frontiers in Human Neuroscience

Front Hum Neurosci. 2010; 4: 25.
Published online 2010 Mar 22. Prepublished online 2009 Dec 23.
doi: 10.3389/fnhum.2010.00025

PMCID: PMC2904053
PMID: 20631856

Prediction, Cognition and the Brain

Andreja Bubic,[1,2,*] D. Yves von Cramon,[1,3] and Ricarda I. Schubotz[1,3]

[1]Department of Cognitive Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[2]University of Leipzig, Leipzig, Germany
[3]Max Planck Institute for Neurological Research, Cologne, Germany

Edited by: Shuhei Yamaguchi, Shimane University, Japan
Reviewed by: Hiroshi Nittono, Hiroshima University, Japan; Jun'ichi Katayama, Kwansei Gakuin University, Japan

*Correspondence: Andreja Bubic, Department of Cognitive Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1A, 04103 Leipzig, Germany. e-mail: bubic@cbs.mpg.de

Received 2009 Sep 22; Accepted 2010 Mar 7.

Copyright © 2010 Bubic, von Cramon and Schubotz.

This is an open-access article subject to an exclusive license agreement between the authors and the Frontiers Research Foundation, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.

Abstract

The term “predictive brain” depicts one of the most relevant concepts in cognitive neuroscience which emphasizes the importance of “looking into the future”, namely prediction, preparation, anticipation, prospection or expectations in various cognitive domains. Analogously, it has been suggested that predictive processing represents one of the fundamental principles of neural computations and that errors of prediction may be crucial for driving neural and cognitive processes as well as behavior. This review discusses research areas which have recognized the importance of prediction and introduces the relevant terminology and leading theories in the field in an attempt to abstract some generative mechanisms of predictive processing. Furthermore, we discuss the process of testing the validity of postulated expectations by matching these to the realized events and compare the subsequent processing of events which confirm to those which violate the initial predictions. We conclude by suggesting that, although a lot is known about this type of processing, there are still many open issues which need to be resolved before a unified theory of predictive processing can be postulated with regard to both cognitive and neural functioning.

Keywords: anticipation, expectation, internal models, novelty detection, prediction, prediction errors, prospection, simulation

Introduction

In 1637 the *Art of Worldly Wisdom* taught us that “even knowledge has to be in the fashion” (Gracian, 1991). The experience of working in neurosciences similarly illustrates that different times bring different research areas to the front, and these, in turn, promote certain concepts and ideas that draw a lot of

attention and drive the field for a certain period of time. Such “knowledge in fashion” typically fundamentally changes the way we conceptualize cognitive and neural processing by opening new paradigms, new interpretations and new perspectives. We suggest that the concept of a predictive brain can be considered as somewhat “in fashion” at the moment. It is an old idea being brought back to life in numerous areas of research where it is triggering small or larger-scale paradigm shifts. Terms such as prediction, prospection, anticipation, expectation, preparation as well as violations of expectations or prediction errors are being more and more present and exploited in our daily science. And, yet, although many researchers agree when emphasizing the pivotal role of prediction in both cognitive and neural processing, few would show the same level of concurrence with respect to the endorsed terminology, definitions, exemplary phenomena considered representative for such processing or the mechanisms suggested to underlie their occurrence. In fact, while some views and frameworks which have been put forward within this field can be considered mutually independent, others could be described as complimentary or occasionally even opposing. Therefore, it is of relevance to try and mutually compare these different views as this may reveal not just the widely emphasized and agreed-upon benefits of predictive processing, but also differences or even inconsistencies between approaches as well as the open questions which remain to be addressed within this field.

The Predictive Brain

In a very broad sense, predictive processing refers to any type of processing which incorporates or generates not just information about the past or the present, but also future states of the body or the environment. Such directedness towards the future has long been recognized as relevant and beneficial for different aspects of information processing, such as perception, motor and cognitive control, decision making, theory of mind and other cognitive processes in humans as well as, in a more rudimentary form, animals. Historically, the investigation of anticipatory mechanisms and representations started almost in parallel in the contexts of perceptual and motor processing. One of the first proposals that expectations are intrinsically related to actions was formulated in the 19th century within the ideomotor principle which has recently been revisited by theories suggesting the existence of shared or common codes between perception and action (James, 1890; Prinz, 1990; Hommel et al., 2001; Stock and Stock, 2004). Several decades later, similar ideas emphasizing a close interaction between sensory and motor processing have been suggested within the field of motor control. This was primarily motivated by initial investigations aimed at resolving a very fundamental question of how our visual world remains stabile despite constant image displacement introduced by eye and head movements. Early in the 20th century, Mach and Uexkuell offered a theoretical solution to this question which suggested that motor activity directly influences sensory processing, an idea which was already present in the thinking of Bell, Purkinje and von Helmholtz (Bridgeman, 2007). This was experimentally confirmed in 1950 when the terms “efference copy” and “corollary discharge”, concepts which are today widely accepted in the context of forward models in the motor and other cognitive domains, have been first introduced (Grüsser, 1995). While these approaches postulated how action selection depends on anticipated action outcomes and demonstrated how motor commands can be directly incorporated into sensory processes, early psychological experiments conducted by Wundt, Lange and James demonstrated how executing such actions becomes more efficient when these are based on appropriate perceptual expectations (LaBerge, 1995). In this view anticipation was treated as a form of attention which was regarded beneficial as it allowed more pertinent reactions in the immediate situation. James even conceptualized sensory anticipation as “pre-perception” of an event, given that it reflects the pre-activation of relevant brain structures which reduces the need for very elaborate processing following the actual event presentation (James, 1890). In addition to this, the current research on prediction in perception has been greatly influenced by von Helmholtz who argued that sensory systems evolved in order to infer the causes of changes in sensory inputs (Friston and Stephan, 2007), thus equating perception with recognition, namely inference about the state of the world. Although it is theoretically conceivable that such inference could be accomplished though different types of computations, it has lately been demonstrated that it, in fact, incorporates predictive mechanisms supported by recurrent neural processing

(Friston et al., 2006; Bar, 2007). During the early decades of psychological research the importance of expectations was also emphasized for other cognitive functions such as, e.g., learning (Tolman, 1948) as well as behavioral sequencing, namely serial ordering of behavior across different hierarchical levels (Lashley, 1951).

Although suggestions regarding the importance of predictive mechanisms originate from very early phases of both psychology and neurosciences, until recently they have not been strongly advocated in the majority of frameworks of cognitive and neural processing. Specifically, typical approaches in delineating cognitive processes postulate a rather serial process, starting with sensory, continuing with executive and "higher cognitive" functions and ending in overt behavior. Such thinking stems from the original behaviorist conceptualizations which emphasized the linear progression from sensory stimulation to overt behavior, a view which was also present in early information-processing cognitivist theories. And, even though the extreme behaviorist stimulus–response view of human behavior is today rarely or almost never advocated, different cognitive processes are today still dominantly studied in isolation. In addition, although rarely explicitly postulated, it is often assumed that a given process of interest starts with the output of some earlier, lower-level process and terminates once it provides input to the next processing stage. For example, one typically assumes that executing movements follows sensory processing and decision making while object recognition occurs once low-level visual processing is finished, providing input for higher-level cognitive functions. Although such hypotheses are not labeled as "reactive" or "non-predictive", they are in many ways challenged by conceptualizations promoted by explicitly defined predictive frameworks. These introduce paradigms and findings which indicate high inter-dependence of processes which are typically investigated in isolation, such as action and perception. In addition, although they do not question the importance of the feedforward information flow, these approaches primarily emphasize the relevance of feedback and recurrent processing. This does not imply that feedback or top-down biases represent exclusively predictive phenomena or that they are incompatible with more traditional views on cognitive or neural processing. After all, the ideas regarding constant exchange between incoming sensory data and existing knowledge used for postulating hypothesis regarding sensations have been advocated in the "analysis by synthesis" views formulated in the early days of cognitive psychology (MacKay, 1956). Regardless of this, however, the importance of top-down processing is emphasized much more within predictive approaches which often aim at determining how previous knowledge influences and guides current event processing. These approaches show how, given the levels of ambiguity and noise which is always present both in the environment and our neural system, such prior biases become crucial for facilitating and optimizing current event processing, regardless of whether it concerns recognizing objects, executing movements or scaling emotional reactions. They allow us to act, not solely react once all relevant information has been presented and fully processed, by making predictions about what to expect next while talking into account the current context and previous experiences integrated across different timescales.

In the previous paragraph, a general predictive approach to cognitive and neural processing has been compared, but not directly contrasted with classical behaviorist and early information- processing views on neural and cognitive processing. In addition to this, predictive frameworks can also be more directly contrasted with approaches which emphasize postdictive or retrospective processing mechanisms. Generally, such views suggest that efficient processing of events in ambiguous contexts does not need to result from effective preparation, but retrospective use of information regarding events which occurred following those of interest. However, although experimental evidence indicates that the brain uses such postdictive mechanisms in certain contexts (Whitney and Murakami, 1998; Eagleman and Sejnowski, 2000; Enns and Lleras, 2008), there is no unified retrospective or postdictive framework which may be contrasted with the predictive ones. And, although the suggested predictive and postdictive mechanisms represent mutually opposing phenomena, there is evidence that both types of processes may coexist in certain situations (e.g., Soga et al., 2009). Such contexts also clearly indicate how prediction may be dissociated from memory, a function with which it is intrinsically deeply related. As suggested before,

prediction crucially depends on previous experience and builds on memories of various kinds. It does not, however, include mnemonic encoding nor can it be reduced to mnemonic recall. Furthermore, it does not necessarily have to occur in all contexts where previous experiences are available for generating expectations about the future. In contrast, predictive processes cooperate and actively build on mnemonic ones and, in that sense, help to generate goal-directed and adapted behavior.

Many terms, how many meanings?

All throughout the history of cognitive (neuro)science, different terms, e.g., anticipation, expectation, prediction, prospection or preparation, have been used with respect to predictive processing. Although these do not necessarily transmit the same meaning, they are rarely clearly differentiated. For example, LaBerge (1995) defined the terms “anticipation” or “preparation” as elevated levels of processing in sensory or motor areas occurring prior and facilitating the processing of the expected perceptual or motor event. In contrast, the term “expectation” reflects a memory component as it refers to an item stored in either working or long-term memory which includes the information regarding the spatial and temporal characteristics of the expected event (LaBerge, 1995). Given that such representations can also be coded in rather abstract or verbal forms they do not necessarily presuppose a pre-activation of the relevant sensory cortices. A somewhat different distinction is offered by Butz et al. (2003) who compared the usage of the terms anticipation and prediction. According to this view, these terms convey partially different meanings: while prediction refers to a representation of an event (potentially comparable to the previous specification of the term expectation), anticipation describes the impact of predictions on current behavior, e.g., decisions and actions based on such predictions. Even though not completely specified, these are still somewhat more clearly defined when compared to the widely used term “prospection”. Gilbert and Wilson (2007, p. 1352) define prospection as an ability to “pre-experience the future by simulating it in our minds” which may, however, lack the detail and richness of genuine perceptions. Given that these simulations often do not mimic events of interest in a very reliable and realistic fashion, they are prone to errors which reflect mistakes in representing the context or content of simulated events. For example, such simulations are typically shortened and essentialized when compared to real events. In addition, while often being based on specific exemplars which are not representative for the target scenario, they also tend to be run in a comparative and decontextualized manner (Gilbert and Wilson, 2007, 2009). Thus, although prospection can include some aspects of both expectation and anticipation, it is not clearly specified in which extent and under which conditions. Therefore, this term is more suited to refer to a more general orientation towards the future in a sense that stored information is constantly used to imagine, simulate and predict future events (Schacter et al., 2007). This specification may, however, not be quite in line with the view endorsed by Schütz-Bosbach and Prinz (2007) who define prospective codes in event production and simulation as representations of present events which contain information pertaining to their future effects or goals. In this sense, prospective codes describe specific effects associated with a certain event and are therefore similar to the previous specification of expectations. However, prospective coding as used in this context should not be confused with the account of predictive coding which describes how causes are mapped to their sensory expressions (e.g., motor command to sensory consequences of such a command), a process not necessarily synonymous with forecasting (Friston et al., 2006; Kilner et al., 2007).

In addition, although some aspects of predictive processing have traditionally been regarded as specific forms of attention, it has recently been suggested that these represent fully distinct phenomena in the sense that expectations guide visual processing based on prior likelihood in contrast to attention which prioritizes sensory processing based on the motivational relevance of presented stimuli (Summerfield and Egner, 2009). Although appealing, such separation may not be warranted as attention unifies a broad range of phenomena aimed not just at selecting motivationally relevant stimuli, but also maintaining relevant activity and, importantly, preparing for the incoming events (LaBerge, 1995). Such broad conceptualization is not just a historical relict, but also reflects the fact that all described phenomena [selective attention, sustained attention (vigilance) as well as preparation], reflect focusing or enhancing

the processing of currently relevant information in contrast to, e.g., arousal or alertness which represent non-specific activation phenomena (Oken et al., 2006). Therefore, when comparing the terms prediction and attention, it may be of use to clearly specify the aspect of attentive processing to which predictive processing is being compared. Among these different aspects, selective attention bears the most similarity with prediction and these can partly overlap in some underlying neural mechanisms, although they may also be characterized by different temporal course, type of information used for biasing information processing or other features.

In summary, although a wide range of terms exist for describing the basic phenomenon of foresight, these are not consistently used and are often interchanged. This is somewhat problematic, as the lack of systematization in terminology may become reflected in the lack of systematization in understanding the phenomenon of interest. Overall, there seems to be an agreement that the term expectation describes the representations of what is predicted to occur in the future. We suggest that the process of formulating and communicating these expectations to the sensory or motor areas which become activated prior to the realized event could, in turn, be summarized under the term anticipatory or preparatory processing. While this type of processing describes expectations formulated on short timescales, consideration of potential distant future events could be termed prospection. The term prediction has previously been used to describe both a single event expected within a certain context, and the overall process of postulating such "single predictions". Given this ambiguity, the term prediction (predictive processing) could preferentially be used for describing the general orientation towards the future which includes a wide range of predictive phenomena, although in some contexts it may also serve as an appropriate synonym for expectation (Figure 1).

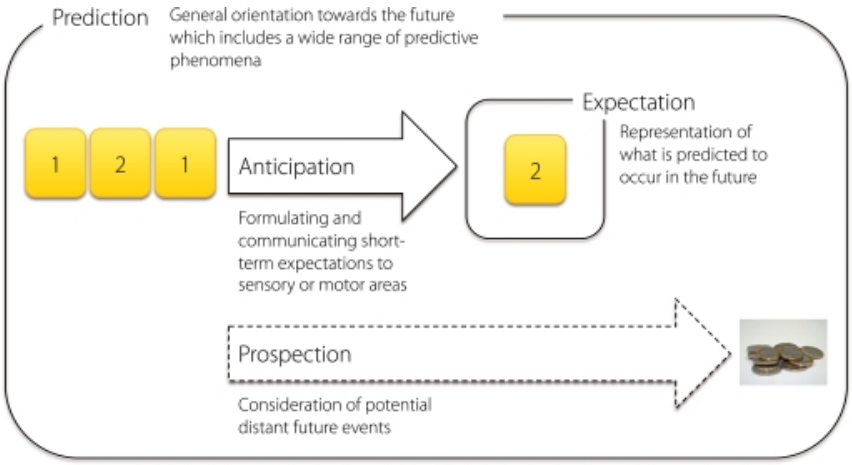

Figure 1

Basic terms that describe different aspects of a general orientation towards the future. When presented with a sequence of regularly alternating numbers "1" and "2", we soon start expecting that a number "2" should occur after the number "1", a process which results in a specific anticipatory activation of the sensory as well as motor cortex. The representation of the specific expected event can, in turn, be termed expectation. While this type of a process occurs on a short timescale, prospection describes the formulation of rather long-term expectations which do not necessarily lead to specific sensory or motor pre-activations, e.g., when a participant in an experiment foresees to receive a payment after the experimental session. Although all these terms describe somewhat different phenomena, they may all be considered as specific aspects of a common directedness towards the future which may be termed prediction.

Nature and strength of prediction

In the previous paragraphs we have introduced the basic terminology which describes different facets of predictive processing and indicated some inconsistencies in its usage. However, even if this terminology was to become more uniform and agreed upon, it would still not solve all existing problems in this area, as a more elaborate systematization which could account for differences between different levels, timescales or types (e.g., implicit and explicit) of predictions would still be lacking. Specifically, the nature and strength of predictions varies greatly in different contexts and may be influenced by different factors, e.g., the strength of the relationship between different events, frequency or context of their occurrence, etc. While in some situations expectations can be formulated in a rather unspecific manner and be restricted to a selected set of event features, e.g., sensory modality or location of an incoming stimulus, in others they may be very specific and pertain to the exact stimulus identity as well as the timing of its appearance. In addition, although a separation between implicit anticipations expressed through habits (behavior) and explicit ones which include representations of the predicted future states has been suggested (Pezzulo, 2008), it is still not clear whether these should be considered as a dichotomy or if it would be more appropriate to posit a continuous distribution of representations characterized by different degree of explicitness. Furthermore, prediction can take place on different temporal scales. First, expectations can be formulated based on the knowledge gained through long-term experience (Bar, 2007) or learning triggered by short-term exposure to non-random patterns (Schubotz, 2007). Second, it is possible to predict events which are expected to occur in different moments in the future, e.g., those expected to occur within seconds-range in contrast to those which may occur in the distant future. Long-term prediction is usually used “offline” and is not necessarily coupled with any immediately relevant or running process in contrast to short-term prediction which is more likely to be used “online” for regulating the ongoing behavior, as exemplified in motor control where it is coupled to the current sensorimotor cycle (Pezzulo et al., 2008). Consequently, prediction occurring on shorter timescales is typically more accurate when compared to long-term planning. In this context it is important to note that the timescale of prediction should not be confused with the concept of temporal expectations, namely a foresight of when something will occur (Nobre et al., 2007), which interact with expectations about other event properties in order to optimize our behavior. In addition, it is possible to generate multiple expectations pertaining to different points in space and time, as done in hierarchical predictive systems (Pezzulo et al., 2008) which capture the hierarchical organization of cognitive processes, the neural system and behavior (Dehaene and Changeux, 1997; Friston et al., 2006; Grafton and Hamilton, 2007; Kiebel et al., 2008, 2009). Furthermore, it has been shown that multiple expectations pertaining to the same event occurring at one point in time may also be formulated across different brain systems. For example, Ritter et al. (1999) demonstrated how rare predictable tones may be classified as violations of expectations at a preattentive, lower level of cognitive processing and, in the same time, as expected events at a higher level. Expectations of such different type and specificity could be mediated through different mechanisms or, alternatively, be based on the same types of processes partially implemented within different brain regions. Understanding how such different types of predictions are coded in the brain will be crucial in understanding their mutual relations and potential interactions. In summary, in describing different aspects of prediction, it is always important to clearly specify as many features of such processing as possible. As indicated in this paragraph and Figure 2 , there exist numerous factors which may be of relevance in this context. Although it may sometimes be difficult to clearly specify all of them, it is nevertheless important to try, as this may substantially aid general understanding and future progress within the field.

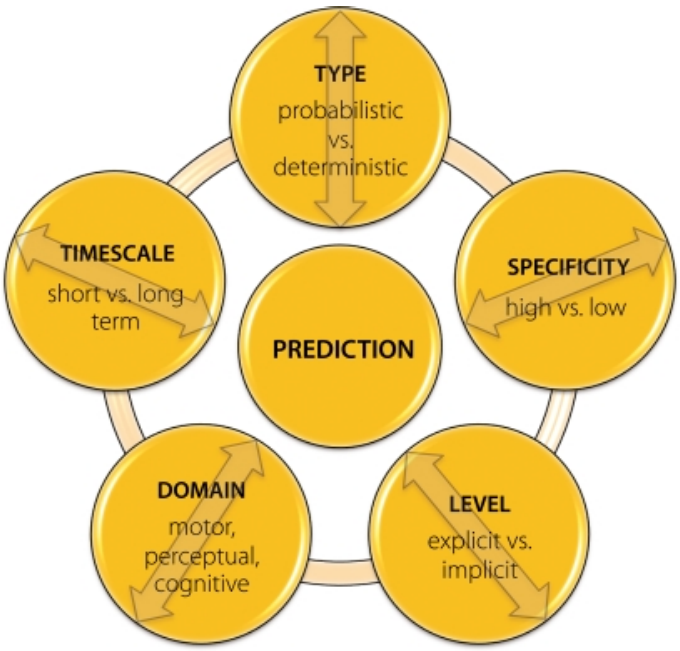

Figure 2

Main factors which specify the nature of predictive processes across different contexts. It is important to note that these do not represent mutually independent and orthogonal dimensions, but may in some contexts strongly overlap and interact. Furthermore, some additional features such as, e.g., type of information used for formulating expectations, might also be of relevance when specifying the exact nature of the predictive phenomenon of interest.

Prerequisites and benefits of prediction

Prerequisites of prediction

Events can be predictable if they occur in a non-random fashion, allowing the brain to extract either deterministic or probabilistic regularity of the relationship between different events. This knowledge can later be used for predicting the occurrences of some events following the presentation of those customarily preceding them. It is important to note that this statement describes only situations which afford predictability and does not say anything about how the brain deals with completely novel or random input. Although such contexts do not promote predictive processing directly, there is evidence which suggests that the brain may still employ similar predictive strategies in an attempt to extract a pattern within the random input (Schubotz and von Cramon, 2002a; Schubotz, 2004) or relate the novel input to familiar knowledge by generating analogies, thus facilitating the processing of new stimuli (Bar, 2007). On the other hand, predictions generated in non-random contexts may be based on learning and identifying associations, especially temporal dependencies between events (Butz et al., 2003; Bar, 2007). This can be accomplished by accumulating information related to statistical regularities while dealing with constant noise and uncertainty in the environment (Kording and Wolpert, 2006), as well as applying inference rules or deducing analogies between events (Pezzulo et al., 2008; Bar, 2009). Within the perceptual domain, rules (i.e., regular relations between events) of different complexity may afford predictability of the incoming stimuli. For example, concrete (so-called first-order) rules defined by constant repetition of a stimulus (or a stimulus feature) can trigger an expectation about the continuation of its appearance in the future (Squires et al., 1976). On the other hand, second-order or even higher-order (contingency) rules

(Näätanen et al., 2001; Shanks, 2007) which require the extraction of relations between specific, mutually non-interchangeable stimuli can underlie expectations related to more complex events, even those which were previously not encountered within the respective context.

One special instance in which the knowledge about the temporal structure of incoming events can be used for predicting upcoming events is serial order processing (serial pattern learning, sequence processing or sequencing) across different domains. Such processing may also differ in its complexity, allowing one to distinguish between simple linear (flat) sequences based on learning local dependencies between neighboring items and non-linear (hierarchical) sequences defined by long-distance dependencies (Conway and Christiansen, 2001; Fitch and Hauser, 2004; Bapi et al., 2005; Opitz and Friederici, 2007). Although most prominent in the language and the motor domain (Cohen et al., 1990; Tanji and Shima, 1994; Clegg et al., 1998; Keele et al., 2003; Ashe et al., 2006), serial order processing has also been studied in the context of artificial grammar learning (Reber, 1967; Cleeremans and McClelland, 1991; Bahlmann et al., 2008), music (Pfordresher et al., 2007), perception (Schubotz and von Cramon, 2001a; Remillard, 2003; Hoen et al., 2006) and executive functions (Koechlin et al., 2000; Jubault et al., 2007). These domains are comparable as they all require ordinal processing of different information types and may share not only functional commonalities, but partly also the underlying neural substrate (Lelekov-Boissard and Dominey, 2002; Janata and Grafton, 2003; Patel, 2003; Fiebach and Schubotz, 2006; Jubault et al., 2007; Opitz and Friederici, 2007). And, although all of them can greatly benefit from predictive processing, the level of anticipation afforded in them may depend on the type of acquired sequence knowledge. Specifically, Willingham et al. (1989) showed that, in comparison to implicit, the explicit knowledge is more likely to allow participants to predict the upcoming stimulus before it appears.

While the majority of previous examples indicate how temporal structure of incoming events affords predictability, spatial or more abstract relations between events can also become important sources of predictions. Such relations can be conceptualized as context frames, namely contextual structures which provide sets of expectations about the identity of stimuli and thus facilitate perception and action (Bar, 2004; Fenske et al., 2006). More elaborate long-term predictions or simulations of the future can, on the other hand, be based on recombinations of past events and concrete episodes contained within the individuals' episodic memory (Schacter et al., 2007) and used for creating "memories for the future" (Ingvar, 1985).

Benefits of prediction

As previously mentioned, the benefits of preparation have been recognized very early both in the motor and the perceptual domain. Behavioral experiments conducted by Wundt showed that attention and expectations related to the upcoming stimulus can shorten perception time, while Lange demonstrated beneficial behavioral effects following the correct anticipation of a response (LaBerge, 1995). Ever since the 19th century, more and more advantages of predictive in contrast to pure reactive processing have been postulated. Llinas (2002) argued that predictions identified at different levels of processing save resources and allow the perceiver to prepare the appropriate reactions. They can lead to faster recognition and interpretation of events encountered in the environment (Bar, 2007) by limiting the repertoire of potential responses to such events. Given that the information relevant for planning and executing appropriate reactions are available sooner, measurable benefits of anticipatory processing include an increase in accuracy, speed or maintenance of information processing (LaBerge, 1995). In addition, expectations allow us to construct a coherent and stable representation of the environment which is usually not easy, given the available, often impoverished (noisy and delayed) information (Kveraga et al., 2007a). Moreover, they may guide top-down deployment of attention, improve information seeking as well as subsequent decision making (Butz and Pezzulo, 2008). Thus, in a way, prediction allows us to act, and not solely react to events occurring around us.

On a more general functional and behavioral level, the ideomotor principle suggested that anticipated

sensory consequences of one's actions can trigger and guide behavior (Hommel et al., 2001), which has been supported by studies showing that representations of events or actions include the anticipated effects of those events as well as intentions behind the actions (Kerzel et al., 2000; Schütz-Bosbach and Prinz, 2007). Thus, Schütz-Bosbach and Prinz (2007) suggested that both perception and production of such events rely on prospective codes which incorporate information about their future states. Kunde et al. (2007) similarly argued that anticipation constitutes a necessary prerequisite for action because any action or response needs to start with a response-related anticipation. In this view, voluntary behavior in general is initiated and controlled by a representation of its expected outcomes (Hoffmann et al., 2007), illustrating how anticipation lies in the foundations of goal-directed behavior (Pezzulo, 2008). The initiation of such behavior could additionally be facilitated by simulations of the expected emotional consequences of actions (premotions) which can also be used as a basis for additional predictions (Gilbert and Wilson, 2009).

Importance of prediction across neurocognitive domains

All of the previously described benefits of prediction indirectly suggest that this type of processing could be useful for numerous cognitive domains and functions. Such a suggestion even has some common sense validity: for example, one does not enter a car to drive many miles because one is hungry, but because one is hungry and expects to find food in a restaurant which can be reached using a car. Therefore, not surprisingly, the fact that we are constantly oriented towards the future, that we formulate intentions and plan our future actions has been acknowledged and widely studied within the fields of, e.g., planning or prospective memory (Winograd, 1988; Friedman and Scholnick, 1997; Morris and Ward, 2005). In recent years, however, a more elaborate view according to which prediction or anticipation represents a fundamental principle of brain functioning which is “at the core of cognition” (Pezzulo et al., 2007, p. 68) has emerged. Illustrating its importance, the anticipatory nature of many cognitive functions and neural systems, such as motor control (Wolpert and Flanagan, 2001), motor imagery and action understanding (Jeannerod, 2001; Kilner et al., 2007), visual processing and attention (Mehta and Schaal, 2002; Enns and Lleras, 2008), language (DeLong et al., 2005), music (Keller and Koch, 2008), emotional processing (Ueda et al., 2003; Nitschke et al., 2006; Herwig et al., 2007; Gilbert and Wilson, 2009), executive functions (Partiot et al., 1995; Baker et al., 1996; Fuster, 2001; Wylie et al., 2006), and the theory of mind (Frith and Frith, 2006) has been suggested and experimentally demonstrated.

However, although predictive accounts of many functions have been posited and are currently well accepted, it is not always easy to show that it is indeed predictive processing which underlies a certain phenomenon. For example, some aspects of two phenomena which are typically considered as representative cases of prediction in vision, namely representational momentum (Kerzel, 2005) and the flash-lag effect (Nijhawan, 1997), could also be explained by retrospective processing (Whitney and Murakami, 1998; Eagleman and Sejnowski, 2000; Enns and Lleras, 2008). Furthermore, it has been shown that the visual extrapolation of the moving object's position in the flash-lag effect seems to occur at a stage when visual input is transformed into information needed for motor response, and not during early stages of visual processing (Kerzel and Gegenfurtner, 2003). In addition, it is important to recognize that prediction in visual perception is often discussed on two different levels. On the one hand, according to the more literate meaning of the concept, predictive processing refers to modulations of brain activity prior to the actual presentation of the stimulus triggered by, e.g., instruction (Carlsson et al., 2000), specific task cue (Simmons et al., 2004), prior presentation of a stimulus which had become associated with the target stimulus through short-term learning within the same (Schubotz and von Cramon, 2001b) or a different (Widmann et al., 2007) modality, as well as prior presentation of objects which are long-term, e.g., semantically or contextually, related to the target event (Kveraga et al., 2007b). However, the term predictive processing is also used to describe top-down predictions which are initialized after stimulus presentation, thus constituting the activate-predict-confirm perceptive cycle (Enns and Lleras, 2008). These expectations are based on some features of the presented stimulus, e.g., low spatial frequencies, which get

processed fast and become the basis for formulating predictions about the objects' more specific properties which are processed slower (Kveraga et al., 2007a,b). All of these examples indicate that even within specific cognitive domains, predictive processing comes in different flavors and may be very difficult to capture if not all factors underlying its specific occurrence are accounted for.

In an attempt to define unifying mechanisms across different systems, simulation theories of cognition have emphasized the role of internal simulations or emulations not just within one, but across many different cognitive domains, arguing that these can mostly be reduced to covert simulations based on internal models (Hesslow, 2002; Grush, 2004). A crucial representational status with respect to such simulation processes, especially with respect to different aspects of social cognition, e.g., action understanding or language, has been proposed for the so-called "mirror neuron system" (Gallese, 2007). Although potentially appealing, many of such large-scale interpretations which incorporate a wide range of cognitive functions are still somewhat underspecified with respect to the mechanisms and neural architecture underlying different types of processes. In addition, even the terminology which these endorse is still not used in a systematic fashion. For example, simulations and emulations have previously been differentiated with respect to the type of phenomenon (goals and/or algorithms used by the original process) (Moulton and Kosslyn, 2009) or process (related to body or the environment) (Poirier and Hardy-Vallee, 2005) they mimic, as well as brain regions and networks they engage (Grush, 2004). Furthermore, simulations based on internal models which are assumed to incorporate all details related to the mimicked events should not be confused with those which underlie general forecasting of (distant) future events (Gilbert and Wilson, 2007). While the first type of simulations might be implemented in the motor system and used, for example, for understanding the actions of others (Jeannerod, 2001), the latter is based on a somewhat flexible recombination of past episodes and primarily relies on processing within memory-related systems (Schacter et al., 2007).

On a comparable scale of relevance, but much more computationally specified, Friston (2005) proposed a pivotal role of expectations across different levels of neural processing: not only is the common code of brain functioning a predictive one, but our predictions act as a form of self-fulfilling prophecy. According to this view, predictive processing is inherent to all levels of our hierarchically organized neural system. In addition, predictions are suggested to drive our perception, cognition and behavior in a sense that we do not only passively match expected to incoming events and objectively evaluate the accuracy of our expectations, but actively try to fulfill those predictions by preferentially sampling corresponding features in the environment. This implies that the suggested impact of predictions may also be disadvantageous as these may constrain our processing and behavior, always keeping us within the limits defined by our previous experiences. If true, this might suggest that a more reactive strategy based on detailed consideration of all incoming information might qualitatively be more advantageous. However, it has been suggested that this is not the case (Friston and Stephan, 2007). Furthermore, exploration and novelty have rewarding properties of their own (Bunzeck and Duzel, 2006; Knutson and Cooper, 2006) which may result in balancing between exploratory and exploitative behavior across different contexts. In summary, it is becoming more and more clear that anticipation and expectations do not just represent isolated phenomena, but one of the main unitary principles of cognition characteristic not just for humans but also animals, e.g., dogs, snakes or insects (Roitblat and Scopatz, 1983; Rainer et al., 1999; Webb, 2004). This may justify recent claims according to which the mind itself can be conceived of as an anticipatory device (Pezzulo et al., 2007). As described in more detail in the Section "Introduction," such conceptualization is far from behaviorist stimulus–response models of human behavior or later cognitivist metaphors of the mind as a computer, namely a highly efficient, primarily feedforward processing machine.

Mechanisms of prediction

Anticipatory or predictive processing is directed towards the future and, at the same time, highly dependent and grounded in the information from the past. This bridging over different temporal points and taking advantage of the past in order to improve behavior in the future is suggested to be the core capacity which

makes our cognitive brain so efficient (Kveraga et al., 2007a). Given that prediction is inherent to many different levels and types of processes, it is not easy to identify common neural mechanisms supporting such processing across all contexts. While some of these effects could be mediated in a somewhat indirect manner through changes in alertness and attention (Brunia, 1999), most of them should be considered direct. Specifically, prediction is associated with a wide range of neural phenomena within different brain networks, e.g., changes of neuronal threshold in sensory cortices (Gomez et al., 2004), long-range phase synchronization (Gross et al., 2006), changes in connectivity across brain regions (O'Reilly et al., 2008) or existence of preparatory-set cells in the prefrontal or parietal cortex (Quintana and Fuster, 1992). Generally, as will be shown in the following sections, predictive processing has been related to almost all brain regions and networks. Such results should not be surprising, given the wide range of contexts which afford predictions.

Sources and sites of predictions in perception

One way to understand predictive processing in perception is to conceptualize anticipation as a bias signal (Rees and Frith, 1998) which improves the computational efficiency of a specific area. This description may be useful, as it points to three elements which need to be specified in order to understand such a phenomenon: brain regions which formulate expectations and impose such a bias (sources), regions which are influenced by it (sites) and a communication mode mediating this process. Within sites of prediction such as, e.g., relevant sensory cortices, modulations of activity occurring in expectation of a stimulus include a reduction of activation threshold and an increase in signal-to-noise ratio which facilitates subsequent stimulus processing (Brunia, 1999; Gomez et al., 2004). These effects are reflected in the elicitation of particular event-related anticipatory components, e.g., contingent negative variation, stimulus preceding negativity or the readiness potential (Brunia, 1999; Praamstra et al., 2006) and the suppression of specific brain rhythms in the sensory cortices (event-related desynchronization; ERD) as measured using electroencephalography (EEG) (Bastiaansen and Brunia, 2001). Evidence for the claim that improved speed and accuracy of processing expected stimuli reflects preparatory effects in the relevant sensory cortices potentially coupled with the inhibitory effects in other sensory modalities (Brunia, 1999) comes from studies which show comparable patterns of activity in stimulus perception and anticipation. For example, findings showing that actual somatosensory stimulation and anticipation of such stimulation engage the same network (Carlsson et al., 2000) suggest a top-down modulated pre-activation of sensory cortex waiting for the stimulus to occur. A similar pre-activation of areas involved in processing relevant events has been show in other domains, e.g., emotion or pain processing (Porro et al., 2002; Ueda et al., 2003), although not consistently (Bermpohl et al., 2006).

In addition to understanding preparatory effects in relevant sensory cortices, it is important to describe how these effects are initiated and controlled. In an attempt to answer this question, Gomez et al. (2004) suggested that frontomedial cortical areas, namely the supplementary motor area and anterior cingulate cortex, represent the best candidate areas responsible for initiating the process of preparing for perception (and action) by recruiting specific sensory (and motor) cortices needed for subsequent processing. An important role of orbitofrontal as well as medial prefrontal cortex in formulating expectations about incoming visual objects which is crucial for object recognition has also been suggested by Kveraga et al. (2007b) and Summerfield et al. (2006). On the other hand, dorsolateral prefrontal cortex was hypothesized to be implicated in sustaining the activation of the sensory (and motor) cortices (Gomez et al., 2004). Similarly, Brunia (1999) suggested a crucial role of prefrontal cortex in organizing anticipatory behavior by activating cortico-cortical and thalamo-cortical loops to sensory (and motor) areas after the preparatory set had been established. Once formed, the preparatory set could be communicated through changes in brain's oscillatory activity, specifically increases in phase synchronization of neuronal populations in executive areas triggering the increased effective synaptic gain of neurons in target sensory population (Engel et al., 2001). Along these lines, Liang et al. (2002) have shown that synchronized activity in prefrontal cortex during anticipation of a visual stimulus predicts characteristics of early visual processing

and behavioral response. Therefore, these authors argued that synchronized oscillations in prefrontal cortex represent a plausible candidate for sustaining visual anticipation, proposing that such anticipatory control develops as a consequence of accumulating prior experience.

Everything previously mentioned within this section suggests that one useful way of conceptualizing predictive processing in perception may include distinguishing between “sources” which formulate expectations and “sites” to which these are then communicated. And, although this distinction may generally prove to be useful, it may not always be easy to incorporate. For example, it is by now well established that predictive processing is not a phenomenon restricted to two levels of brain processing (one source and one site), but one which occurs across multiple levels of hierarchy. According to the predictive coding model, visual processing can be described as an integration of top-down expectations and bottom-up stimulus information occurring across multiple levels within a hierarchical architecture (Rao and Ballard, 1999; Friston, 2005; Friston and Kiebel, 2009). In this view, top-down expectancy biases are communicated through cortical feedback connections, while feedforward ones convey error signals which indicate the goodness-of-fit of predictions and incoming stimulus information, namely the difference (residual error) between top-down and bottom-up signals. While the importance of prediction errors in this framework will be described in the later sections, at this point it is important to notice that most levels within such a hierarchy can be considered as both sources and sites of predictions. Furthermore, it is not always so that high-level associative areas have to be responsible for generating predictions about the incoming input. As indicated by research related to the auditory mismatch negativity event-related component, predictions generated by representations of rules could also be formulated within a sensory, in this case auditory, system itself (Schröger, 2007; Winkler, 2007).

Internal models mediate prediction across many domains

In addition to the general account presented above, more specific suggestions emphasizing a crucial role of certain systems and regions of the brain, primarily the motor system and especially the cerebellum (Jeannerod, 2001; Wolpert and Flanagan, 2001; Wolpert et al., 2003; Schubotz, 2007), in predictive processing have also been proposed. Functionally, it has been suggested that the prediction of future states of the body or the environment arises from mimicking their respective dynamics through the use of internal models (Johnson-Laird, 1983; Wolpert et al., 1995; Grush, 2004). The internal model approach was originally developed in the motor domain where it went beyond explaining the release of motor commands acting on the musculoskeletal system and introduced another level of computations which essentially entail internal simulations of different aspects of sensorimotor processing (Wolpert et al., 2003) accomplished by internal models. The initial development of the internal model framework was motivated by demonstrations from the experimental work of Sperry, who proposed that a corollary discharge from an action command modulates the visual perception of movement (Sperry, 1950), as well as from von Holst and Mittelstaedt who first described how the discrimination of self- produced and externally applied stimuli may occur through the interaction between sensory feedback signals following an action and an efference copy of the action command (von Holst and Mittelstaedt, 1950). Although addressing somewhat different issues and introducing different terminology, these two findings were the first to demonstrate how the system predicts self-generated sensory signals, an idea which has been greatly pursued in the last decades within the framework of internal models. These models simulate the dynamics of the motor system in order to, in case of inverse models, deduce the motor command which lead to a certain outcome or, in case of forward models, predict the expected sensory consequences of the executed movement (Wolpert and Miall, 1996). The predictive process is initiated by a copy of the motor command, i.e., an efference copy, while the term corollary discharge is typically used to describe the output of the predictor, namely the expected sensory consequences of the produced action. In this context, it has been experimentally demonstrated that expected sensory consequences of self-generated movements get processed in an attenuated fashion both in the auditory and the somatosensory domain (Martikainen et al., 2005; Bäß et al., 2009; Hesse et al., 2010). In contrast, sensory outcomes of self- generated actions which

violate expectations formulated based on motor signals elicit deviance-related event-related potentials of the EEG and cause behavioral delay (Waszak and Herwig, 2007; Iwanaga and Nittono, 2010), indicating that they are processed as deviant events. Importantly, although these effects occur as responses to violations related to different types of movements, it has recently been demonstrated that they are especially accentuated in cases of voluntary actions (Nittono, 2006; Adachi et al., 2007). On a more general level, it has been demonstrated that internal models are, in essence, predictive (Bays et al., 2006) and, in addition to distinguishing between self-generated and externally produced movements, may be used to estimate the current or predict the future state of the system (Miall and Wolpert, 1996) as well as estimate more general context variables (Wolpert and Flanagan, 2001). Computationally, the expectations formulized within the internal models could be optimized in a Bayesian fashion, through weighted combinations of priors and sensory likelihoods (Kording and Wolpert, 2006) and subsequently evaluated through a comparison with the actual sensory input available after the movement (Figure 3).

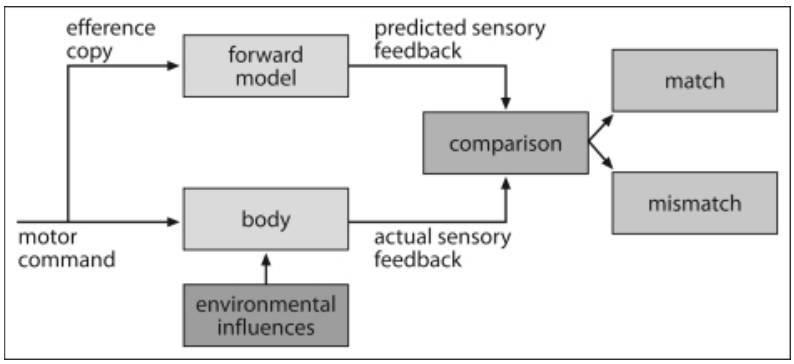

Figure 3

Prediction in motor control. Based on the efference copy of the motor command, a forward model is formulated and used for predicting the consequences of one's own actions. These predictions are compared with the incoming sensory input which can result either in a “match” in case predictions were correctly formulated, or a “mismatch”, signaling an error in prediction.

Although the internal model framework has originally been developed within the motor domain, it has in recent decades proved to be useful for explaining different phenomena well beyond this field. For example, it was recognized rather early that one class of forward models can mimic or approximate some aspects of the environment using the collected sensory knowledge such as, e.g., a trajectory of an already moving object, for predicting its future behavior (Miall and Wolpert, 1996; Wolpert and Kawato, 1998). In a similar fashion, initially motivated by findings implicating the motor system in some forms of perceptual processing (Schubotz and von Cramon, 2002a,b,c, 2003), Schubotz and von Cramon (2003) suggested a joint, so-called sensorimotor forward model, account unifying the perceptual and motor domain. In this view, prediction underlies both motor and perceptual processes in which the brain can emulate expected events, regardless of whether these constitute sensory consequences of one's own actions or expected sensory stimuli. Such emulation is enabled by the creation of internal, sensorimotor or even amodal forward models which can be exploited for making predictions about future states of the modeled space, be that the body or the environment (Schubotz, 2007). Although suggesting that the prediction of both internal and external events can be supported through highly comparable computations implemented within the motor system, this view does not automatically assume that the models supporting perceptual and motor processing should be completely identical. While motor processing requires development of highly accurate and precise models (Blakemore et al., 1998; Miall, 2003), in perception such high precision may

be either unnecessary since accurate prediction can often rely only on relational properties of external events, or even disadvantageous because it occurs in a noisy system and environment. This suggestion is in line with the hyper-MOSAIC model (Wolpert et al., 2003) which proposes an architecture containing several levels of forward models differing in the level of specificity and function, thus providing a general framework for understanding prediction in a wide range or high level cognitive functions including action observation, imitation, mental practice, social interaction and the theory of mind.

The brain as a prediction device

In the previous sections different parts of the brain have been associated with predictive processing, specifically different sensory cortices, the thalamus, the prefrontal cortex and the motor system. These sections described only a subset of contexts which afford predictability, most of which were limited to short timescales. In addition, it is important to mention that a pivotal role in prediction on longer timescales can be associated with the prefrontal cortex which is, together with medial temporal regions, especially the hippocampus (Eichenbaum and Fortin, 2009; Lisman and Redish, 2009), and posterior cerebral cortices (including the lateral parietal and temporal regions, the precuneus and the retrosplenial cortex), crucial for imagining the future as well as remembering the past (Schacter et al., 2007; Schacter and Addis, 2009). However, although this region is also typically considered as the key region implicated in planning (Fuster, 1997), the contributions of the parietal cortices should also be acknowledged in this context (Ruby et al., 2002). Furthermore, a more central role of lateral parietal, together with premotor regions can be posited for formulating temporal expectations (Coull and Nobre, 2008; Coull, 2009). In addition, it is important to mention other brain regions which have been associated with predictive processing, e.g., the basal ganglia (Schultz and Dickinson, 2000; Flesischer, 2007; Kotz et al., 2009) and especially the ventral striatum in reward prediction (Knutson and Cooper, 2005) or amygdala, insula and the anterior cingulate cortex in pain or emotional processing (Ploghaus et al., 1999; Porro et al., 2003; Ueda et al., 2003). Not questioning the validity of these or accounts previously specified, it is still important to note one danger which can be associated with considering all of these accounts together, without clearly specifying the type of predictive processing they refer to. Specifically, if one was to try and summarize all brain areas which have so far been mentioned as incorporating some aspect of predictive processing, these would include: unimodal sensory cortices, lateral and medial parietal and temporal areas, orbitofrontal, medial frontal and dorsolateral prefrontal cortex, premotor cortex, insula, cerebellum, basal ganglia, amygdala and thalamus (Figure 4). In other words, the whole brain. And, while it may be true that different aspects of prediction can be captured across the whole brain or nervous system itself, it does not imply that they share an equivalent role.

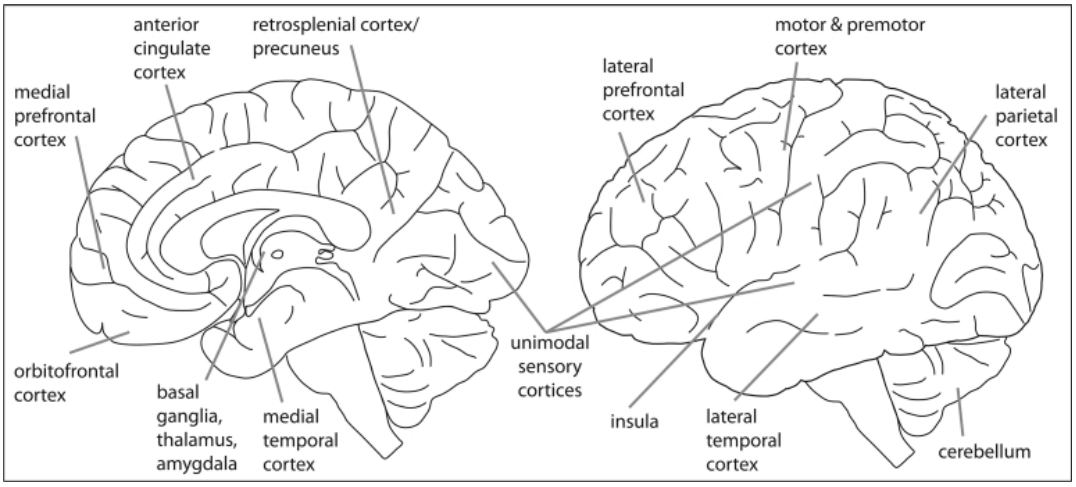

Figure 4

The predictive brain. Numerous brain areas and networks have previously been associated with some aspect of prediction, thus illustrating the notion of the "predictive brain". It is important to note that this outline is simplified as it lacks both anatomical and functional specificity. Nevertheless, it may be useful as an illustration and an invitation to further elaborate specific roles of the marked regions.

In summary, there are different ways of conceptualizing and differentiating the role of different brain areas in prediction. One way is to differentiate between sources and sites of predictions, as shown in the example of perception. In this view, higher-level areas such as lateral, medial, orbital prefrontal and premotor regions could be considered as sources which formulate expectations and communicate them to lower-level, typically sensory areas. However, as previously elaborated, this view can only be considered as a very rough simplification. Furthermore, it may be useful to consider numerous dimensions which have previously been discussed and suggested to be relevant in defining the nature of predictive phenomena. This includes, for example, the timescale of prediction, as some regions may be more relevant for short-term, e.g., the premotor cortex, in contrast to others which are important for prediction across different timescales, such as, e.g., the prefrontal cortex. In addition, the cognitive domain, e.g., emotional, perceptual or motor, or the nature of predicted features, e.g., object identity or the context within which it is usually encountered, can be considered as crucial in determining the sources within which such expectations are formulated. An alternative way of determining which levels and types of predictions are associated with certain brain areas is to specify more holistic models. Previously described predictive coding model can be viewed as such. Within this framework the brain is seen as a "Bayesian inference machine", constantly building models of the environment and the body, allowing the brain to predict their respective future states (Knill and Pouget, 2004; Friston and Stephan, 2007). Importantly, such general nature of brain processing can then account for many phenomena across domains and processes, e.g., perception, attention, action or learning (Friston, 2005; Friston and Stephan, 2007). An important aspect of this and other models of prediction relates to testing the validity of posited expectations by comparing them to the realized events. Potential outcomes of such testing will be described in the following section.

When Predictions Meet Reality

Although the process of formulating expectations is interesting in its own right, it is also quite fascinating to consider what happens once the external event occurs, especially in cases where it does not meet the initial expectations. In the previous sections it was suggested that expected stimuli (matches) are processed

in a more efficient manner than the unexpected ones (mismatches), as indicated by more accurate and faster reactions to these events. However, efficiency should not be confused with relevance or associated priority. On the contrary, given that these represent pure confirmations of correctly formulated expectations and signal correct learning, matches carry little informational value and are therefore not relevant for the system. Consequently, an expected event does not need to be explicitly represented or communicated to higher cortical areas which have processed all of its relevant features prior to its occurrence. In contrast, errors of prediction have much greater value (Friston and Stephan, 2007), as they may signal unsuccessful learning, a major change in the surroundings or noise and smaller changes in the body or the environment, corresponding to normal (plant or world) drifts which typically occur over time (Grush, 2004). Therefore, registering and further processing events which deviate from predictions is important, but “costly” as they draw attentional resources needed in order to check their behavioral relevance (Corbetta et al., 2002) which will determine their subsequent treatment. For example, errors of prediction which are irrelevant for the current mental set or reflect noise in the environment can be registered and ignored, allowing the individual to reorient himself to the task at hand (Escera et al., 2000; Corbetta et al., 2002). However, when these are relevant and informative, e.g., in situations where one fails to learn the relevant contingencies or the environment suddenly changes, they can trigger an update of one's knowledge (Winkler et al., 1996; Winkler and Czigler, 1998) and behavioral adaptations. Therefore, the cost associated with processing these events may in the end turn to be beneficial, as it can lead to an adaptive reaction to the changing environment. Such significance of deviant events for cognitive processing and behavior is reflected on the level of our nervous system which is highly sensitive to novel events, changes in the environment and other types of errors in prediction (Corbetta and Shulman, 2002; Friston et al., 2006).

Importantly, not only are novel or unexpected events preferentially detected, but also encoded, as demonstrated by the identified novelty advantage in memory (Knight and Nakada, 1998; Kishiyama et al., 2009). This may explain why prediction errors or the discrepancies between expected and realized events have been postulated as one of the main learning forces. Specifically, associative learning theories (Rescorla and Wagner, 1972; Schultz et al., 1997) describe how taking into account differences between the predicted and actual outcomes promotes learning and postulate how the size of the prediction error affects the rate of forming associations between events. At the neuronal level, these discrepancies can be translated into changes in synaptic weights using specific learning computational rules, leading to changes in the model and subsequent more accurate predictions (Wolpert et al., 2003). Neurons in different brain structures have been shown to code prediction errors stemming from different sources, e.g., rewards, punishment, external stimuli and own behavior, a process which in some contexts may be mediated through the dopaminergic and norepinergic pathways (Schultz and Dickinson, 2000). In addition to this direct link between prediction errors and learning, a somewhat more indirect one may be mediated through increased attentional resources being diverted towards the perceived prediction error (Wills et al., 2007) or their high emotional significance (Frey et al., 2009). On a somewhat different note, although it has previously been suggested that errors in behavior could be organized hierarchically (Krigolson and Holroyd, 2007), it is not clear what such hierarchical structure includes and whether different levels of hierarchy may somehow interact. Interestingly, it has also been shown that the detection of semantic violations in language might be somehow restricted by the processing of syntactic structure (Friederici et al., 1999) and that different types of deviants in visual sequences may be processed through different mechanisms (Koester and Prinz, 2007). This line of research comparing and mutually relating different types and sources of errors will surely become more and more important in the future as it may reveal interesting and important insights about both regular and violated predictive processing within and across different contexts.

The question of how errors of prediction are processed online relates strongly to the general issue of the integration of top-down and bottom-up information which has been posited to rely on error-minimalization mechanisms (Grossberg, 1980; Mumford, 1992; Ullman, 1995; Friston, 2005; Kveraga et al., 2007b). According to the predictive view, expectations mediated through feedback connections represent top-down

information which are compared and integrated with bottom-up signals communicated through feedforward connections, a process accomplished through specific synchronization patterns visible across different levels of the hierarchy (Kveraga et al., 2007b) and changes in connectivity between relevant regions (den Ouden et al., 2009). It has already been described that mismatches which are detected through such a comparison elicit more pronounced responses which get communicated to the next level in the hierarchy using feedforward connections. The size of such mismatches (prediction error) is suggested to reflect surprise which the brain tries to minimize in order to maintain present and future stability (Friston and Stephan, 2007). In contrast, matches produce non-salient responses and their overall processing is suppressed. In this view, postulated predictions act as a form of perceptual filter, as their accuracy determines which information is suppressed at an earlier processing stage (match) and what is communicated to a higher level (mismatch). It has been suggested that this conceptualization may be incompatible with current theories of attention which posit an enhancement of stimulus-driven activity that it is consistent with top-down bias communicated through feedback connections (Desimone and Duncan, 1995; Summerfield and Egner, 2009). However, it has recently been demonstrated that this may not be the case, as the predictive coding model can be considered mathematically equivalent with a particular form of biased competition model of attention (Spratling, 2008a,b).

It is plausible to expect that the near future will being a formulation of an unifying framework bridging seemingly contradictory attentional and predictive phenomena, given that both of these reflect comparable processing biases implemented within the same hierarchical brain architecture. An additional open issue concerns the differences in dynamics of processing events which confirm and violate previous expectations. Although it was previously mentioned that more elaborate processing should follow the presentation of mismatches, Summerfield and Koechlin (2008) have suggested a more refined hypothesis according to which match-suppression should occur in lower-level hierarchical areas in contrast to match-enhancement which is to be expected in higher-level regions. In accordance with this, these authors demonstrated how processing expected stimuli preferentially engages ventral prefrontal and orbitofrontal cortex (Summerfield and Koechlin, 2008). However, the importance of these regions has previously been identified in the completely opposite context of detecting violations of expectations (Nobre et al., 1999; Petersson et al., 2004; Petrides, 2007), leaving this issue to be settled in the future. And, although it has been suggested that the search of representation (prediction/match relevant) and error (mismatch signaling) neurons will be an important challenge for the future (Summerfield and Egner, 2009), it is still not clear whether error-codes signaling a breach of expectations necessarily have to be implemented within single neurons, or if such signals could be implemented within dendrites of certain neuronal populations (Spratling, 2008b). In addition to identifying neural regions which preferentially process matches and mismatches, future research may benefit from investigating neural synchrony between relevant cortical regions across different levels of hierarchy (Kveraga et al., 2007b) as a potential complementary signature of (un)successful matching between top-down and bottom-up information. Clearly, much more research will be needed in order to clarify these issues.

Concluding Remarks

In conclusion, anticipatory or predictive processing potentially reflects one of the core, fundamental principles of brain functioning which justifies the notion of “the predictive brain”. Even if this statement is too strong, the relevance of prediction in cognitive and neural processing can still not be overestimated. Prediction allows us to direct our behavior towards the future, while remaining well-grounded and guided by the information pertaining to the present and the past. Furthermore, predictive processing represents one of the key features of many cognitive functions and is mediated through a wide selection of mechanisms expressed in numerous cortical and subcortical levels. Such benefits are widely acknowledged and have in recent years been greatly investigated.

However, although a lot is known about this type of processing, numerous open questions remain. Many of these can be identified with respect to each specific account or model incorporating or positing predictive

mechanisms. Even more importantly, it seems even more difficult to reconciliate manifold views and theories which emphasize the importance of different brain systems implicated in predictive processing or bring together findings stemming from different domains, especially those aimed at exploring expectations of various types, specificities and temporal structure. An even bigger challenge will include comparing and reconciling predictive with non-predictive mechanisms of brain and cognitive functioning: as mentioned before, predictive mechanisms are rarely contrasted with clearly opposing approaches and some predictive phenomena may be rather similar to functions which have traditionally not been seen as predictive. Bringing these together will be an important experimental and theoretical task for the future.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

The authors thank Heike Schmidt-Duderstedt for help with the figures.

References

1. Adachi S., Morikawa K., Nittono H. (2007). Event-related potentials elicited by unexpected visual stimuli after voluntary actions. Int. J. Psychophysiol. 66, 238–24310.1016/j.ijpsycho.2007.08.004 [PubMed] [CrossRef] [Google Scholar]
2. Ashe J., Lungu O. V., Basford A. T., Lu X. (2006). Cortical control of motor sequences. Curr. Opin. Neurobiol. 16, 213–22110.1016/j.conb.2006.03.008 [PubMed] [CrossRef] [Google Scholar]
3. Bäß P., Widmann A., Roye A., Schroger E., Jacobsen T. (2009). Attenuated human auditory middle latency response and evoked 40-Hz response to self-initiated sounds. Eur. J. Neurosci. 29, 1514–152110.1111/j.1460-9568.2009.06683.x [PubMed] [CrossRef] [Google Scholar]
4. Bahlmann J., Schubotz R. I., Friederici A. D. (2008). Hierarchical artificial grammar processing engages Broca's area. Neuroimage 42, 525–53410.1016/j.neuroimage.2008.04.249 [PubMed] [CrossRef] [Google Scholar]
5. Baker S. C., Rogers R. D., Owen A. M., Frith C. D., Dolan R. J., Frackowiak R. S., Robbins T. W. (1996). Neural systems engaged by planning: a PET study of the Tower of London task. Neuropsychologia 34, 515–52610.1016/0028-3932(95)00133-6 [PubMed] [CrossRef] [Google Scholar]
6. Bapi R. S., Pammi V. S. C., Ahmed M. K. P. (2005). Investigation of sequence processing: a cognitive and computational neuroscience perspective. Curr. Sci. 89, 1690–1698 [Google Scholar]
7. Bar M. (2004). Visual objects in context. Nat. Rev. Neurosci. 5, 617–62910.1038/nrn1476 [PubMed] [CrossRef] [Google Scholar]
8. Bar M. (2007). The proactive brain: using analogies and associations to generate predictions. Trends Cogn. Sci. 11, 280–28910.1016/j.tics.2007.05.005 [PubMed] [CrossRef] [Google Scholar]
9. Bar M. (2009). The proactive brain: memory for predictions. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 364, 1235–124310.1098/rstb.2008.0310 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
10. Bastiaansen M. C., Brunia C. H. (2001). Anticipatory attention: an event-related desynchronization approach. Int. J. Psychophysiol. 43, 91–10710.1016/S0167-8760(01)00181-7 [PubMed] [CrossRef] [Google Scholar]
11. Bays P. M., Flanagan J. R., Wolpert D. M. (2006). Attenuation of self- generated tactile sensations is predictive, not postdictive. PLoS Biol. 4, e28.10.1371/journal.pbio.0040028 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
12. Bermpohl F., Pascual-Leone A., Amedi A., Merabet L. B., Fregni F., Gaab N., Alsop D., Schlaug G., Northoff G. (2006). Dissociable networks for the expectancy and perception of emotional stimuli in

the human brain. Neuroimage 30, 588–60010.1016/j.neuroimage.2005.09.040 [PubMed] [CrossRef] [Google Scholar]
13. Blakemore S. J., Rees G., Frith C. D. (1998). How do we predict the consequences of our actions? A functional imaging study. Neuropsychologia 36, 521–52910.1016/S0028-3932(97)00145-0 [PubMed] [CrossRef] [Google Scholar]
14. Bridgeman B. (2007). Efference copy and its limitations. Comput. Biol. Med. 37, 924–92910.1016/j.compbiomed.2006.07.001 [PubMed] [CrossRef] [Google Scholar]
15. Brunia C. H. M. (1999). Neural aspects of anticipatory behavior. Acta Psychol. 101, 213–24210.1016/S0001-6918(99)00006-2 [PubMed] [CrossRef] [Google Scholar]
16. Bunzeck N., Duzel E. (2006). Absolute coding of stimulus novelty in the human substantia nigra/VTA. Neuron 51, 369–37910.1016/j.neuron.2006.06.021 [PubMed] [CrossRef] [Google Scholar]
17. Butz M. V., Pezzulo G. (2008). “Benefits of anticipation in cognitive agents,” in The Challenge of Anticipation, eds Pezzulo G., Butz M. V., Castelfranchi C., editors. (Berlin: Springer;), 45–62 [Google Scholar]
18. Butz M. V., Sigaud O., Gerard P. (2003). “Anticipatory behavior: exploring knowledge about the future to improve current behavior,” in Anticipatory Behavior in Adaptive Learning Systems, eds Butz M. V., Sigaud O., Gerard P., editors. (Berlin: Springer;), 1–10 [Google Scholar]
19. Carlsson K., Petrovic P., Skare S., Petersson K. M., Ingvar M. (2000). Tickling anticipations: neural processing in anticipation of a sensory stimulus. J. Cogn. Neurosci. 12, 691–70310.1162/089892900562318 [PubMed] [CrossRef] [Google Scholar]
20. Cleeremans A., McClelland J. L. (1991). Learning the structure of event sequences. J. Exp. Psychol. Gen. 120, 235–25310.1037/0096-3445.120.3.235 [PubMed] [CrossRef] [Google Scholar]
21. Clegg B. A., DiGirolamo G. J., Keele S. W. (1998). Sequence learning. Trends Cogn. Sci. 2, 275–28110.1016/S1364-6613(98)01202-9 [PubMed] [CrossRef] [Google Scholar]
22. Cohen A., Ivry R. I., Keele S. W. (1990). Attention and structure in sequence learning. J. Exp. Psychol. Learn Mem. Cogn. 16, 17–3010.1037/0278-7393.16.1.17 [CrossRef] [Google Scholar]
23. Conway C. M., Christiansen M. H. (2001). Sequential learning in non-human primates. Trends Cogn. Sci. 5, 539–54610.1016/S1364-6613(00)01800-3 [PubMed] [CrossRef] [Google Scholar]
24. Corbetta M., Kincade J. M., Shulman G. L. (2002). Neural systems for visual orienting and their relationships to spatial working memory. J. Cogn. Neurosci. 14, 508–52310.1162/089892902317362029 [PubMed] [CrossRef] [Google Scholar]
25. Corbetta M., Shulman G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. Nat. Rev. Neurosci. 3, 201–21510.1038/nrn755 [PubMed] [CrossRef] [Google Scholar]
26. Coull J. T. (2009). Neural substrates of mounting temporal expectation. PLoS Biol. 7, e1000166.10.1371/journal.pbio.1000166 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
27. Coull J. T., Nobre A. C. (2008). Dissociating explicit timing from temporal expectation with fMRI. Curr. Opin. Neurobiol. 18, 137–14410.1016/j.conb.2008.07.011 [PubMed] [CrossRef] [Google Scholar]
28. Dehaene S., Changeux J. P. (1997). A hierarchical neuronal network for planning behavior. Proc. Natl. Acad. Sci. U.S.A. 94, 13293–1329810.1073/pnas.94.24.13293 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
29. DeLong K. A., Urbach T. P., Kutas M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. Nat. Neurosci. 8, 1117–112110.1038/nn1504 [PubMed] [CrossRef] [Google Scholar]
30. den Ouden H. E., Friston K. J., Daw N. D., McIntosh A. R., Stephan K. E. (2009). A dual role for prediction error in associative learning. Cereb. Cortex 19, 1175–118510.1093/cercor/bhn161 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
31. Desimone R., Duncan J. (1995). Neural mechanisms of selective visual attention. Annu. Rev. Neurosci. 8, 193–22210.1146/annurev.ne.18.030195.001205 [PubMed] [CrossRef] [Google Scholar]

32. Eagleman D. M., Sejnowski T. J. (2000). Motion integration and postdiction in visual awareness. Science 287, 2036–203810.1126/science.287.5460.2036 [PubMed] [CrossRef] [Google Scholar]
33. Eichenbaum H., Fortin N. J. (2009). The neurobiology of memory based predictions. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 364, 1183–119110.1098/rstb.2008.0306 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
34. Engel A. K., Fries P., Singer W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. Nat. Rev. Neurosci. 2, 704–71610.1038/35094565 [PubMed] [CrossRef] [Google Scholar]
35. Enns J. T., Lleras A. (2008). What's next? New evidence for prediction in human vision. Trends Cogn. Sci. 12, 327–33310.1016/j.tics.2008.06.001 [PubMed] [CrossRef] [Google Scholar]
36. Escera C., Alho K., Schröger E., Winkler I. (2000). Involuntary attention and distractibility as evaluated with event-related brain potentials. Audiol. Neurootol. 5, 151–16610.1159/000013877 [PubMed] [CrossRef] [Google Scholar]
37. Fenske M. J., Aminoff E., Gronau N., Bar M. (2006). Top-down facilitation of visual object recognition: object-based and context-based contributions. Prog. Brain Res. 155, 3–2110.1016/S0079-6123(06)55001-0 [PubMed] [CrossRef] [Google Scholar]
38. Fiebach C. J., Schubotz R. I. (2006). Dynamic anticipatory processing of hierarchical sequential events: a common role for Broca's area and ventral premotor cortex across domains? Cortex 42, 499–50210.1016/S0010-9452(08)70386-1 [PubMed] [CrossRef] [Google Scholar]
39. Fitch W. T., Hauser M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. Science 303, 377–38010.1126/science.1089401 [PubMed] [CrossRef] [Google Scholar]
40. Flesischer J. G. (2007). “Neural correlates of anticipation in cerebellum, basal ganglia, and hippocampus,” in Anticipatory Behavior in Adaptive Learning Systems: From Brains to Individual and Social Behavior, eds Butz M. V., Siguard O., Baldassarre G., Pezzulo G., editors. (Heidelberg: Springer;), 19–34 [Google Scholar]
41. Frey S., Zlatkina V., Petrides M. (2009). Encoding touch and the orbitofrontal cortex. Hum. Brain Mapp. 30, 650–65910.1002/hbm.20532 [PubMed] [CrossRef] [Google Scholar]
42. Friederici A. D., Steinhauer K., Frisch S. (1999). Lexical integration: Sequential effects of syntactic and semantic information. Mem. Cogn. 27, 438–453 [PubMed] [Google Scholar]
43. Friedman S. L., Scholnick E. K. (1997). The Developmental Psychology of Planning: Why, How, and When Do We Plan? Mahwah, NJ: Lawrence Erlbaum Associates [Google Scholar]
44. Friston K. J. (2005). A theory of cortical responses. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 360, 815–83610.1098/rstb.2005.1622 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
45. Friston K. J., Kiebel S. J. (2009). Predictive coding under the free-energy principle. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 364, 1211–122110.1098/rstb.2008.0300 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
46. Friston K. J., Kilner J., Harrison L. (2006). A free energy principle for the brain. J. Physiol. Paris 100, 70–8710.1016/j.jphysparis.2006.10.001 [PubMed] [CrossRef] [Google Scholar]
47. Friston K. J., Stephan K. E. (2007). Free-energy and the brain. Synthese 159, 417–45810.1007/s11229-007-9237-y [PMC free article] [PubMed] [CrossRef] [Google Scholar]
48. Frith C. D., Frith U. (2006). How we predict what other people are going to do. Brain Res. 1079, 36–4610.1016/j.brainres.2005.12.126 [PubMed] [CrossRef] [Google Scholar]
49. Fuster J. M. (1997). The Prefrontal Cortex. Anatomy, Physiology, and Neuropsychology of the Frontal Lobe. Philadelphia: Lippincott Raven [Google Scholar]
50. Fuster J. M. (2001). The prefrontal cortex – an update: time is of the essence. Neuron 30, 319–33310.1016/S0896-6273(01)00285-9 [PubMed] [CrossRef] [Google Scholar]
51. Gallese V. (2007). Before and below ‘theory of mind’: embodied simulation and the neural correlates of social cognition. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 362, 659–66910.1098/rstb.2006.2002 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
52. Gilbert D. T., Wilson T. D. (2007). Prospection: experiencing the future. Science 317,

1351–135410.1126/science.1144161 [PubMed] [CrossRef] [Google Scholar]
53. Gilbert D. T., Wilson T. D. (2009). Why the brain talks to itself: sources of error in emotional prediction. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 364, 1335–134110.1098/rstb.2008.0305 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
54. Gomez C. M., Vaquero E., Vazquez-Marrufo M. (2004). A neurocognitive model for short-term sensory and motor preparatory activity in humans. Psicologica 25, 217–229 [Google Scholar]
55. Gracian B. (1991). The Art of Worldly Wisdom. New York: Broadway Business [Google Scholar]
56. Grafton S. T., Hamilton A. F. (2007). Evidence for a distributed hierarchy of action representation in the brain. Mov. Sci. 26, 590–61610.1016/j.humov.2007.05.009 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
57. Gross J., Schmitz F., Schnitzler I., Kessler K., Shapiro K., Hommel B., Schnitzler A. (2006). Anticipatory control of long-range phase synchronization. Eur. J. Neurosci. 24, 2057–206010.1111/j.1460-9568.2006.05082.x [PubMed] [CrossRef] [Google Scholar]
58. Grossberg S. (1980). How does a brain build a cognitive code? Psychol. Rev. 87, 1–5110.1037/0033-295X.87.1.1 [PubMed] [CrossRef] [Google Scholar]
59. Grush R. (2004). The emulation theory of representation: motor control, imagery, and perception. Behav. Brain Sci. 27, 377–39610.1017/S0140525X04000093 [PubMed] [CrossRef] [Google Scholar]
60. Grüsser O. J. (1995). On the history of the ideas of efference copy and reafference. Clio Med. 33, 35–55 [PubMed] [Google Scholar]
61. Herwig U., Baumgartner T., Kaffenberger T., Bruhl A., Kottlow M., Schreiter-Gasser U., Abler B., Jancke L., Rufer M. (2007). Modulation of anticipatory emotion and perception processing by cognitive control. Neuroimage 37, 652–66210.1016/j.neuroimage.2007.05.023 [PubMed] [CrossRef] [Google Scholar]
62. Hesse M. D., Nishitani N., Fink G. R., Jousmaki V., Hari R. (2010). Attenuation of somatosensory responses to self-produced tactile stimulation. Cereb. Cortex 20, 425–43210.1093/cercor/bhp110 [PubMed] [CrossRef] [Google Scholar]
63. Hesslow G. (2002). Conscious thought as simulation of behavior and perception. Trends Cogn. Sci. 6, 242–24710.1016/S1364-6613(02)01913-7 [PubMed] [CrossRef] [Google Scholar]
64. Hoen M., Pachot-Clouard M., Segebarth C., Dominey P. F. (2006). When Broca experiences the Janus syndrome: an ER-fMRI study comparing sentence comprehension and cognitive sequence processing. Cortex 42, 605–62310.1016/S0010-9452(08)70398-8 [PubMed] [CrossRef] [Google Scholar]
65. Hoffmann J., Berner M., Butz M. V., Herbort O., Kiesel A., Kunde W., Lenhard A. (2007). Explorations of anticipatory behavioral control (ABC): a report from the Cognitive Psychology Unit of the University of Wurzburg. Cogn. Process. 8, 133–14210.1007/s10339-007-0166-y [PubMed] [CrossRef] [Google Scholar]
66. Hommel B., Musseler J., Aschersleben G., Prinz W. (2001). The theory of event coding (TEC): a framework for perception and action planning. Behav. Brain Sci. 24, 849–878; discussion 878–937. [PubMed] [Google Scholar]
67. Ingvar D. H. (1985). “Memory of the future”: an essay on the temporal organization of conscious awareness. Hum. Neurobiol. 4, 127–136 [PubMed] [Google Scholar]
68. Iwanaga M., Nittono H. (2010). Unexpected action effects elicit deviance-related brain potentials and cause behavioral delay. Psychophysiology 47, 281–28810.1111/j.1469-8986.2009.00937.x [PubMed] [CrossRef] [Google Scholar]
69. James W. (1890). The Principles of Psychology, Vol. 1 New York: Dover Publications [Google Scholar]
70. Janata P., Grafton S. T. (2003). Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. Nat. Neurosci. 6, 682–68710.1038/nn1081 [PubMed] [CrossRef] [Google Scholar]

71. Jeannerod M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. Neuroimage 14, S103–S10910.1006/nimg.2001.0832 [PubMed] [CrossRef] [Google Scholar]
72. Johnson-Laird P. (1983). Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness. Cambridge: Cambridge University Press/Harvard University Press [Google Scholar]
73. Jubault T., Ody C., Koechlin E. (2007). Serial organization of human behavior in the inferior parietal cortex. J. Neurosci. 27, 11028–1103610.1523/JNEUROSCI.1986-07.2007 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
74. Keele S. W., Ivry R., Mayr U., Hazeltine E., Heuer H. (2003). The cognitive and neural architecture of sequence representation. Psychol. Rev. 110, 316–33910.1037/0033-295X.110.2.316 [PubMed] [CrossRef] [Google Scholar]
75. Keller P. E., Koch I. (2008). Action planning in sequential skills: relations to music performance. Q. J. Exp. Psychol. 61, 275–29110.1080/17470210601160864 [PubMed] [CrossRef] [Google Scholar]
76. Kerzel D. (2005). Representational momentum beyond internalized physics – embodied mechanisms of anticipation cause errors in visual short-term memory. Curr. Dir. Psychol. Sci. 14, 180–18410.1111/j.0963-7214.2005.00360.x [CrossRef] [Google Scholar]
77. Kerzel D., Bekkering H., Wohlschlager A., Prinz W. (2000). Launching the effect: Representations of causal movements are influenced by what they lead to. Q. J. Exp. Psychol. 53, 1163–118510.1080/02724980050156344 [PubMed] [CrossRef] [Google Scholar]
78. Kerzel D., Gegenfurtner K. R. (2003). Neuronal processing delays are compensated in the sensorimotor branch of the visual system. Curr. Biol. 13, 1975–197810.1016/j.cub.2003.10.054 [PubMed] [CrossRef] [Google Scholar]
79. Kiebel S. J., Daunizeau J., Friston K. J. (2008). A hierarchy of time-scales and the brain. PLoS Comput. Biol. 5, e1000209.10.1371/journal.pcbi.1000209 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
80. Kiebel S. J., Daunizeau J., Friston K. J. (2009). Perception and hierarchical dynamics. Front. Neuroinform. 3:20.10.3389/neuro.11.020.2009 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
81. Kilner J. M., Friston K. J., Frith C. D. (2007). Predictive coding: an account of the mirror neuron system. Cogn. Process. 8, 159–16610.1007/s10339-007-0170-2 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
82. Kishiyama M. M., Yonelinas A. P., Knight R. T. (2009). Novelty enhancements in memory are dependent on lateral prefrontal cortex. J. Neurosci. 29, 8114–811810.1523/JNEUROSCI.5507-08.2009 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
83. Knight R. T., Nakada T. (1998). Cortico-limbic circuits and novelty: a review of EEG and blood flow data. Rev. Neurosci. 9, 57–70 [PubMed] [Google Scholar]
84. Knill D. C., Pouget A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci. 27, 712–71910.1016/j.tins.2004.10.007 [PubMed] [CrossRef] [Google Scholar]
85. Knutson B., Cooper J. C. (2005). Functional magnetic resonance imaging of reward prediction. Curr. Opin. Neurol. 18, 411–41710.1097/01.wco.0000173463.24758.f6 [PubMed] [CrossRef] [Google Scholar]
86. Knutson B., Cooper J. C. (2006). The lure of the unknown. Neuron 51, 280–28210.1016/j.neuron.2006.07.017 [PubMed] [CrossRef] [Google Scholar]
87. Koechlin E., Corrado G., Pietrini P., Grafman J. (2000). Dissociating the role of the medial and lateral anterior prefrontal cortex in human planning. Proc. Natl. Acad. Sci. U.S.A. 97, 7651–765610.1073/pnas.130177397 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
88. Koester D., Prinz W. (2007). Capturing regularities in event sequences: evidence for two mechanisms. Brain Res. 1180, 59–7710.1016/j.brainres.2007.08.056 [PubMed] [CrossRef] [Google Scholar]

89. Kording K. P., Wolpert D. M. (2006). Bayesian decision theory in sensorimotor control. Trends Cogn. Sci. 10, 319–32610.1016/j.tics.2006.05.003 [PubMed] [CrossRef] [Google Scholar]
90. Kotz S. A., Schwartze M., Schmidt-Kassow M. (2009). Non-motor basal ganglia functions: a review and proposal for a model of sensory predictability in auditory language perception. Cortex 45, 982–99010.1016/j.cortex.2009.02.010 [PubMed] [CrossRef] [Google Scholar]
91. Krigolson O. E., Holroyd C. B. (2007). Hierarchical error processing: different errors, different systems. Brain Res. 1155, 70–8010.1016/j.brainres.2007.04.024 [PubMed] [CrossRef] [Google Scholar]
92. Kunde W., Elsner K., Kiesel A. (2007). No anticipation-no action: the role of anticipation in action and perception. Cogn. Process. 8, 71–7810.1007/s10339-007-0162-2 [PubMed] [CrossRef] [Google Scholar]
93. Kveraga K., Boshyan J., Bar M. (2007a). Magnocellular projections as the trigger of top-down facilitation in recognition. J. Neurosci. 27, 13232–1324010.1523/JNEUROSCI.3481-07.2007 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
94. Kveraga K., Ghuman A. S., Bar M. (2007b). Top-down predictions in the cognitive brain. Brain Cogn. 65, 145–16810.1016/j.bandc.2007.06.007 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
95. LaBerge D. (1995). Attentional Processing: The Brain's Art of Mindfulness. Cambridge: Harvard University Press [Google Scholar]
96. Lashley K. S. (1951). “The problem of serial order in behavior,” in Cerebral Mechanisms in Behavior, ed. Jeffries L. A., editor. (New York, NY: John Wiley & Sons;), 112–136 [Google Scholar]
97. Lelekov-Boissard T., Dominey P. F. (2002). Human brain potentials reveal similar processing of non- linguistic abstract structure and linguistic syntactic structure. Clin. Neurophysiol. 32, 72–8410.1016/S0987-7053(01)00291-X [PubMed] [CrossRef] [Google Scholar]
98. Liang H., Bressler S. L., Ding M., Truccolo W. A., Nakamura R. (2002). Synchronized activity in prefrontal cortex during anticipation of visuomotor processing. Neuroreport 13, 2011–201510.1097/00001756-200211150-00004 [PubMed] [CrossRef] [Google Scholar]
99. Lisman J., Redish A. D. (2009). Prediction, sequences and the hippocampus. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 364, 1193–120110.1098/rstb.2008.0316 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
100. Llinas R. R. (2002). I of the Vortex: From Neurons to Self. Cambridge: MIT Press [Google Scholar]
101. MacKay D. (1956). Towards an information-flow model of human behaviour. Br. J. Psychol. 43, 30–43 [PubMed] [Google Scholar]
102. Martikainen M. H., Kaneko K., Hari R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. Cereb. Cortex 15, 299–30210.1093/cercor/bhh131 [PubMed] [CrossRef] [Google Scholar]
103. Mehta B., Schaal S. (2002). Forward models in visuomotor control. J. Neurophysiol. 88, 942–953 [PubMed] [Google Scholar]
104. Miall R. C. (2003). Connecting mirror neurons and forward models. Neuroreport 14, 2135–213710.1097/00001756-200312020-00001 [PubMed] [CrossRef] [Google Scholar]
105. Miall R. C., Wolpert D. M. (1996). Forward models for physiological motor control. Neural. Netw. 9, 1265–127910.1016/S0893-6080(96)00035-4 [PubMed] [CrossRef] [Google Scholar]
106. Morris R., Ward G. (2005). The Cognitive Psychology of Planning. Hove: Psychology Press [Google Scholar]
107. Moulton S. T., Kosslyn S. M. (2009). Imagining predictions: mental imagery as mental emulation. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 364, 1273–128010.1098/rstb.2008.0314 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
108. Mumford D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. Biol. Cybern. 66, 241–25110.1007/BF00198477 [PubMed] [CrossRef]

[Google Scholar]
109. Näätanen R., Tervaniemi M., Sussman E., Paavilainen P., Winkler I. (2001). “Primitive intelligence” in the auditory cortex. Trends Neurosci. 24, 283–28810.1016/S0166-2236(00)01790-2 [PubMed] [CrossRef] [Google Scholar]
110. Nijhawan R. (1997). Visual decomposition of colour through motion extrapolation. Nature 386, 66–6910.1038/386066a0 [PubMed] [CrossRef] [Google Scholar]
111. Nitschke J. B., Sarinopoulos I., Mackiewicz K. L., Schaefer H. S., Davidson R. J. (2006). Functional neuroanatomy of aversion and its anticipation. Neuroimage 29, 106–11610.1016/j.neuroimage.2005.06.068 [PubMed] [CrossRef] [Google Scholar]
112. Nittono H. (2006). Voluntary stimulus production enhances deviance processing in the brain. Int. J. Psychophysiol. 59, 15–2110.1016/j.ijpsycho.2005.06.008 [PubMed] [CrossRef] [Google Scholar]
113. Nobre A. C., Correa A., Coull J. T. (2007). The hazards of time. Curr. Opin. Neurobiol. 17, 465–47010.1016/j.conb.2007.07.006 [PubMed] [CrossRef] [Google Scholar]
114. Nobre A. C., Coull J. T., Frith C. D., Mesulam M. M. (1999). Orbitofrontal cortex is activated during breaches of expectation in tasks of visual attention. Nat. Neurosci. 2, 11–1210.1038/4513 [PubMed] [CrossRef] [Google Scholar]
115. Oken B. S., Salinsky M. C., Elsas S. M. (2006). Vigilance, alertness, or sustained attention: physiological basis and measurement. Clin Neurophysiol 117, 1885–190110.1016/j.clinph.2006.01.017 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
116. Opitz B., Friederici A. D. (2007). Neural basis of processing sequential and hierarchical syntactic structures. Hum. Brain Mapp. 28, 585–59210.1002/hbm.20287 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
117. O'Reilly J. X., Mesulam M. M., Nobre A. C. (2008). The cerebellum predicts the timing of perceptual events. J. Neurosci. 28, 2252–226010.1523/JNEUROSCI.2742-07.2008 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
118. Partiot A., Grafman J., Sadato N., Wachs J., Hallett M. (1995). Brain activation during the generation of non-emotional and emotional plans. Neuroreport 6, 1397–140010.1097/00001756-199507100-00009 [PubMed] [CrossRef] [Google Scholar]
119. Patel A. D. (2003). Language, music, syntax and the brain. Nat. Neurosci. 6, 674–68110.1038/nn1082 [PubMed] [CrossRef] [Google Scholar]
120. Petersson K. M., Forkstam C., Ingvar M. (2004). Artificial syntactic violations activate Broca's region. Cogn. Sci. 28, 383–407 [Google Scholar]
121. Petrides M. (2007). The orbitofrontal cortex: novelty, deviation from expectation, and memory. Ann. N. Y. Acad. Sci. 1121, 33–5310.1196/annals.1401.035 [PubMed] [CrossRef] [Google Scholar]
122. Pezzulo G. (2008). Coordinating with the future: the anticipatory nature of representation. Minds Mach. 18, 179–22510.1007/s11023-008-9095-5 [CrossRef] [Google Scholar]
123. Pezzulo G., Butz M. V., Castelfranchi C. (2008). “The anticipatory approach: definitions and taxonomies,” in The Challenge of Anticipation, eds Pezzulo G., Butz M. V., Castelfranchi C., Falcone R., editors. (Berlin: Springer;), 23–43 [Google Scholar]
124. Pezzulo G., Hoffmann J., Falcone R. (2007). Anticipation and anticipatory behavior. Cogn. Process. 8, 67–7010.1007/s10339-007-0173-z [CrossRef] [Google Scholar]
125. Pfordresher P. Q., Palmer C., Jungers M. K. (2007). Speed, accuracy, and serial order in sequence production. Cogn. Sci. 31, 63–9810.1207/s15516709cog3101_3 [PubMed] [CrossRef] [Google Scholar]
126. Ploghaus A., Tracey I., Gati J. S., Clare S., Menon R. S., Matthews P. M., Rawlins J. N. (1999). Dissociating pain from its anticipation in the human brain. Science 284, 1979–198110.1126/science.284.5422.1979 [PubMed] [CrossRef] [Google Scholar]
127. Poirier P., Hardy-Vallee B. (2005). “Structured thoughts: the spatial-motor view,” in The Compositionality of Meaning and Content, Vol. II *Applications to Linguistics, Psychology and Neuroscience*, eds Machery E., Werning M., Gerhard S., editors. (Frankfurt: Ontos Verlag;).

[Google Scholar]
128. Porro C. A., Baraldi P., Pagnoni G., Serafini M., Facchin P., Maieron M., Nichelli P. (2002). Does anticipation of pain affect cortical nociceptive systems? J. Neurosci. 22, 3206–3214 [PMC free article] [PubMed] [Google Scholar]
129. Porro C. A., Cettolo V., Francescato M. P., Baraldi P. (2003). Functional activity mapping of the mesial hemispheric wall during anticipation of pain. Neuroimage 19, 1738–174710.1016/S1053-8119(03)00184-8 [PubMed] [CrossRef] [Google Scholar]
130. Praamstra P., Kourtis D., Kwok H. F., Oostenveld R. (2006). Neurophysiology of implicit timing in serial choice reaction-time performance. J. Neurosci. 26, 5448–545510.1523/JNEUROSCI.0440-06.2006 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
131. Prinz W. (1990). “A common coding approach to perception and action,” in Relationships Between Perception and Action, eds Neumann O., Prinz W., editors. (Berlin: Springer;), 167–201 [Google Scholar]
132. Quintana J., Fuster J. M. (1992). Mnemonic and predictive functions of cortical neurons in a memory task. Neuroreport 3, 721–72410.1097/00001756-199208000-00018 [PubMed] [CrossRef] [Google Scholar]
133. Rainer G., Rao S. C., Miller E. K. (1999). Prospective coding for objects in primate prefrontal cortex. J. Neurosci. 19, 5493–5505 [PMC free article] [PubMed] [Google Scholar]
134. Rao R. P., Ballard D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. 2, 79–8710.1038/4580 [PubMed] [CrossRef] [Google Scholar]
135. Reber A. S. (1967). Implicit learning of artificial grammars. J. Verbal Learn. Verbal Behav. 6, 855–86310.1016/S0022-5371(67)80149-X [CrossRef] [Google Scholar]
136. Rees G., Frith C. D. (1998). How do we select perceptions and actions? Human brain imaging studies. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 353, 1283–129310.1098/rstb.1998.0283 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
137. Remillard G. (2003). Pure perceptual-based sequence learning. J. Exp. Psychol. 29, 581–597 [PubMed] [Google Scholar]
138. Rescorla R. A., Wagner A. R. (1972). “A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement,” in Classical Conditioning, Vol. I, eds Black A. H., Prokasy W. F., editors. (New York: Appleton-Century-Crofts;), 64–99 [Google Scholar]
139. Ritter W., Sussman E., Deacon D., Cowan N., Vaughan H. G., Jr. (1999). Two cognitive systems simultaneously prepared for opposite events. Psychophysiology 36, 835–83810.1017/S0048577299990248 [PubMed] [CrossRef] [Google Scholar]
140. Roitblat H. L., Scopatz R. A. (1983). Sequential effects in pigeon delayed matching-to-sample performance. J. Exp. Psychol. Anim. Behav. Process. 9, 202–22110.1037/0097-7403.9.2.202 [PubMed] [CrossRef] [Google Scholar]
141. Ruby P., Sirigu A., Decety J. (2002). Distinct areas in parietal cortex involved in long-term and short-term action planning: a PET investigation. Cortex 38, 321–33910.1016/S0010-9452(08)70663-4 [PubMed] [CrossRef] [Google Scholar]
142. Schacter D. L., Addis D. R. (2009). On the nature of medial temporal lobe contributions to the constructive simulation of future events. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 364, 1245–125310.1098/rstb.2008.0308 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
143. Schacter D. L., Addis D. R., Buckner R. L. (2007). Remembering the past to imagine the future: the prospective brain. Nat. Rev. Neurosci. 8, 657–66110.1038/nrn2213 [PubMed] [CrossRef] [Google Scholar]
144. Schröger E. (2007). Mismatch negativity: a microphone into auditory memory. J. Psychophysiol. 21, 138–146 [Google Scholar]
145. Schubotz R. I. (2004). Human Premotor Cortex: Beyond Motor Performance. Leipzig: Max Planck

Institute for Human Cognitive and Brain Sciences [Google Scholar]
146. Schubotz R. I. (2007). Prediction of external events with our motor system: towards a new framework. Trends Cogn. Sci. 11, 211–21810.1016/j.tics.2007.02.006 [PubMed] [CrossRef] [Google Scholar]
147. Schubotz R. I., von Cramon D. Y. (2001a). Functional organization of the lateral premotor cortex: fMRI reveals different regions activated by anticipation of object properties, location and speed. Brain Res. Cogn. Brain Res. 11, 97–11210.1016/S0926-6410(00)00069-0 [PubMed] [CrossRef] [Google Scholar]
148. Schubotz R. I., von Cramon D. Y. (2001b). Interval and ordinal properties of sequences are associated with distinct premotor areas. Cereb. Cortex 11, 210–22210.1093/cercor/11.3.210 [PubMed] [CrossRef] [Google Scholar]
149. Schubotz R. I., von Cramon D. Y. (2002a). A blueprint for target motion: fMRI reveals perceived sequential complexity to modulate premotor cortex. Neuroimage 16, 920–93510.1006/nimg.2002.1183 [PubMed] [CrossRef] [Google Scholar]
150. Schubotz R. I., von Cramon D. Y. (2002b). Dynamic patterns make the premotor cortex interested in objects: influence of stimulus and task revealed by fMRI. Brain Res. Cogn. Brain Res. 14, 357–36910.1016/S0926-6410(02)00138-6 [PubMed] [CrossRef] [Google Scholar]
151. Schubotz R. I., von Cramon D. Y. (2002c). Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: an fMRI study. Neuroimage 15, 787–79610.1006/nimg.2001.1043 [PubMed] [CrossRef] [Google Scholar]
152. Schubotz R. I., von Cramon D. Y. (2003). Functional–anatomical concepts of human premotor cortex: evidence from fMRI and PET studies. Neuroimage 20 (Suppl. 1), S120–S13110.1016/j.neuroimage.2003.09.014 [PubMed] [CrossRef] [Google Scholar]
153. Schultz W., Dayan P., Montague P. R. (1997). A neural substrate of prediction and reward. Science 275, 1593–159910.1126/science.275.5306.1593 [PubMed] [CrossRef] [Google Scholar]
154. Schultz W., Dickinson A. (2000). Neuronal coding of prediction errors. Annu. Rev. Neurosci. 23, 473–50010.1146/annurev.neuro.23.1.473 [PubMed] [CrossRef] [Google Scholar]
155. Schütz-Bosbach S., Prinz W. (2007). Prospective coding in event representation. Cogn. Process. 8, 93–10210.1007/s10339-007-0167-x [PubMed] [CrossRef] [Google Scholar]
156. Shanks D. R. (2007). Associationism and cognition: human contingency learning at 25. Q. J. Exp. Psychol. 60, 291–30910.1080/17470210601000581 [PubMed] [CrossRef] [Google Scholar]
157. Simmons A., Matthews S. C., Stein M. B., Paulus M. P. (2004). Anticipation of emotionally aversive visual stimuli activates right insula. Neuroreport 15, 2261–226510.1097/00001756-200410050-00024 [PubMed] [CrossRef] [Google Scholar]
158. Soga R., Akaishia R., Sakai K. (2009). Predictive and postdictive mechanisms jointly contribute to visual awareness. Conscious. Cogn. 18, 578–59210.1016/j.concog.2009.04.007 [PubMed] [CrossRef] [Google Scholar]
159. Sperry R. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion, J. Comp. Psychol. Physiol. 43, 482–48910.1037/h0055479 [PubMed] [CrossRef] [Google Scholar]
160. Spratling M. W. (2008a). Predictive coding as a model of biased competition in visual attention. Vision Res. 48, 1391–140810.1016/j.visres.2008.03.009 [PubMed] [CrossRef] [Google Scholar]
161. Spratling M. W. (2008b). Reconciling predictive coding and biased competition models of cortical function. Front. Comput. Neurosci. 2:4.10.3389/neuro.10.004.2008 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
162. Squires K. C., Wickens C., Squires N. K., Donchin E. (1976). The effect of stimulus sequence on the waveform of the cortical event-related potential. Science 193, 1142–114610.1126/science.959831 [PubMed] [CrossRef] [Google Scholar]
163. Stock A., Stock C. (2004). A short history of ideo-motor action. Psychol. Res. 68, 176–18810.1007/s00426-003-0154-5 [PubMed] [CrossRef] [Google Scholar]

164. Summerfield C., Egner T. (2009). Expectation (and attention) in visual cognition. Trends Cogn. Sci. 13, 403–40910.1016/j.tics.2009.06.003 [PubMed] [CrossRef] [Google Scholar]
165. Summerfield C., Egner T., Greene M., Koechlin E., Mangels J., Hirsch J. (2006). Predictive codes for forthcoming perception in the frontal cortex. Science 314, 1311–131410.1126/science.1132028 [PubMed] [CrossRef] [Google Scholar]
166. Summerfield C., Koechlin E. (2008). A neural representation of prior information during perceptual inference. Neuron 59, 336–34710.1016/j.neuron.2008.05.021 [PubMed] [CrossRef] [Google Scholar]
167. Tanji J., Shima K. (1994). Role for supplementary motor area cells in planning several movements ahead. Nature 371, 413–41610.1038/371413a0 [PubMed] [CrossRef] [Google Scholar]
168. Tolman E. C. (1948). Cognitive maps in rats and men. Psychol. Rev. 55, 189–20810.1037/h0061626 [PubMed] [CrossRef] [Google Scholar]
169. Ueda K., Okamoto Y., Okada G., Yamashita H., Hori T., Yamawaki S. (2003). Brain activity during expectancy of emotional stimuli: an fMRI study. Neuroreport 14, 51–5510.1097/00001756-200301200-00010 [PubMed] [CrossRef] [Google Scholar]
170. Ullman S. (1995). Sequence seeking and counter streams: a computational model for bidirectional information flow in the visual cortex. Cereb. Cortex 5, 1–1110.1093/cercor/5.1.1 [PubMed] [CrossRef] [Google Scholar]
171. von Holst E., Mittelstaedt H. (1950). Das Reafferenzprinzip. Naturwissenschaften 37, 464–47610.1007/BF00622503 [CrossRef] [Google Scholar]
172. Waszak F., Herwig A. (2007). Effect anticipation modulates deviance processing in the brain. Brain Res. 1183, 74–8210.1016/j.brainres.2007.08.082 [PubMed] [CrossRef] [Google Scholar]
173. Webb B. (2004). Neural mechanisms for prediction: do insects have forward models? Trends Neurosci. 27, 278–28210.1016/j.tins.2004.03.004 [PubMed] [CrossRef] [Google Scholar]
174. Whitney D., Murakami I. (1998). Latency difference, not spatial extrapolation. Nat. Neurosci. 1, 656–65710.1038/3659 [PubMed] [CrossRef] [Google Scholar]
175. Widmann A., Gruber T., Kujala T., Tervaniemi M., Schroger E. (2007). Binding symbols and sounds: evidence from event-related oscillatory gamma-band activity. Cereb. Cortex 17, 2696–270210.1093/cercor/bhl178 [PubMed] [CrossRef] [Google Scholar]
176. Willingham D. B., Nissen M. J., Bullemer P. (1989). On the development of procedural knowledge. J. Exp. Psychol. 15, 1047–1060 [PubMed] [Google Scholar]
177. Wills A. J., Lavric A., Croft G. S., Hodgson T. L. (2007). Predictive learning, prediction errors, and attention: evidence from event-related potentials and eye tracking. J. Cogn. Neurosci. 19, 843–85410.1162/jocn.2007.19.5.843 [PubMed] [CrossRef] [Google Scholar]
178. Winkler A. S. (2007). Interpreting the mismatch negativity. J. Psychophysiol. 21, 147–16310.1027/0269-8803.21.34.147 [CrossRef] [Google Scholar]
179. Winkler I., Czigler I. (1998). Mismatch negativity: deviance detection or the maintenance of the ‘standard’. Neuroreport 9, 3809–3813 [PubMed] [Google Scholar]
180. Winkler I., Karmos G., Näätanen R. (1996). Adaptive modeling of the unattended acoustic environment reflected in the mismatch negativity event-related potential. Brain Res. 742, 239–25210.1016/S0006-8993(96)01008-6 [PubMed] [CrossRef] [Google Scholar]
181. Winograd E. (1988). “Some observations on prospective remembering,” in Practical Aspects of Memory: Current Research and Issues, eds Gruneberg M. M., Morris P. E., Sykes R. N., editors. (Chichester: Wiley;), 348–353 [Google Scholar]
182. Wolpert D. M., Doya K., Kawato M. (2003). A unifying computational framework for motor control and social interaction. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 358, 593–60210.1098/rstb.2002.1238 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
183. Wolpert D. M., Flanagan J. R. (2001). Motor prediction. Curr. Biol. 11, R729–R73210.1016/S0960-9822(01)00432-8 [PubMed] [CrossRef] [Google Scholar]
184. Wolpert D. M., Ghahramani Z., Jordan M. I. (1995). An internal model for sensorimotor integration.

Science 269, 1880–188210.1126/science.7569931 [PubMed] [CrossRef] [Google Scholar]
185. Wolpert D. M., Kawato M. (1998). Multiple paired forward and inverse models for motor control. Neural. Netw. 11, 1317–132910.1016/S0893-6080(98)00066-5 [PubMed] [CrossRef] [Google Scholar]
186. Wolpert D. M., Miall R. C. (1996). Forward models for physiological motor control. Neural. Netw. 9, 1265–127910.1016/S0893-6080(96)00035-4 [PubMed] [CrossRef] [Google Scholar]
187. Wylie G. R., Javitt D. C., Foxe J. J. (2006). Jumping the gun: is effective preparation contingent upon anticipatory activation in task-relevant neural circuitry? Cereb. Cortex 16, 394–40410.1093/cercor/bhi118 [PubMed] [CrossRef] [Google Scholar]

Articles from Frontiers in Human Neuroscience are provided here courtesy of **Frontiers Media SA**