ubiquitous in a world full of living entities, including ourselves. However, in modern times, inanimate motion may become fairly relevant (not only when we watch a soccer game). Premotor networks subserve dynamic prediction in both perception and perception-inducing movement. (d) When we witness actions, engage in conversation, or listen to a concert, multiple memory systems, and hence large networks including prefrontal sites, have to be concurrently exploited to run complex predictions.

effect seems to depend on several factors, including stimulus recognition, attention, expectation, and learning. de Gardelle, Waszczuk, Egner, and Summerfield (2012) found that within the same functional area, two-thirds of the voxels showed repetition suppression, whereas one-third of the voxels showed repetition enhancement. Authors argue that this mixture of reverse effects meets the hypotheses derived from predictive coding, as stimulus repetition should not only decrease surprise (i.e., prediction error) but also enhance prediction signals in the very same area. However, alternative models can explain for either repetition suppression (Grill-Spector et al., 2006) or repetition enhancement (Henson, Shallice, & Dolan, 2000; James & Gauthier, 2006).

Attention modulates the previously mentioned effects of prediction in an overtly complex way (e.g., Larsson and Smith, 2012; Murray & Wojciulik, 2004; Yi & Chun, 2005), and to neatly determine and to understand how prediction interacts with attention are ongoing endeavors. Both attention and prediction are considered to promote perceptual detection and recognition but presumably in different ways. One suggestion is that attention prioritizes processing of currently relevant information on the basis of motivation, whereas expectation biases the interpretation of percepts on the basis of memories (Summerfield & Egner, 2009). This difference is reflected by the effect that attention and expectation have on perceptual responses: While attended stimuli lead to increased perceptual responses, expected stimuli lead to decreased responses. Notably, increased responses to unexpected stimuli do not necessarily depend on attention, as demonstrated by the largely attention-independent MMN component. Critically, some robust findings on attentional modulation appear difficult to reconcile with a predictive coding account: For instance, frequently occurring targets elicit large ERPs, whereas infrequently (and hence unpredicted) distractors do not (cf. Bowman, Filetti, Wyble, & Olivers, 2013; further studies reviewed in Rauss et al., 2011). One solution to this problem may be to consider attention as a means by which prediction error is given increased influence or weight in the interplay of top-down and bottom-up flow of information (a mechanism also called *precision weighting*) (Feldman & Friston, 2010; Friston, 2009). This view may explain, for instance, why the neural responses to predicted stimuli are silenced when unattended, but not so when attended (Kok, Rahnev, Jehee, Lau, & de Lange, 2012). While there is some evidence that attentional prioritization promotes prediction errors from attended stimuli or features (Brown & Friston, 2013; Chennu et al., 2013; Jiang, Summerfield, & Egner, 2013; Vossel et al., 2014), future studies have to evaluate whether the precision weighting interpretation of attention can account for, and indeed is to be invoked for, the variety of findings on attention (cf. Larsson and Smith, 2012).

Predicting Upcoming States

The previously considered examples refer to prediction as mechanism of perception of single, static entities in the presence. Here, the 'pre' in 'prediction' simply means that the process of recognition does not start with the *full* stimulus, but with an internal representation retrieved from memory that already provides most of the information needed to identify the stimulus. That perception is not caused by information flowing in from our eyes, ears, and skin, as we formerly thought, certainly takes a little getting used to.

However, in everyday language, the notion of prediction refers to something in the future that is not yet here. So, how do we estimate what will be next (and exactly when), given the presence? Depending on the situation or experimental task, either formal transition rules (Janacsek & Nemeth, 2012), naive physics (Zago & Lacquaniti, 2005), or probability distributions (Fiser, Berkes, Orbán, & Lengyel, 2010) can serve to tune expectation for upcoming stimuli. However, a straightforward way to classify prediction is by its memory base. We will first consider prediction that newly arises ad hoc from information in a limited time window of a few seconds, before turning to prediction that exploits long-term memories.

Ad hoc prediction of local change

When we watch a movie or read a text while ignoring a sequence of auditory stimuli concurrently presented to us, rare (and hence unforeseen) deviant stimuli trigger the previously mentioned MMN ERP component. The MMN witnesses that the sequence of standard tones established the automatic expectation of an underlying regularity, be it with regard to tonal frequency, duration, intensity, or even abstract patterns (Näätänen et al., 1978, 2005, 2007). The MMN has a source in the auditory cortex (superior temporal planes), while further contributions from frontal sites, particularly the inferior frontal gyrus, but possibly also Anterior Cingulate Cortex (ACC) and/ or Supplementary Motor Area (SMA), remain controversial and variable across methods and implementations of the paradigm (for a review of frontal findings, see Deouell, 2007). The MMN has been discussed in favor of different predictive models of perception (Friston, 2005a, 2005b; Garrido, Kilner, Stephan, & Friston, 2009; Lieder, Stephan, Daunizeau, Garrido, & Friston, 2013; Schröger et al., 2013; Winkler, 2007). From the predictive coding perspective, the MMN is embraced as signature of a prediction error delivered from the primary auditory cortex via the superior temporal gyrus to inferior frontal sites so as to modify the feedback prediction signal (Friston, 2005a; Garrido et al., 2009). Analogue ideas have been put forward in case of the visual MMN (Kimura, 2012; Kimura, Schröger, & Czigler, 2011).

Prediction of global change

To be sure, the MMN is a very prominent but not the only ERP component signaling breaches of expectation. Several ERP components such as the P300 and the N400 have already been or will be incorporated in upcoming predictive models of perception (e.g., Bekinschtein et al., 2009; Chennu et al., 2013; Federmeier, 2007). In contrast to the MMN, however, these components are discussed as possible signatures of hierarchically nested and parallel predictions that arise from more global, contextual structure of stimulus sequences that are operative in parallel to local predictions (Vuust, Ostergaard, Pallesen, Bailey, & Roepstorff, 2009). As for Bar's (2003, 2004) accounts on object perception, long-term memories come into play here.

Actually, most everyday predictions of future states exploit not only experience gathered the last few seconds but also contingencies and probabilities that have been learned and stored in long-term memory. Different prediction types can

be related to different underlying long-term memory systems. O'Reilly, Jbabdi, Rushworth, and Behrens (2013) reported that in one and the same situation, two distinct types of predictions can be concurrently applied: Subjects estimated future states either probabilistically, that is, based on a history of probabilistically distributed events, or dynamically, that is, by extrapolation from current states and changes (Figure 1(b) and 1(c), respectively). Probabilistic prediction relied on the OFC, whereas dynamic (or change-based) prediction relied on a ventrolateral premotor and anterior intraparietal cortex, cerebellum, and dorsal striatum. These kinds of prediction will be shortly sketched in the following.

Styles and Sources of Prediction

Probabilistic Prediction

Contingencies we experience are stochastic, and thus, we learn to expect them with certain probabilities. When we engage in probabilistic prediction, we apply this mostly implicit knowledge about stochastic structures in the world. Due to their vital role in all animals, learning of reward contingencies has been extensively investigated in behavioral neuroscience. In the reversal-learning paradigm, reward contingencies might reverse on each trial. Based on this prediction error, the prediction can be immediately adapted in the forthcoming trial. The brain site that is most strongly associated with reversal learning is the OFC (Lucantonio, Stalnaker, Shaham, Niv, & Schoenbaum, 2012; Schoenbaum, Roesch, Stalnaker, & Takahashi, 2009).

In a similar way, probabilistic reinforcement learning requires subjects to integrate outcomes from many preceding trials to estimate the amount and probability of reward that is associated with a given stimulus or choice. A discrepancy between expected reward and actual experience is called *reward prediction error*. As to the cerebral implementation, reward prediction and evaluation are mediated by the dopaminergic system, comprising the basal ganglia, the anterior insula, and three sites in the frontomedian cortex: the OFC, the anterior cingulate cortex, and ventromedial prefrontal cortex (Garrison, Erdeniz, & Done, 2013; Glimcher, 2011; Rushworth & Behrens, 2008; Schultz, 2013). However, it remains subject to ongoing research which of these areas mediate the distinct steps of reward prediction, selection and execution of the 'winner' stimulus/action, signaling of the prediction error, and learning from the prediction error to improve future predictions (Daw & Doya, 2006).

Stating that the described network is engaged in predicting the subjective value of future events should not lead to the conclusion that it is exclusively reward- or value-related. Actually, the conviction that the OFC is related to reward could be a side effect of studying animals whose learning behavior could not be otherwise controlled. Evidence has accumulated that the striatum more generally provides a prediction error in response to changes in context, rules, and contingencies (Grahn, Parkinson, & Owen, 2008; Schiffer & Schubotz, 2011). Likewise, the role of the OFC currently undergoes a stark reconception (McDannald, Jones, Takahashi, & Schoenbaum, 2014) that leads away from a specialization in value and reward coding. Rather, the OFC is proposed to

exploit memory systems to represent hidden states of the environment in order to disambiguate current, context-dependent states that cannot be distinguished perceptually (Wilson, Takahashi, Schoenbaum, & Niv, 2014; note how well this fits with Bar's notion that the OFC disambiguates competing or noisy visual object information).

Dynamic Prediction

In contrast to probabilistic predictions, 'dynamic forward' or change predictions are applied when the relevant information is about ongoing observable change, that is, events. In order to serve as a source for dynamic prediction models, the respective cortical area would be expected to code for transitions or transformations in space–time. The premotor system mediates such constantly updating predictions not only when we move our body but also when we attend to 'dynamic forward' change in the environment (Schubotz, 2007). Premotor engagement in this kind of prediction has often been suggested to express a mobilization in virtue of response preparation (because premotor areas are part of the motor system) or was mistakenly attributed to the prefrontal cortex (because one rather expected prefrontal engagement in prediction, or cognition in general). The particular function that is mediated by premotor, but not prefrontal cortex, was thus not sufficiently appreciated until recently. It shows that instructed prediction of dynamically changing stimuli (coined 'serial prediction task') leads to premotor engagement, whereas attention to targets or serial match to sample in the same sequences does not (e.g., Bubic, von Cramon, Jacobsen, Schröger, & Schubotz, 2009). Thus, the particular role of premotor areas in perceptual prediction can be encapsulated by reference to *change* prediction: Target detection and serial match to sample are geared to the single stimulus (even when this task is performed on a sequence of stimuli), whereas the serial prediction task relies on processing stimulus–stimulus transitions. Depending on what stimulus property prediction is directed at, different subregions of the lateral premotor cortex are more engaged than others (e.g., Schubotz & von Cramon, 2001, 2002; Schubotz, von Cramon, & Lohmann, 2003; further studies reviewed in Schubotz & von Cramon, 2003; Schubotz, 2004, 2007), reflecting a default mapping (Wolfensteller, Schubotz, & von Cramon, 2007) that can be reshaped by sensorimotor training (Wolfensteller, Schubotz, & von Cramon, 2004). Integrating these findings, the Habitual Pragmatic Event Map (HAPEM) framework proposes that the premotor subregions are tuned to different styles of afferent transformation, such as rotation, deformation, or acceleration, to predictively guide both action and event perception (Schubotz, 2007).

The premotor loops include subcortical components, particularly dorsal striatal and cerebellar areas. Among these, the cerebellum has also been proposed a 'predictive device,' particularly with regard to interval timing (Ivry & Keele, 1989). Predictive accounts greatly appreciate the cerebellar contribution to predictive motor control (Bastian, 2006; Ito, 2008; Miall, Weir, Wolpert, & Stein, 1993; Wolpert, Miall, & Kawato, 1998), though largely ignoring the cerebellum's long-established role in perceptual and cognitive timing (Ivry & Keele, 1989; Salman, 2002).

High-Performance Prediction

Predictive signals from the premotor cortex may provide a bias for afferent–afferent spatiotemporal transitions, but not more. In contrast, networks including regions of the prefrontal cortex can be suggested to provide predictive signals based on not only statistical knowledge (as outlined for the OFC mentioned earlier) but also semantic and episodic memories. Generalizing the proposal by O'Reilly and colleagues (2013), such different types of long-term memory-based prediction presumably run in parallel to optimize preparedness for highly complex, variable, multidimensional, and multimodal events that we witness in everyday life: speech, action, and music.

While speech, action, and music are dynamic events, certain aspects of which can be dynamically predicted by extrapolation of current states and changes, further memory systems have to be used to approximate an acceptable level of predictive success (**Figure 1(d)**). For instance, prediction in music does not simply amount to the prediction of melodies and rhythms (Rohrmeier & Koelsch, 2012). Similarly, it is highly controversial to what extend prediction in speech perception can be construed in close analogy to prediction in action observation (as put forward by Pickering and Garrod, 2013). Speech comprehension is difficult to imagine without a role of lexical and conceptual expectations (Hickok, 2012) and tests of predictive models based on semantic memory yield fascinating results (Sohoglu, Peelle, Carlyon, & Davis, 2012). Likewise, action is more than motor control, and intentions cannot be reduced to future motor states (Pacherie & Haggard, 2010; Thinnes-Elker et al., 2012). Thus, while predictive models have been very successfully applied to motor control (Friston, Mattout, & Kilner, 2011; Kilner, Friston, & Frith, 2007), we are far from the whole picture of actions. When we consider action observation in its natural, semantically highly loaded environment and context, its *situatedness* (e.g., Rohlfing, Rehm, & Goecke, 2003), it is evident that predictions have to be fed by more than 'dynamic forward' models (which suffice to track hand transport or estimate target location in goal-directed grasping). Rather, it has to be augmented by predictions that exploit semantic, episodic, normative, and social memories. Recent findings indicate that the brain gathers and explores a vast amount of semantic and episodic information during action observation that is irrelevant to decode task-relevant instant goals, but highly relevant to disambiguate overarching goals. Thus, actions seem by default predictively perceived against the backdrop of the observer's semantic and episodic knowledge (Hrkać, Wurm, & Schubotz, 2013; Schiffer, Ahlheim, Ulrichs, & Schubotz, 2013; Schubotz, Korb, Schiffer, Stadler, & von Cramon, 2012; Wurm, Cramon, & Schubotz, 2012; Wurm & Schubotz, 2012; Wurm, von Cramon, & Schubotz, 2011). A major challenge for future ecologically valid accounts on prediction and expectation is to understand and integrate the multiple memory sources that feed several levels of interacting predictions while we perceive, cogitate, and act.

Abnormal and Pathological Prediction

Prediction relies on neural communication that is vulnerable to disturbances and pathologies. Obviously, hallucinations, illusions, and delusions are candidate phenomena for predictions going astray, as one may think of wrong percepts as generated by deficient integration of predicted and experienced sensations, resulting in incorrect prediction errors (Adams, Stephan, Brown, Frith, & Friston, 2013; Corlett, Honey, Krystal, & Fletcher, 2011; Friston, 2005b; Geisler & Kersten, 2002; Kumar et al., 2014). Patients suffering from schizophrenia show these symptoms and have been investigated with regard to deficient prediction. Aberrant prediction error signals are also indicated by ERP findings in schizophrenia, reporting changes in the P300 (Demiralp et al., 2002; van der Stelt, Frye, Lieberman, & Belger, 2004) and the MMN (Baldeweg, 2004; Umbricht & Krljes, 2005).

One hypothesis here is that the long-distance connections underlying predictive signals are impaired (Friston & Frith, 1995; Stephan, Friston, & Frith, 2009), resulting in uncompensated prediction error signals (cf. Yamashita & Tani, 2012). Recently, Jardri and Denève (2013) rather proposed that hallucinations arise from an altered local cortical microcircuitry due to imbalance between excitatory transmitter signaling and inhibitory transmitter signaling, as evident in schizophrenia (Lewis, Hashimoto, & Volk, 2005; O'Donnell, 2011) and psychosis (Corlett et al., 2011). In particular, circular belief propagation is suggested when bottom-up and top-down signals are reverberated and thus taken into account multiple times. Such fundamental disturbances in long-distance connections and local circuitry have been discussed not only with regard to schizophrenic's perceptual lapses but also with regard to their symptoms in motor control (Ford & Mathalon, 2012) and derogated sense of agency (Ford, Palzes, Roach, & Mathalon, 2013).

Outlook

In neuroscientific terms, predictions are action potentials that travel from one neuron to another to generate there an excitation state that otherwise could be generated by input from receptors. In a psychological sense, this amounts to a memory coming true. However, the world (including our body) is not always as memorized, and prediction errs. Resulting prediction errors hence travel back to train our memories. This learning changes our expectations, that is, the brain's context-dependent bias to deliver certain, that is, predictions.

The phenomenon of prediction has gathered great interest in neuroscience, and we start to appreciate in which way novel insights into this issue change long-established notions of perception and other mental functions. A challenge is to understand which memory systems are recruited to provide generative models for prediction in real life. Presumably, natural environments and situations draw on several styles of predictions and hence multiple memory systems at the same time (e.g., see **Figure 1**, upper panel). However, to find out whether there is a limit to the number of concurrently doable predictions, and how these predictions interact, awaits research.

Predictive coding has been recently further spelled out, for instance, with regard to possible biological implementations (Bastos et al., 2012) in cortical microcircuits (Douglas & Martin, 2004; Douglas, Martin, & Whitteridge, 1989) and

with regard to the role of neural oscillations in prediction of not only what but also when a perturbation will occur (Arnal & Giraud, 2012). However, while predictive coding provides a framework to accommodate many recent and older findings, more than a few issues and problems remain to be investigated and resolved (Clark, 2013). Among the most interesting challenges is a paradox that later has been coined the *dark room dilemma*: If the brain is doing nothing but trying to minimize its prediction errors (or 'minimize its entropy,' following the current predictive coding account), animals should go and seek a dark room and stay there – forever, in an 'ultimate stable state' that is perfectly predictable and absolutely boring (Mumford, 1992). This odd behavior appears to logically result from the mission to minimize the long-term average of prediction errors or *entropy* (hence to improve models of states of the world and body). However, that is obviously not the case: We go out for a walk, we join a human crowd on a festival, and we try to learn Greek at an evening course. Predictive coding leaves not much space for *exploration*, which is the natural observable counterpart to *exploitation*.

An elegant solution has been put forward by Little and Sommer (2013), who proposed to relinquish the motive of entropy and, staying with information theoretical concepts, to resort to *mutual information*, that is, shared information between the internal models and sensations. While both minimization of entropy and maximization of mutual information reduce prediction errors, only the latter (cf. Still, 2009) results in creatures that we are – exploring the world to learn while exploiting what we have learned.

See also: **INTRODUCTION TO ANATOMY AND PHYSIOLOGY**: Basal Ganglia; Cerebellum: Anatomy and Physiology; Lateral and Dorsomedial Prefrontal Cortex and the Control of Cognition; Motor Cortex; **INTRODUCTION TO CLINICAL BRAIN MAPPING**: Schizophrenia; **INTRODUCTION TO COGNITIVE NEUROSCIENCE**: A Frontostriatal Circuit for Timing the Duration of Events; Hierarchical Reinforcement Learning; Music; Neuroimaging Studies of Reinforcement-Learning; Rule Representation; Speech Perception; Statistical Learning; The Medial Temporal Lobe and Episodic Memory; Top-Down Suppression; **INTRODUCTION TO SOCIAL COGNITIVE NEUROSCIENCE**: Action Perception and the Decoding of Complex Behavior; Biological Motion; Mentalizing; **INTRODUCTION TO SYSTEMS**: Cortical Action Representations; Early Auditory Processing; Motion Perception; Reward; Visuospatial Attention.

References

Adams, R. A., Shipp, S., & Friston, K. J. (2013). Predictions not commands: Active inference in the motor system. *Brain Structure and Function*, 218(3), 611–643.

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), 1–26.

Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus predictability reduces responses in primary visual cortex. *The Journal of Neuroscience*, 30, 2960–2966.

Arnal, L. H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, 16(7), 390–398.

Baldeweg, T. (2004). Mismatch negativity potentials and cognitive impairment in schizophrenia. *Schizophrenia Research*, 69, 203–217.

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience*, 15(4), 600–609.

Bar, M. (2004). Visual objects in context. *Nature Reviews. Neuroscience*, 5(8), 617–629.

Bastian, A. J. (2006). Learning to predict the future: The cerebellum adapts feedforward movement control. *Current Opinion in Neurobiology*, 16(6), 645–649.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Perspective canonical microcircuits for predictive coding. *Neuron*, 76, 695–711.

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences of the United States of America*, 106(5), 1672–1677.

Bowman, H., Filetti, M., Wyble, B., & Olivers, C. (2013). Attention is more than prediction precision. *Behavioral and Brain Sciences*, 36(3), 206–208.

Brown, H. R., & Friston, K. J. (2013). The functional anatomy of attention: A DCM study. *Frontiers in Human Neuroscience*, 7(784), 1–10.

Bubic, A., von Cramon, D. Y., Jacobsen, T., Schröger, E., & Schubotz, R. I. (2009). Violation of expectation: Neural correlates reflect bases of prediction. *Journal of Cognitive Neuroscience*, 21(1), 155–168.

Bubic, A., von Cramon, D. Y., & Schubotz, R. I. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience*, 4(25), 1–15.

Bullier, J. (2001). Integrated model of visual processing. *Brain Research Reviews*, 36(2–3), 96–107.

Chennu, S., Noreika, V., Gueorguiev, D., Blenkmann, A., Kochen, S., Ibáñez, A., et al. (2013). Expectation and attention in hierarchical auditory prediction. *The Journal of Neuroscience*, 33(27), 11194–11205.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36, 181–253.

Corlett, P. R., Honey, G. D., Krystal, J. H., & Fletcher, P. C. (2011). Glutamatergic model psychoses: Prediction error, learning, and inference. *Neuropsychopharmacology*, 36(1), 294–315.

Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Current Opinion in Neurobiology*, 16, 199–204.

de Gardelle, V., Waszczuk, M., Egner, T., & Summerfield, C. (2012). Concurrent repetition enhancement and suppression responses in extrastriate visual cortex. *Cerebral Cortex*, 23(9), 2235–2244.

Demiralp, T., Ucok, A., Devrim, M., Isoglu-Alkac, U., Tecer, A., & Polich, J. (2002). N2 and P3 components of event-related potential in first-episode schizophrenic patients: Scalp topography, medication, and latency effects. *Psychiatry Research*, 111, 167–179.

Deouell, L. Y. (2007). The frontal generator of the mismatch negativity revisited. *Journal of Psychophysiology*, 21, 188–203.

Douglas, R. J., & Martin, K. A. C (2004). Neuronal circuits of the neocortex. *Annual Reviews of Neuroscience*, 27, 419–451.

Douglas, R. J., Martin, K. A. C, & Whitteridge, D. (1989). A canonical microcircuit for neocortex. *Neural Computation*, 1, 480–488.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44(4), 491–505.

Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, 4(215), 1–23.

Fenske, M. J., Aminoff, E., Gronau, N., & Bar, M. (2006). Top-down facilitation of visual object recognition: Object-based and context-based contributions. *Progress in Brain Research*, 155, 3–21.

Fiser, J., Berkes, P., Orbán, G., & Lengyel, M. (2010). Statistically optimal perception and learning: From behavior to neural representations. *Trends in Cognitive Sciences*, 14(3), 119–130.

Ford, J. M., & Mathalon, D. H. (2012). Anticipating the future: Automatic prediction failures in schizophrenia. *International Journal of Psychophysiology*, 83(2), 232–239.

Ford, J. M., Palzes, V. A., Roach, B. J., & Mathalon, D. H. (2013). Did I do that? Abnormal predictive processes in schizophrenia when button pressing to deliver a tone. *Schizophrenia Bulletin*, 40, 804–812.

Friston, K. J. (2005a). A theory of cortical responses. *Philosophical Transactions of the Royal Society B*, 360(1456), 815–836.

Friston, K. J. (2005b). Hallucinations and perceptual inference. *Behavioral and Brain Sciences*, 28, 764–766.

Friston, K. J. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, 13(7), 293–301.

Friston, K. J. (2011). What is optimal about motor control? *Neuron*, 72, 488–498.

Friston, K. J., & Frith, C. D. (1995). Schizophrenia: A disconnection syndrome? *Clinical Neuroscience*, 3, 89–97.

Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics*, 104(1–2), 137–160.

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology*, *120*, 453–463.

Garrison, J., Erdeniz, B., & Done, J. (2013). Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, *37*, 1297–1310.

Geisler, W. S., & Kersten, D. (2002). Illusions, perception and Bayes. *Nature Neuroscience*, *5*, 508–510.

Gilbert, C. D., & Sigman, M. (2007). Brain states: Top-down influences in sensory processing. *Neuron*, *54*, 677–696.

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(3), 15647–15654.

Grahn, J., Parkinson, J., & Owen, A. M. (2008). The cognitive functions of the caudate nucleus. *Progress in Neurobiology*, *86*(3), 141–155.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*, 14–23.

Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception. *The Behavioral and Brain Sciences*, *27*(3), 377–442.

Helmholtz, H. (1860/1962). *Handbuch der physiologischen Optik* (3rd ed.). New York: Dover.

Henson, R., Shallice, T., & Dolan, R. (2000). Neuroimaging evidence for dissociable forms of repetition priming. *Science*, *287*, 1269–1272.

Hickok, G. (2012). The cortical organization of speech processing: Feedback control and predictive coding the context of a dual-stream model. *Journal of Communication Disorders*, *45*(6), 393–402.

Hrkać, M., Wurm, M. F., & Schubotz, R. I. (2013). Action observers implicitly expect actors to act goal-coherently, even if they do not: An fMRI study. *Human Brain Mapping*, *1–13*, .

Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews. Neuroscience*, *9*(4), 304–313.

Ivry, R. B., & Keele, S. W. (1989). Timing functions of the cerebellum. *Journal of Cognitive Neuroscience*, *1*(2), 136–152.

James, T. W., & Gauthier, I. (2006). Repetition-induced changes in BOLD response reflect accumulation of neural activity. *Human Brain Mapping*, *27*(1), 37–46.

Janacsek, K., & Nemeth, D. (2012). Predicting the future: From implicit learning to consolidation. *International Journal of Psychophysiology*, *83*(2), 213–221.

Jardri, R., & Denève, S. (2013). Circular inferences in schizophrenia. *Brain*, *136*, 3227–3241.

Jiang, J., Summerfield, C., & Egner, T. (2013). Attention sharpens the distinction between expected and unexpected percepts in the visual brain. *The Journal of Neuroscience*, *33*(47), 18438–18447.

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: An account of the mirror neuron system. *Cognitive Processing*, *8*(3), 159–166.

Kimura, M. (2012). Visual mismatch negativity and unintentional temporal-context-based prediction in vision. *International Journal of Psychophysiology*, *83*, 144–155.

Kimura, M., Schröger, E., & Czigler, I. (2011). Visual mismatch negativity and its importance in visual cognitive sciences. *Neuroreport*, *22*(14), 669–673.

Kok, P., Rahnev, D., Jehee, J. F.M, Lau, H. C., & de Lange, F. P. (2012). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex*, *22*(9), 2197–2206.

Kumar, S., & Schönwiesner, M. (2012). Mapping human pitch representation in a distributed system using depth-electrode recordings and modeling. *The Journal of Neuroscience*, *32*(39), 13348–13351.

Kumar, S., Sedley, W., Barnes, G. R., Teki, S., Friston, K. J., & Griffiths, T. D. (2014). A brain basis for musical hallucinations. *Cortex*, *52*, 86–97.

Kveraga, K., Ghuman, A. S., & Bar, M. (2007). Top-down predictions in the cognitive brain. *Brain and Cognition*, *65*, 145–168.

Larsson, J., & Smith, A. T. (2012). fMRI repetition suppression: Neuronal adaptation or stimulus expectation? *Cerebral Cortex*, *22*, 567–576.

Lewis, D. A., Hashimoto, T., & Volk, D. W. (2005). Cortical inhibitory neurons in schizophrenia. *Nature Reviews. Neuroscience*, *6*, 312–324.

Lieder, F., Stephan, K. E., Daunizeau, J., Garrido, M. I., & Friston, K. J. (2013). A neurocomputational model of the mismatch negativity. *PLoS Computational Biology*, *9*(11), e1003288.

Little, D. Y., & Sommer, F. T. (2013). Maximal mutual information, not minimal entropy, for escaping the "Dark Room". *Behavioral and Brain Sciences*, *36*(3), 220–221.

Lucantonio, F., Stalnaker, A., Shaham, Y., Niv, Y., & Schoenbaum, G. (2012). The impact of orbitofrontal dysfunction on cocaine addiction. *Nature Publishing Group*, *15*(3), 358–366.

McDannald, M. A., Jones, J. L., Takahashi, Y. K., & Schoenbaum, G. (2014). Learning theory: A driving force in understanding orbitofrontal function. *Neurobiology of Learning and Memory*, *108*, 22–27.

Miall, R. C., Weir, D. J., Wolpert, D. M., & Stein, J. F. (1993). Is the cerebellum a smith predictor? *Journal of Motor Behavior*, *25*(3), 203–216.

Mumford, D. (1992). On the computational architecture of the neocortex. The role of cortico-cortical loops. *Biological Cybernetics*, *66*, 241–251.

Murray, S. O., Kersten, D., Olshausen, B. A., Schrater, P., & Woods, D. L. (2002). Shape perception reduces activity in human primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 15164–15169.

Murray, S. O., & Wojciulik, E. (2004). Attention increases neural selectivity in the human lateral occipital complex. *Nature Neuroscience*, *7*, 70–74.

Näätänen, R., Gaillard, A., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, *42*, 313–329.

Näätänen, R., Jacobsen, T., & Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): A review of the evidence. *Psychophysiology*, *42*(1), 25–32.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*, 2544–2590.

O'Donnell, P. (2011). Adolescent onset of cortical disinhibition in schizophrenia: Insights from animal models. *Schizophrenia Bulletin*, *37*, 484–492.

O'Reilly, J. X., Jbabdi, S., Rushworth, M. F.S, & Behrens, T. E.J (2013). Brain systems for probabilistic and dynamic prediction: Computational specificity and integration. *PLoS Biology*, *11*(9), e1001662.

Pacherie, E., & Haggard, P. (2010). What are intentions? In L. Nadel & W. Sinnott-Armstrong (Eds.), *Conscious will and responsibility. A tribute to Benjamin Libet* (pp. 70–84). Oxford: Oxford University Press.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(4), 329–347.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87.

Rauss, K., & Pourtois, G. (2013). What is the bottom-up and what is the top-down in predictive coding? *Frontiers in Psychology*, *4*(276), 1–8.

Rauss, K., Schwartz, S., & Pourtois, G. (2011). Top-down effects on early visual processing in humans: A predictive coding framework. *Neuroscience and Biobehavioral Reviews*, *35*(5), 1237–1253.

Rohlfing, K. J., Rehm, M., & Goecke, K. U. (2003). Situatedness: The interplay between context(s) and situation. *Journal of Cognition and Culture*, *3*(2), 132–156.

Rohrmeier, M. A., & Koelsch, S. (2012). Predictive information processing in music cognition. A critical review. *International Journal of Psychophysiology*, *83*(2), 164–175.

Rushworth, M. F.S, & Behrens, E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience*, *11*(4), 389–397.

Salman, M. S. (2002). The cerebellum: It's about time! But timing is not everything-new insights into the role of the cerebellum in timing motor and cognitive tasks. *Journal of Child Neurology*, *17*(1), 1–9.

Schiffer, A. M., Ahlheim, C., Ulrichs, K., & Schubotz, R. I. (2013). Neural changes when actions change: Adaptation of strong and weak expectations. *Human Brain Mapping*, *34*(7), 1713–1727.

Schiffer, A. M., & Schubotz, R. I. (2011). Caudate nucleus signals for breaches of expectation in a movement observation paradigm. *Frontiers in Human Neuroscience*, *5*(38), 1–12.

Schoenbaum, G., Roesch, M. R., Stalnaker, T. A., & Takahashi, Y. K. (2009). A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nature Reviews. Neuroscience*, *10*(12), 885–892.

Schröger, E., Bendixen, A., Denham, S. L., Mill, R. W., Bőhm, T. M., & Winkler, I. (2013). Predictive regularity representations in violation detection and auditory stream segregation: From conceptual to computational models. *Brain Topography*, *27*, 565–577.

Schubotz, R. I. (2004). Human premotor cortex: Beyond motor performance. In *MPI series in human cognitive and brain sciences* (vol. 50. Leipzig: Max Planck Institute for Human Cognitive and Brain Sciences.

Schubotz, R. I. (2007). Prediction of external events with our motor system: Towards a new framework. *Trends in Cognitive Sciences*, *11*(5), 211–218.

Schubotz, R. I., Korb, F. M., Schiffer, A. M., Stadler, W., & von Cramon, D. Y. (2012). The fraction of an action is more than a movement: Neural signatures of event segmentation in fMRI. *NeuroImage*, *61*(4), 1195–1205.

Schubotz, R. I., & von Cramon, D. Y. (2001). Functional organization of the lateral premotor cortex: fMRI reveals different regions activated by anticipation of object properties, location and speed. *Cognitive Brain Research*, *11*(1), 97–112.

Schubotz, R. I., & von Cramon, D. Y. (2002). Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: An fMRI study. *NeuroImage*, *15*(4), 787–796.

Schubotz, R. I., & von Cramon, Y. (2003). Functional-anatomical concepts of human premotor cortex: Evidence from fMRI and PET studies. *NeuroImage*, *20*, 120–131.

Schubotz, R. I., von Cramon, D. Y., & Lohmann, G. (2003). Auditory what, where, and when: A sensory somatotopy in lateral premotor cortex. *NeuroImage*, *20*(1), 173–185.

Schultz, W. (2013). Updating dopamine reward signals. *Current Opinion in Neurobiology*, *23*, 229–238.

Segaert, K., Weber, K., de Lange, F. P., Petersson, K. M., & Hagoort, P. (2013). The suppression of repetition enhancement: A review of fMRI studies. *Neuropsychologia*, *51*(1), 59–66.

Shadmehr, R., Smith, M. A., & Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annual Review of Neuroscience*, *33*, 89–108.

Sohoglu, E., Peelle, J. E., Carlyon, R. P., & Davis, M. H. (2012). Predictive top-down integration of prior knowledge during speech perception. *The Journal of Neuroscience*, *32*(25), 8443–8453.

Spratling, M. W. (2010). Predictive coding as a model of response properties in cortical area V1. *The Journal of Neuroscience*, *30*(9), 3531–3543.

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*, 509–527.

Still, S. (2009). Information theoretic approach to interactive learning. *Europhysics Letters*, *85*, 28005.

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, *13*, 403–409.

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M. M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience*, *11*(9), 1004–1006.

Thinnes-Elker, F., Iljina, O., Apostolides, J. K., Kraemer, F., Schulze-Bonhage, A., Aertsen, A., et al. (2012). Intention concepts and brain-machine interfacing. *Frontiers in Psychology*, *3*(455), 1–10.

Todorovic, A., & de Lange, F. P. (2012). Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *The Journal of Neuroscience*, *32*(39), 13389–13395.

Umbricht, D., & Krljes, S. (2005). Mismatch negativity in schizophrenia: A metaanalysis. *Schizophrenia Research*, *76*, 1–23.

van der Stelt, O., Frye, J., Lieberman, J. A., & Belger, A. (2004). Impaired P3 generation reflects high-level and progressive neurocognitive dysfunction in schizophrenia. *Archives of General Psychiatry*, *61*, 237–248.

Vetter, P., Grosbras, M. H., & Muckli, L. (2013). TMS over V5 disrupts motion prediction. *Cerebral Cortex*, *4*(2), 90–98.

Vossel, S., Mathys, C., Daunizeau, J., Bauer, M., Driver, J., Friston, K. J., et al. (2014). Spatial attention, precision, and Bayesian inference: A study of saccadic response speed. *Cerebral Cortex*, *24*, 1436–1450.

Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C., & Roepstorff, A. (2009). Predictive coding of music-brain responses to rhythmic incongruity. *Cortex*, *45*, 80–92.

Wacogne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors inhuman cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(51), 20754–20759.

Wilson, R. C., Takahashi, Y. K., Schoenbaum, G., & Niv, Y. (2014). Viewpoint orbitofrontal cortex as a cognitive map of task space. *Neuron*, *81*, 267–279.

Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology*, *21*(3–4), 147–163.

Wolfensteller, U., Schubotz, R. I., & von Cramon, D. Y. (2004). "What" becoming "where": Functional magnetic resonance imaging evidence for pragmatic relevance driving premotor cortex. *The Journal of Neuroscience*, *24*(46), 10431–10439.

Wolfensteller, U., Schubotz, R. I., & von Cramon, D. Y. (2007). Understanding non-biological dynamics with your own premotor system. *NeuroImage*, *36*(2), 33–43.

Wolpert, D. M. (2007). Probabilistic models in human sensorimotor control. *Human Movement Science*, *26*(4), 511–524.

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, *11*(18), 729–732.

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, *2*(9), 338–347.

Wurm, M. F., Cramon, D. Y., & Schubotz, R. I. (2012). The context-object-manipulation triad: Cross talk during action perception revealed by fMRI. *Journal of Cognitive Neuroscience*, *24*(7), 1548–1559.

Wurm, M. F., & Schubotz, R. I. (2012). Squeezing lemons in the bathroom: Contextual information modulates action recognition. *NeuroImage*, *59*(2), 1551–1559.

Wurm, M. F., von Cramon, D. Y., & Schubotz, R. I. (2011). Do we mind other minds when we mind other minds' actions? A functional magnetic resonance imaging study. *Human Brain Mapping*, *32*(12), 2141–2150.

Yamashita, Y., & Tani, J. (2012). Spontaneous prediction error generation in schizophrenia. *PLoS One*, *7*(5), e37843.

Yi, D. J., & Chun, M. M. (2005). Attentional modulation of learning-related repetition attenuation effects in human parahippocampal cortex. *The Journal of Neuroscience*, *25*, 3593–3600.

Zago, M., & Lacquaniti, F. (2005). Cognitive, perceptual and action-oriented representations of falling objects. *Neuropsychologia*, *43*(2), 178–188.