

The proactive brain: memory for predictions

Moshe Bar*

*Martinos Center for Biomedical Imaging at MGH, Harvard Medical School, Charlestown,
MA 02129, USA*

It is proposed that the human brain is proactive in that it continuously generates predictions that anticipate the relevant future. In this proposal, analogies are derived from elementary information that is extracted rapidly from the input, to link that input with the representations that exist in memory. Finding an analogical link results in the generation of focused predictions via associative activation of representations that are relevant to this analogy, in the given context. Predictions in complex circumstances, such as social interactions, combine multiple analogies. Such predictions need not be created afresh in new situations, but rather rely on existing *scripts* in memory, which are the result of real as well as of previously imagined experiences. This cognitive neuroscience framework provides a new hypothesis with which to consider the purpose of memory, and can help explain a variety of phenomena, ranging from recognition to first impressions, and from the brain's 'default mode' to a host of mental disorders.

Keywords: expectation; anticipation; foresight; mindset; memory; default network

1. INTRODUCTION: THE GENERAL FRAMEWORK

A common belief that might be implicit but not entirely accurate is that there is a clear boundary between perception and cognition. According to this artificial boundary 'perception' pertains to analysis of the information about the physical world around us, conveyed by the senses and analysed by the sensory cortex, while 'cognition' refers to the analysis and operations that are taking place beyond what is required for perceiving the input; faculties such as attention, memory and categorization. For example (the examples here will tend to focus on the world of visual object recognition, because it is a field in which intense research activity has been invested for several decades now, and because recognition is unique in how it straddles the definitions of perception and cognition by relying on both to be accomplished), the assumption is that when we encounter an object in our environment, our primary goal is to understand what it is (i.e. recognize it). However, I propose that rather than seeing the challenge of vision (or any of the other senses) as answering the question 'what is this?', we should look at the goal more as linking the input with an analogous representation in memory, and simultaneously with the information associated with it, by asking instead 'what is this *like*?' Such recognition-by-analogy, and the concomitant associative activation, does not require an exhaustive analysis of the input's properties (e.g. colour, contour, texture) nor rely exclusively on a bottom-up, sensory-triggered, flow of information. This modified perspective puts emphasis on how we use experience by affording immediate access to analogies and associated representations. In other words, our perception of the environment relies on memory as much as it does on incoming

information, which blurs the border between perception and cognition. Support for this notion that sensory information interacts with existing representations for the purpose of perception proper is related to the concept of embodied cognition (Barsalou 2003; Noe 2005), and has been recently supported by functional magnetic resonance imaging studies (Wang & Jiang 2008).

At the heart of the proposed framework lies the idea that we do not interpret our world merely by analysing incoming information, but rather we try to understand it using a proactive link of incoming features to existing, familiar information (e.g. objects, people, situations). More specifically, when encountering a novel input (and all inputs are novel to some degree because we never encounter anything twice under exactly the same conditions), our brains 'ask' what does this input resemble from things with which we are already familiar? Critically, this question is being asked actively using top-down guesses that could be based on rudimentary input information (Bar *et al.* 2006). Once an analogy is found (e.g. 'this is the driver's seat of a car'), associated representations are activated rapidly, a process that provides the platform for predictions that pre-sensitize the representations of what is most likely to occur and be encountered next (e.g. 'this must be the stick-shift, this must be the headlight switch, and I am sure there is a cup holder here somewhere').

This process of analogy → associations → predictions can be seen as providing the basis for a universal principle. Naturally, our environment and needs involve circumstances that are considerably more complex than the basic elemental process described here, but this process can be expanded to any level of complexity. Specifically, multiple analogies can be found in parallel, relating to multiple aspects of the input, and they can be combined to generate compound predictions that go beyond simple perception and cognition and are as complex as those

*bar@nmr.mgh.harvard.edu

One contribution of 18 to a Theme Issue 'Predictions in the brain: using our past to prepare for the future'.

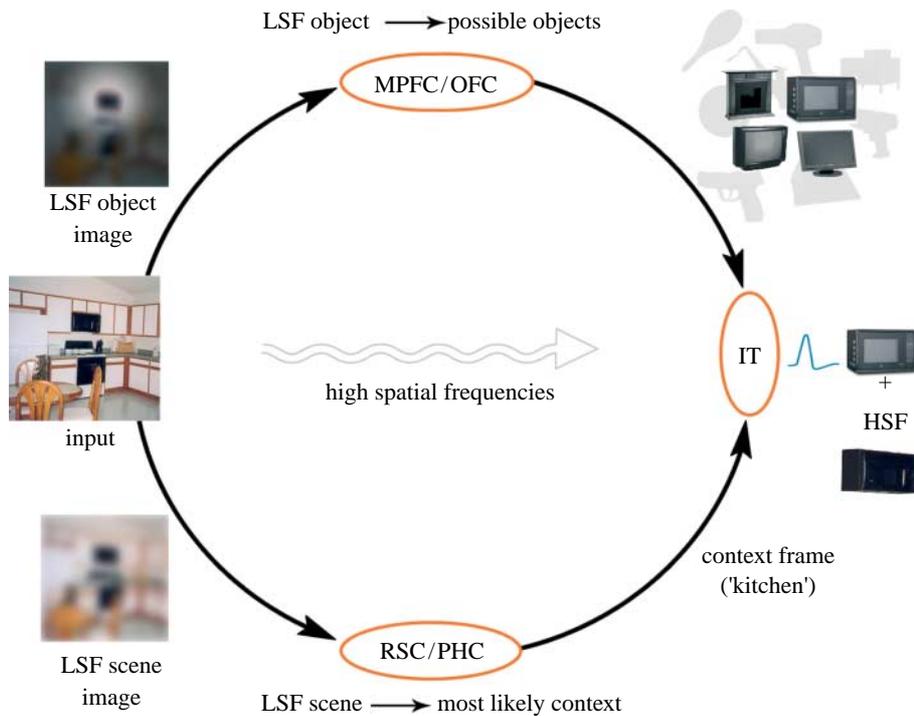


Figure 1. Combining object and context information to find a quick and reliable analogy. In parallel to the bottom-up systematic progression of the image details along the visual pathways, there are quick projections of low spatial frequency (LSF) information, possibly via the magnocellular pathway. This coarse but rapid information is sufficient for generating an 'initial guess' about the context and about objects in it. These context-based predictions are validated and refined with the gradual arrival of higher spatial frequencies (HSFs) (Bar 2004). (MPFC, medial prefrontal cortex; OFC, orbital frontal cortex; RSC, retrosplenial complex; PHC, parahippocampal cortex; IT, inferior temporal cortex.) The arrows are unidirectional in the figure to emphasize the flow during the proposed analysis, but all these connections are bidirectional in nature. Adapted from Bar (2004).

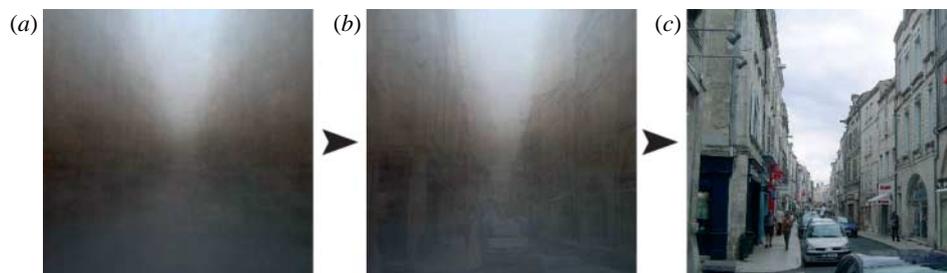


Figure 2. When seeing a novel image of a scene, its LSF version rapidly activates a prototypical context frame (a) in memory (i.e. an analogy). The relatively slower and gradual arrival of detail in the HSFs (b) results in a more episodic context frame (c), with scene-specific information. Image in (a) courtesy of A. Torralba.)

required, for example, in social interactions. In other words, predictions do not merely rely on reactivation of previously experienced environments and situations, but rather can be derived from a combination of representations in memory to generate a novel mental scenario.

Analogies, associations and predictions are all issues that have been studied extensively in the past, individually. The present proposal brings them together by synthesizing previous findings and building upon them to provide an integrative framework for how the brain generates proactive predictions that facilitate our interactions with the environment, where associations serve as the building blocks and analogies provide the trigger. The next sections elaborate on these individual components and on their integration.

2. ANALOGIES: THE TRIGGER

Analogy is typically seen as a sophisticated cognitive tool used in specific types of problem-solving (Gentner

1983; Muter & Snowling 1994; Holyoak & Thagard 1997). Here the term is used instead simply to refer to the link of a novel input to a similar representation existing in memory: What is this like? (A metaphor could have been another choice.) Analogical mapping facilitates interpretation, but more importantly connects with a set of associated representations that provide a platform for predictions. Analogies can be based on similarity in various levels; perceptual, conceptual, semantic, functional and so on. Indeed, analogies as considered here can facilitate anything from the recognition of a cat never seen before to helping us decide what to pack when going on holiday in a new destination.

To demonstrate how analogy can help interpret the input rapidly using only rudimentary information, consider a proposed model of how a visual object can be recognized rapidly using only global information about its general appearance and about the context in

which it is embedded (figure 1). The top part is derived from an earlier proposal (Bar 2003) and supporting evidence (Bar *et al.* 2006; Kveraga & Boshyan 2007), whereby in addition to the systematic analysis of detail along the ventral visual pathway, global appearance properties, conveyed by low spatial frequencies (LSFs), are projected directly from early visual areas to the orbital frontal cortex (OFC). The OFC then triggers the activation of the most probable object identities that share the same global appearance as the input target object with which an analogy in memory is sought. This proposal and corresponding supportive evidence suggest that OFC represents visual information, although the level and specificity of object representations in OFC is still under investigation.

In parallel (bottom part of figure 1), an LSF image of the input scene is typically sufficient for extracting the context in which the target object appears (Oliva & Torralba 2001; Torralba & Oliva 2003; Bar 2004), thereby activating the representations of objects and relationships that are common to that specific context (*context frames*; Bar & Ullman 1996; Bar 2004), using a cortical network involving the parahippocampal cortex (PHC), the retrosplenial complex (RSC) and to some extent also the medial prefrontal cortex (MPFC; Bar & Aminoff 2003; Bar 2004). Together, a simple intersection of the candidate object interpretations (figure 1, top) with the objects that typically appear in such context (figure 1, bottom) yields a quick recognition of the target object on a basic level. This basic-level recognition provides the analogy (i.e. input–memory link) that triggers the associative generation of predictions.

Such analogy matching requires generic representations in memory, which, in addition to being typical of the exemplar that they represent, should be invariant to variations that are not meaningful to the analogy. (In the research of analogy, as defined traditionally, there have been numerous arguments and demonstrations that individual exemplars can actually be more appropriate for category representations than prototypes (e.g. Allen & Brooks 1991), but this is beyond the scope of the current discussion and the different manner by which the term analogy is considered here.) How are these generic representations accomplished? This has proved to be an extremely difficult problem to answer. One proposal is based on the observation that various instances of the same exemplar (e.g. object or context) can vary dramatically in their details, but they nevertheless share global properties. In the realm of visual cognition, we can consider LSF scene images as prototypical representations of contexts, because they contain only such instance-invariant visual features, and do not represent details that vary from one instance to another. An LSF image of an input scene typically will match and activate one such global/prototypical context frame in memory (i.e. the analogy). This is sufficient in most cases to generate rapid predictions that guide our pressing goals, such as navigation and avoidance. The arrival of the specific detail, with the high spatial frequencies (HSFs), gradually makes the initially activated prototypical representation more

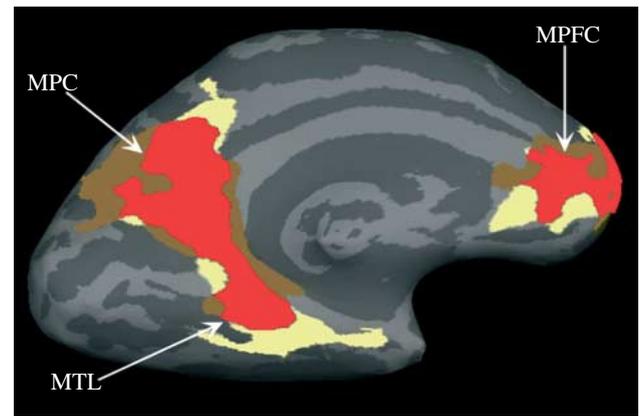


Figure 3. Medial view of the left hemisphere, demonstrating the overlap between the default network (defined, for example, using the contrast between activation at ‘rest’ baseline versus activation in an n -back working memory task) and the associative predictions network (defined, for example, using the contrast between recognition of highly associative objects versus recognition of only weakly associative objects). Brown area, contextual associations network; yellow area, default network; red area, overlap. MTL, medial temporal lobe; MPC, medial parietal cortex; MPFC, medial prefrontal cortex.

episodic, in that it ‘fills’ the LSF blobs with image-specific details (figure 2).

One possibility for the development of prototypical representations in memory (e.g. the context frame) with experience is by averaging accumulated instances of the same context, which could be seen as analogous to LSF filtering. Indeed, the street prototype (figure 2a) is a result of averaging over 100 street images (Torralba & Oliva 2003). In addition to providing the trigger for the generation of predictions, analogies that are based on such ‘averaged’ representations can be used as a powerful tool for accommodating the infinite variations in the appearance of the physical world around us, where things tend to be familiar but not identical, as has been demonstrated impressively in computer vision (Jenkins & Burton 2008). In other words, a global, coarse, representation of a certain concept can be activated by different exemplars of that concept, regardless of changing details, and thus affords both analogy and a somewhat invariant recognition.

Everyday situations tend to carry additional regularities and diagnostic properties beyond their ‘average’ physical appearance. Consider, for example, your actions when you pick up the phone to discover (i.e. by analogy with your previous experience) that it is a telemarketing call. The appearance of the phone and the room you are in, as well as the voice and even the content of the conversation, do not influence your reaction; you largely execute a familiar response, with only minor variations, based on knowledge that is considerably higher level than simple perceptual properties. In summary, when entities in our environment occur frequently enough, with repeating diagnostic properties, these regularities can be used to link an input, via analogy, to similar representations in memory, which then allows using associated information as predictions that facilitate our interaction with the environment.

3. ASSOCIATIONS: THE BUILDING BLOCKS

The importance of associations to our mental lives has been widely recognized and has been discussed already by early philosophers, dating back to Plato, Aristotle and Vives. Associations provide the vehicle for memory encoding and retrieval, but here they are also proposed to serve as the building blocks of predictions: by activating a certain analogy, information that is associated with this analogy in memory is triggered, generating an 'expectation' by becoming pre-sensitized.

Seeing the brain as proactive implies that, by 'default', when we are not engaged in some demanding and all-consuming task, the brain generates predictions. Therefore, if generating predictions is a continuous operation of the brain, and predictions rely on associative activation, one needs to show that associative activation is an ongoing process in human thought. We have recently provided such demonstration (Bar *et al.* 2007) by showing a striking overlap between the cortical network that mediates contextual associative processing and the cortical network that has been termed the brain's 'default network' (Raichle *et al.* 2001).

The default network is believed to subserve the mental processes that take place in the brain when subjects are not engaged in a specific goal-oriented task (Binder *et al.* 1999; Mazoyer *et al.* 2001; Raichle *et al.* 2001). The major aspects of this default network overlap remarkably with the same regions that we have found as directly related to contextual associative activation (figure 3; Bar & Aminoff 2003; Aminoff *et al.* 2007; Bar *et al.* 2007). This overlap between the default network and the network subserving associative processing of contextually related information is taken as the cortical manifestation that associative predictions are crucial, ongoing, elements of natural thought. This account further allows a more specific ascription of a cognitive function to the brain's default activity (Bar *et al.* 2007).

Interestingly, the regions of the contextual associations network and of the default network that exhibit the greatest overlap—MPC (medial parietal cortex; termed also RSC above), structures in the medial temporal lobe (MTL) and the MPFC—have been reported to be activated in an exceptionally wide variety of studies: navigation and spatial processing (O'Craven & Kanwisher 2000; Maguire 2001); execution errors and planning of saccadic eye movements (Polli & Barton 2005); episodic memory (Ranganath *et al.* 2004; Wagner & Shannon 2005); decision-making (Fleck *et al.* 2006); emotional processing (Maddock 1999); self-referential processing (Kelley *et al.* 2002; Macrae & Moran 2004); social interactions (Iacoboni *et al.* 2004); and mental state attribution (i.e. theory of mind; Saxe & Kanwisher 2003; Frith & Frith 2006). How could the same cortical regions apparently mediate so many different functions? Our proposal has been that the basic operation which is shared by all these diverse processes is the reliance on associations and the generation of predictions (Bar *et al.* 2007). In fact, it is hard to imagine another role that could be assigned to this network that can bridge such a wide variety of mental processes.

Our first approximation of the division of labour between the primary three components of this central

cortical network is that the MTL holds an episodic, physically specific representation of an immediate context, MPC/RSC contains prototypical representations for typical contexts (e.g. kitchen, theatre, beach) including general but diagnostic information on each, and the MPFC uses this associative information to generate predictions (Bar 2007; Bar *et al.* 2007; Aminoff *et al.* 2008). Indeed, the MPFC, the MPC and the MTL have all been reported to show selective activation in recent studies of thinking about the past and the future (Burgess *et al.* 2001; Okuda & Fujii 2003; Addis *et al.* 2007; Buckner & Carroll 2007; Szpunar *et al.* 2007), consistent with the roles we have assigned here to this network.

It is important to note that the term associations is fairly non-specific in that it is certain to involve more than one type of mechanism and representations, and therefore more regions are expected to be involved in the processing of additional types of associations. For example, the striatum, the caudate nucleus and the cerebellum have all been shown to be involved in various paradigms related to associative processing (Pasupathy & Miller 2005).

In the framework proposed here, associations can vary in the automaticity of their activation. On one extreme, associations are automatic, simple and unique, similar to a basic Hebbian association where associated concepts are co-activated when the representation of one of them is activated. Such associations have been the primary focus of research, and have largely been shown to be mediated by the MTL (Petrides 1985; Schacter 1987; Miyashita 1993; Shallice *et al.* 1994; Eichenbaum & Bunsey 1995; Suzuki & Eichenbaum 2000; Stark & Squire 2001; Sperling & Chua 2003; Ranganath *et al.* 2004; Aminoff *et al.* 2007). Other associations are more deliberative, and their selective co-activation depends on the specific context in which they are embedded (Bar 2004). Yet more complex types of associations might combine the output of different modules performing mental simulations (Barsalou 2009; Hassabis & Maguire 2009; Moulton & Kosslyn 2009) and other relatively higher level operations. It is proposed here that even complex mental experiences are derived from associations with simpler, although not necessarily automatically activated, elements. These associations are formed through experience based on similarity and frequent co-occurrence in space and time.

The concept of associative activation is powerful beyond the generation of predictions about what other objects, people and events to expect in a given context. With the broad activations afforded by associative structuring of memory, associative activations can be seen as guiding our behaviour more globally by determining a 'mindset'. This idea is still in its infancy, but it is interesting to consider here, at least one extreme example of how such associative activations affect somewhat unexpected aspects of behaviour. Specifically, Bargh and his colleagues (Bargh & Chen 1996; Dijksterhuis *et al.* 2000) have shown that associative activations, which their stereotype priming techniques can be seen as, can have surprisingly powerful effects on behaviour. For example, priming rudeness in participants made them later interrupt the experimenter more frequently, and priming subjects with elderly traits

(simply by exposing them to a trait-related collection of words) made them subsequently walk slower in the hallway when leaving the laboratory. While these links might seem far-fetched at first, they suggest that priming even a very high-level concept can activate associative predictions and action patterns that together constitute a mindset that is congruent with the prime.

The proposed concept of mindset can be seen as composed of a broad set of predictions, a repertoire of what is expected in the given context and what is not, which constitutes a *state* for guiding behaviour and for tuning our perceptions and cognitions based on the prime that has elicited that mindset (e.g. acting 'old'). Naturally, a mindset can be modified based on ongoing circumstances. One prediction that stems from such a mindset concept is that our response to a certain stimulus is not absolute, but depends on the state determined by the mindset. This has been demonstrated repeatedly in the realm of emotion and affect, but is valid in other settings as well. For example, our response to a sight of a pizza will be profoundly different if we are in a hungry mindset compared with when we have just finished lunch and are rushing to a meeting. A second prediction is that a great deal of the ubiquitous activations seen in the default network can be explained as mindset, and such default activity differs qualitatively when subjects are in different mindsets (e.g. in a 'dance club' mindset versus in a 'job interview' mindset). We tend to think that cortical representations are triggered either by perception or internally retrieved with recall, imagery and simulations. But mindsets would imply that we have a sustained (though updateable) list of needs, goals, desires, predictions, context-sensitive conventions and attitudes. This sustained list constitutes our mindset, and could be continuously maintained and imposed by activity in the contextual/default network. Mindsets can be suspended temporarily for performing an immediate task or achieving a short-term goal. It is possible that mindsets can also be suspended for longer durations, or even emptied, by methods such as meditation. In summary, the concept of mindsets can provide a global framework with which to consider context-based behaviour.

Finally, the central role of associations in our mental lives has far-reaching implications to clinical mental disorders. Particularly interesting to note is that the pattern of activity typically observed at 'rest' (i.e. 'default mode') in healthy individuals differs in patients with major depression. For example, the MPFC exhibits hypoactivity during periods of rest in depressed individuals (Drevets *et al.* 1992; Soares & Mann 1997). Furthermore, the structure, function and connectivity of the same default/associative network are compromised in depression (Mayberg & Liotti 1999; Drevets 2000; Anand *et al.* 2005), and its integrity is improved with antidepressant-related clinical improvement (Mayberg & Liotti 1999; Anand *et al.* 2005). These findings provide critical support to a novel hypothesis that mood can be modulated by associative processing (Bar submitted). That foresight is severely impaired in depressed individuals (Williams *et al.* 1996) provides encouraging support for this hypothesis.

4. PREDICTIONS AND SOME IMPLICATIONS

The proposal that memories are encoded in an associative manner (e.g. context frames) and that their holistic activation generates predictions gives rise to several interesting hypotheses. First, one might wonder why our brain is investing energy in mind wandering, fantasizing and revisiting (and modifying) existing memories. After all, if these are not geared directly towards a specific goal (such as in concrete planning) why not just rest? I propose that a central role of what seems random thoughts and aimless mental simulations is to create 'memories' (Bar 2007). Information encoded in our memory guides and sometimes dictates our future behaviour. One can look at our experience as stored in memory as scripts. The notion of such scripts is similar to proposals in language and artificial intelligence (Schank 1975), and similar to the idea of stored motor plans (Schubotz & von Cramon 2002; Mushiake *et al.* 2006), which contain information on what was the proper response and expectation under similar conditions in the past (see also Barsalou (2009) and Barbey *et al.* (2009) for a similar concept). The idea of behaviour scripts existing in the cortex is supported to some extent by findings of behaviour 'segments' in Broca's area (Koechlin & Jubault 2006). The generation of predictions based on associations with past experiences and memories is further related to Ingvar's 'memory of the future' (Ingvar 1985). But why should these scripts only be derived from real experiences, given that our mind is powerful enough to generate simulated experiences that did not happen but could happen in the future? Unlike real memories, these simulation-based memories have not really taken place, but we benefit from them just as we do from memories that did occur previously. Therefore, one primary role of memory is to guide our behaviour in the future based on analogies, and this memory can be a result of real as well as imagined experience.

An example might help make this point more vivid. You are waiting for your turn for a haircut, and with nothing good to read but an old shampoo catalogue, you let your mind wander. You imagine a scenario of an earthquake: 'what if a powerful earthquake erupts?' Your thoughts can become very specific about your actions in this hypothetical case: how you will locate your family; how you will be ready to help other patrons in that hairdresser's place; and so on. Now it is a 'memory'. A future memory of an event that has never happened, but has some chance, even if slim, of happening, can help facilitate your actions if it ever happens in reality, just as a real memory that is based on actual experience. Interestingly, not only are such 'memories' helpful in guiding our actions in various situations, they can suffer the same types of memory distortion to which real memories are prone. For example, when you could swear you had actually written an email that in reality you had only planned to, in detail, while stuck in traffic. In summary, this perspective promotes considering imagined scenarios, and what might appear as mind wandering, as beneficial to our learning and future behaviour as much as real experiences.

A second interesting issue concerns the question that, while memory is clearly used for generating predictions, what is the influence of predictions on

memory? A proposal that resonates with the framework described here is that aspects of new experiences which meet our expectations are less likely to be encoded. We primarily encode what differed from our memory and predictions: surprises as well as details if they are important/diagnostic enough. Some have argued that such deviations from predictions provide the basis for learning (Schultz & Dickinson 2000). The issue of what is left in memory from a given experience is interesting and is not fully understood, but it is worth noting how many of the known criteria for memory encoding, such as saliency, emotional value, surprise and novelty, relate to predictions.

A third topic pertains to the intriguing interplay between the generation of predictions and the allocation of attention. One of the foremost benefits afforded by accurate predictions is that they rid us of the need to exert mental effort and allocate attention towards predictable aspects of the environment. Of course, when the predictions relate to upcoming items and events that are relevant for accomplishing a specific task at hand (e.g. looking for a friend in a restaurant), predictions directly guide our attention (e.g. spatial locations as well as identity features). Indeed, performance errors can be predicted by failures of preparatory attentional allocation (e.g. Weissman *et al.* 2006) and studies demonstrated the role of predictive associations and memory in the allocation of attention (e.g. Moores *et al.* 2003; Summerfield & Lepsien 2006).

The power of predictions is that we can anticipate some context-specific aspects, to which we do not have to allocate as much attention, and therefore remain with the resources to explore our environment for novelties from which we can learn, and surprises we should avoid. Generating predictability (and thus reducing uncertainties) based on our experience is therefore a powerful tool for detecting the unexpected. One way to look at this issue is to consider predictions as top-down influences, that bias the pre-sensitization of certain representations based on what is most likely to be relevant in that specific situation. In most typical environments, these predictions are met with corresponding incoming (bottom-up) information, which helps us 'not worry' about those aspects of the environment. But in some cases, the bottom-up information ascending from the senses does not meet our expectations (e.g. a shoe hanging from a tree branch, a technology gadget we have never seen before or an unexpected sound of an explosion). Those unpredictable incoming aspects that do not meet the possibilities offered by the top-down predictions (i.e. mindset) can provide a signal both for attentional allocation as well as for subsequent memory encoding. An issue that will not be covered here but is nevertheless related and interesting to consider in this context pertains to the act of balancing our need to learn by exploring new information and our need for certainty afforded by seeking the familiar (Cohen & Aston-Jones 2005; Daw *et al.* 2006).

Inhibition is another issue that bears direct relationship to predictions. Similarly to the implication of the prefrontal cortex (PFC) in predictions, the PFC has been implicated in many other high-level functions,

including executive control, attentional guidance, response to uncertainty and reward value, understanding consequences, self-reference and inhibition. With the exception of inhibition, all these functions can be directly seen as involving associations-based predictions. It is interesting then to consider what might be the role that inhibition possibly plays in the generation of predictions. One might see the generation of predictions that activate specific representations as confining the alternatives by inhibiting other representations that are less likely to be relevant in the given context. Of course, this does not mean that all representations we have in memory are actively inhibited with each predictive signal. Instead, certain aspects of the environment might give rise to associations that might result in somewhat irrelevant or even misleading predictions. For example, an image of a towel might be associated in memory (e.g. MTL or visual cortex) with many other objects; some are relevant in a bath context, some in a beach context, and so on. Activating the representation of a towel might cause all these associates to be activated as well. But only some of them are relevant in the given context; activating automatically the representation of a soap bar when seeing a towel on the beach, rather than in a bath context, will be wasteful and misleading. Note that the inhibition proper does not need to take place in PFC. Once predictions are conveyed to the relevant lower cortex, these predictions can start local processes that inhibit contextually incongruent associative activations as necessary. Taken together, inhibition can play a powerful role in helping the selective activation of only the most relevant representations as predictions, therefore maximizing their usefulness. A lack of inhibition might cause overly broad associative activations and therefore unhelpful, non-specific, predictions. Abnormally increased level of inhibition, on the other hand, might prevent associative activations, which could result in mood disorders (Bar *submitted*).

One everyday manifestation of the proposed framework of using rudimentary information to find an analogy that then serves the generation of predictions is this of first impressions and stereotypes. Consider first impression about someone's personality traits. It has been shown that such judgements can be formed with extremely brief exposures (Bar & Neta 2006; Willis & Todorov 2006), and that they rely on the rapidly arriving global details conveyed by LSF (Bar & Neta 2006). By extracting those rudimentary properties and forming an opinion, our brains actually find an analogy ('this guy looks *similar to Zach*') that then serves as a platform for predictions ('so he must be also frugal and a music expert'), which will directly influence our interactions with that newly introduced person. Similarly, whatever we might be judging based on body language, possibly even without the awareness neither of the perceiver nor of the 'transmitter', could be seen as a symbol (e.g. standing very close) that connects with analogy and predictions ('she must really like me. Or she has hearing problems ...'). Such predictive first impressions might be useful and possibly somewhat accurate when the judged traits directly pertain to our survival (e.g. potential threat) and less so for other traits (e.g. intelligence; Bar & Neta 2006). A related

example is people's response when listening to an exceptionally gifted speaker with oratorical talents that make the listeners automatically assign credibility and authority to that speaker, with less emphasis on the content compared with their reaction to the message carried by more ordinary speakers. Such examples might all be seen as types of prejudiced predictions that are based on global, rudimentary information.

In summary, the generality of information within an activated associative context frame permits it to be applied to new instances of the relevant context such that experiences in memory can help guide new experiences (Bartlett 1932; Brewer & Nakamura 1984; Bar 2007). In most cases such vagueness about details is beneficial for our ability to generalize, although in the cases of prejudice and stereotypes, details and a more bottom-up perception are necessary to avoid undesirable judgements.

5. CONCLUSIONS

Predictions, and their constructions from associations, span a wide range of complexity and function: from basic Hebbian-like associative activation to complex scenarios involving simulations and integration of multiple elements. Their application is correspondingly highly versatile, from simple and procedural automatic learning to language and social interactions.

There is clearly a need for developing models of how this framework of analogies → associations → predictions is realized, a task that is undoubtedly complex. The framework proposed here includes various facets. First, the brain is proactive in generating predictions. Second, interpretation via analogies is meant to answer the question 'what is this *like*?' Third, associations play a central role in foresight. Fourth, the information stored in our memory exerts its contribution to behaviour by way of predictions. This information in memory may be represented as scripts for guiding behaviour, some of which are acquired from actual experience and some are a result of mental simulations. Fifth, inhibition shapes and fine-tunes the selectivity and thus the usefulness of the predictions activated in a given context. Finally, the web of activations that is elicited in a certain situation provides a set of predictions that determines a mindset, which can dictate our responses globally in accordance with the 'prime' that has elicited that mindset.

At a given moment, we integrate information from multiple points in time. We are rarely in the 'now', but rather combine past and present in anticipation of the future. How the brain integrates representations from different points in time, while still distinguishing among them, is an important question for future research. It is also important to understand the largely understudied issue of the computational operations, and the underlying cortical mechanisms, mediating the transformation of a past memory into a future thought. Furthermore, it is possible to imagine how principle properties of an image could be represented in memory (e.g. by way of LSFs that exclude details but preserve diagnostic, global properties), but it is unclear how our experience with more complex situations (e.g. travelling) can be encoded at a gist level in a way that can

easily be applied in novel, analogous circumstances. Revealing the cognitive and neural mechanisms underlying these important issues will fundamentally promote our understanding of how our brains rely on predictions as a universal principle for its operation.

The author would like to thank L. Barrett, L. Barsalou, J. Boshyan, K. Kveraga, B. R. Rosen, K. Shepherd and C. Thomas for their constructive comments and help with the manuscript. This work was supported by National Institute of Neurological Disorders and Stroke grant NS050615 and NIH National Center for Research Resources grant 5P41RR014075.

REFERENCES

- Addis, D. R., Wong, A. T. & Schacter, D. L. 2007 Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. *Neuropsychologia* **45**, 1363–1377. (doi:10.1016/j.neuropsychologia.2006.10.016)
- Allen, S. W. & Brooks, L. R. 1991 Specializing the operation of an explicit rule. *J. Exp. Psychol. Gen.* **120**, 3–19. (doi:10.1037/0096-3445.120.1.3)
- Aminoff, E., Gronau, N. & Bar, M. 2007 The parahippocampal cortex mediates spatial and nonspatial associations. *Cereb. Cortex* **17**, 1493–1503. (doi:10.1093/cercor/bhl078)
- Aminoff, E., Schacter, D. L. & Bar, M. 2008 The cortical underpinnings of context-based memory distortion. *J. Cogn. Neurosci.* **20**, 2226–2237. (doi:10.1162/jocn.2008.20156)
- Anand, A., Li, Y., Wang, Y., Wu, J., Gao, S., Bukhari, L., Mathews, V. P., Kalnin, A. & Lowe, M. J. 2005 Activity and connectivity of brain mood regulating circuit in depression: a functional magnetic resonance study. *Biol. Psychiatry* **57**, 1079–1088. (doi:10.1016/j.biopsych.2005.02.021)
- Bar, M. 2003 A cortical mechanism for triggering top-down facilitation in visual object recognition. *J. Cogn. Neurosci.* **15**, 600–609. (doi:10.1162/089892903321662976)
- Bar, M. 2004 Visual objects in context. *Nat. Rev. Neurosci.* **5**, 617–629. (doi:10.1038/nrn1476)
- Bar, M. 2007 The proactive brain: using analogies and associations to generate predictions. *Trends Cogn. Sci.* **11**, 280–289. (doi:10.1016/j.tics.2007.05.005)
- Bar, M. Submitted A cognitive neuroscience hypothesis of mood and depression.
- Bar, M. & Aminoff, E. 2003 Cortical analysis of visual context. *Neuron* **38**, 347–358. (doi:10.1016/S0896-6273(03)00167-3)
- Bar, M. & Ullman, S. 1996 Spatial context in recognition. *Perception* **25**, 343–352. (doi:10.1068/p250343)
- Bar, M. & Neta, M. 2006 Very first impressions. *Emotion* **6**, 269–278. (doi:10.1037/1528-3542.6.2.269)
- Bar, M. *et al.* 2006 Top-down facilitation of visual recognition. *Proc. Natl Acad. Sci. USA* **103**, 449–454. (doi:10.1073/pnas.0507062103)
- Bar, M., Aminoff, E., Mason, M. & Fenske, M. 2007 The units of thought. *Hippocampus* **17**, 420–428. (doi:10.1002/hipo.20287)
- Barbey, A. K., Krueger, F. & Grafman, J. 2009 Structured event complexes in the medial prefrontal cortex support counterfactual representations for future planning. *Phil. Trans. R. Soc. B* **364**, 1291–1300. (doi:10.1098/rstb.2008.0315)
- Bargh, J. A. & Chen, M. 1996 Automaticity of social behavior: direct effects of trait construct and stereotype activation on action. *J. Pers. Soc. Psychol.* **71**, 230–244. (doi:10.1037/0022-3514.71.2.230)

- Barsalou, L. W. 2003 Abstraction in perceptual symbol systems. *Phil. Trans. R. Soc. B* **358**, 1177–1187. (doi:10.1098/rstb.2003.1319)
- Barsalou, L. W. 2009 Simulation, situated conceptualization, and prediction. *Phil. Trans. R. Soc. B* **364**, 1281–1289. (doi:10.1098/rstb.2008.0319)
- Bartlett, F. C. 1932 *Remembering: a study in experimental and social psychology*. Cambridge, UK: Cambridge University Press.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Rao, S. M. & Cox, R. W. 1999 Conceptual processing during the conscious resting state. A functional MRI study. *J. Cogn. Neurosci.* **11**, 80–95. (doi:10.1162/089892999563265)
- Brewer, W. F. & Nakamura, G. V. 1984 The nature and functions of schemas. In *Handbook of social cognition* (eds R. S. Wyer & T. K. Srull), pp. 119–160. Hillsdale, NJ: Erlbaum.
- Buckner, R. L. & Carroll, D. C. 2007 Self-projection and the brain. *Trends Cogn. Sci.* **11**, 49–57. (doi:10.1016/j.tics.2006.11.004)
- Burgess, N., Maguire, E. A., Spiers, H. J. & O'Keefe, J. 2001 A temporoparietal and prefrontal network for retrieving the spatial context of lifelike events. *Neuroimage* **14**, 439–453. (doi:10.1006/nimg.2001.0806)
- Cohen, J. D. & Aston-Jones, G. 2005 Cognitive neuroscience: decision amid uncertainty. *Nature* **436**, 471–472. (doi:10.1038/436471a)
- Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B. & Dolan, R. J. 2006 Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876–879. (doi:10.1038/nature04766)
- Dijksterhuis, A., Aarts, H., Bargh, J. A. & Van Knippenberg, A. 2000 On the relation between associative strength and automatic behavior. *J. Exp. Soc. Psychol.* **36**, 531–544. (doi:10.1006/jesp.2000.1427)
- Drevets, W. C. 2000 Functional anatomical abnormalities in limbic and prefrontal cortical structures in major depression. *Prog. Brain Res.* **126**, 413–431. (doi:10.1016/S0079-6123(00)26027-5)
- Drevets, W. C., Videen, T. O., Price, J. L., Preskorn, S. H., Carmichael, S. T. & Raichle, M. E. 1992 A functional anatomical study of unipolar depression. *J. Neurosci.* **12**, 3628–3641.
- Eichenbaum, H. & Bunsey, M. 1995 On the binding of associations in memory-clues from studies on the role of the hippocampal region in paired-associate learning. *Curr. Dir. Psychol. Sci.* **4**, 19–23. (doi:10.1111/1467-8721.ep10770954)
- Fleck, M. S., Daselaar, S. M., Dobbins, I. G. & Cabeza, R. 2006 Role of prefrontal and anterior cingulate regions in decision-making processes shared by memory and non-memory tasks. *Cereb. Cortex* **16**, 1623–1630. (doi:10.1093/cercor/bhj097)
- Frith, C. D. & Frith, U. 2006 The neural basis of mentalizing. *Neuron* **50**, 531–534. (doi:10.1016/j.neuron.2006.05.001)
- Gentner, D. 1983 Structure-mapping: a theoretical framework for analogy. *Cogn. Sci.* **7**, 155–170. (doi:10.1016/S0364-0213(83)80009-3)
- Hassabis, D. & Maguire, E. A. 2009 The construction system of the brain. *Phil. Trans. R. Soc. B* **364**, 1263–1271. (doi:10.1098/rstb.2008.0296)
- Holyoak, K. J. & Thagard, P. 1997 The analogical mind. *Am. Psychol.* **52**, 35–44. (doi:10.1037/0003-066X.52.1.35)
- Iacoboni, M., Lieberman, M. D., Lieberman, M. D., Knowlton, B. J., Molnar-Szakacs, I., Moritz, M., Jason Throop, C. & Page Fiske, A. 2004 Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. *Neuroimage* **21**, 1167–1173. (doi:10.1016/j.neuroimage.2003.11.013)
- Ingvar, D. H. 1985 'Memory of the future': an essay on the temporal organization of conscious awareness. *Hum. Neurobiol.* **4**, 127–136.
- Jenkins, R. & Burton, A. M. 2008 100% accuracy in automatic face recognition. *Science* **319**, 435. (doi:10.1126/science.1149656)
- Kelley, W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S. & Heatherton, T. F. 2002 Finding the self? An event-related fMRI study. *J. Cogn. Neurosci.* **14**, 785–794. (doi:10.1162/08989290260138672)
- Koechlin, E. & Jubault, T. 2006 Broca's area and the hierarchical organization of human behavior. *Neuron* **50**, 963–974. (doi:10.1016/j.neuron.2006.05.017)
- Kveraga, K. & Boshyan, J. 2007 Magnocellular projections as the trigger of top-down facilitation in recognition. *J. Neurosci.* **27**, 13 232–13 240. (doi:10.1523/JNEUROSCI.3481-07.2007)
- Macrae, C. N., Moran, J. M., Heatherton, T. F., Banfield, J. F. & Kelley, W. M. 2004 Medial prefrontal activity predicts memory for self. *Cereb. Cortex* **14**, 647–654. (doi:10.1093/cercor/bhh025)
- Maddock, R. J. 1999 The retrosplenial cortex and emotion: new insights from functional neuroimaging of the human brain. [see comments.]. *Trends Neurosci.* **22**, 310–316. (doi:10.1016/S0166-2236(98)01374-5)
- Maguire, E. A. 2001 The retrosplenial contribution to human navigation: a review of lesion and neuroimaging findings. *Scand. J. Psychol.* **42**, 225–238. (doi:10.1111/1467-9450.00233)
- Mayberg, H. S. & Liotti, M. 1999 Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. *Am. J. Psychiatry* **156**, 675–682.
- Mazoyer, B. *et al.* 2001 Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Res. Bull.* **54**, 287–298. (doi:10.1016/S0361-9230(00)00437-8)
- Miyashita, Y. 1993 Inferior temporal cortex: where visual perception meets memory. *Annu. Rev. Neurosci.* **16**, 245–269. (doi:10.1146/annurev.ne.16.030193.001333)
- Moores, E., Laiti, L. & Chelazzi, L. 2003 Associative knowledge controls deployment of visual selective attention. *Nat. Neurosci.* **6**, 182–189. (doi:10.1038/nn996)
- Moulton, S. T. & Kosslyn, S. M. 2009 Imagining predictions: mental imagery as mental emulation. *Phil. Trans. R. Soc. B* **364**, 1273–1280. (doi:10.1098/rstb.2008.0314)
- Mushiaki, H., Saito, N., Sakamoto, K., Itoyama, Y. & Tanji, J. 2006 Activity in the lateral prefrontal cortex reflects multiple steps of future events in action plans. *Neuron* **50**, 631–641. (doi:10.1016/j.neuron.2006.03.045)
- Muter, V. & Snowling, M. 1994 Orthographic analogies and phonological awareness: their role and significance in early reading development. *J. Child Psychol. Psychiatry* **35**, 293–310. (doi:10.1111/j.1469-7610.1994.tb01163.x)
- Noe, A. 2005 *Action in perception (Representation and mind)*. Cambridge, MA: The MIT Press.
- O'Craven, K. M. & Kanwisher, N. 2000 Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J. Cogn. Neurosci.* **12**, 1013–1023. (doi:10.1162/08989290051137549)
- Okuda, J. & Fujii, T. 2003 Thinking of the future and past: the roles of the frontal pole and the medial temporal lobes. *Neuroimage* **19**, 1369–1380. (doi:10.1016/S1053-8119(03)00179-4)
- Oliva, A. & Torralba, A. 2001 Modeling the shape of a scene: a holistic representation of the spatial envelope. *Int. J. Comp. Vision* **42**, 145–175. (doi:10.1023/A:1011139631724)

- Pasupathy, A. & Miller, E. K. 2005 Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* **433**, 873–876. (doi:10.1038/nature03287)
- Petrides, M. 1985 Deficits on conditional associative-learning tasks after frontal- and temporal-lobe lesions in man. *Hum. Neurobiol.* **4**, 137–142.
- Polli, F. E. & Barton, J. 2005 Rostral and dorsal anterior cingulate cortex make dissociable contributions during antisaccade error commission. *Proc. Natl Acad. Sci. USA* **102**, 15 700–15 705. (doi:10.1073/pnas.0503657102)
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A. & Shulman, G. L. 2001 A default mode of brain function. *Proc. Natl Acad. Sci. USA* **98**, 676–682. (doi:10.1073/pnas.98.2.676)
- Ranganath, C., Cohen, M. X., Daw, C. & D'Esposito, M. 2004 Inferior temporal, prefrontal, and hippocampal contributions to visual working memory maintenance and associative memory retrieval. *J. Neurosci.* **24**, 3917–3925. (doi:10.1523/JNEUROSCI.5053-03.2004)
- Saxe, R. & Kanwisher, N. 2003 People thinking about thinking people. The role of the temporo-parietal junction in 'theory of mind'. *Neuroimage* **19**, 1835–1842. (doi:10.1016/S1053-8119(03)00230-1)
- Schacter, D. L. 1987 Memory, amnesia, and frontal lobe dysfunction. *Psychobiology* **15**, 21–36.
- Schank, R. C. 1975 Using knowledge to understand. In *Theoretical issues in natural language processing* (eds R. C. Schank & B. Nash-Weber), Arlington, VA: Tinlap Press.
- Schubotz, R. I. & von Cramon, D. Y. 2002 Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: an fMRI study. *Neuroimage* **15**, 787–796. (doi:10.1006/nimg.2001.1043)
- Schultz, W. & Dickinson, A. 2000 Neuronal coding of prediction errors. *Annu. Rev. Neurosci.* **23**, 473–500. (doi:10.1146/annurev.neuro.23.1.473)
- Shallice, T., Fletcher, P., Frith, C. D., Grasby, P., Frackowiak, R. S. J. & Dolan, R. J. 1994 Brain regions associated with acquisition and retrieval of verbal episodic memory. *Nature* **368**, 633–635. (doi:10.1038/368633a0)
- Soares, J. C. & Mann, J. J. 1997 The anatomy of mood disorders: review of structural neuroimaging studies. *Biol. Psychiatry* **41**, 86. (doi:10.1016/S0006-3223(96)00006-6)
- Sperling, R. & Chua, E. 2003 Putting names to faces: successful encoding of associative memories activates the anterior hippocampal formation. *Neuroimage* **20**, 1400–1410. (doi:10.1016/S1053-8119(03)00391-4)
- Stark, C. E. & Squire, L. R. 2001 Simple and associative recognition memory in the hippocampal region. *Learn. Mem.* **8**, 190–197. (doi:10.1101/lm.40701)
- Summerfield, J. J. & Lepsien, J. 2006 Orienting attention based on long-term memory experience. *Neuron* **49**, 905–916. (doi:10.1016/j.neuron.2006.01.021)
- Suzuki, W. A. & Eichenbaum, H. 2000 The neurophysiology of memory. *Ann. NY Acad. Sci.* **911**, 175–191.
- Szpunar, K. K., Watson, J. M. & McDermott, K. B. 2007 Neural substrates of envisioning the future. *Proc. Natl Acad. Sci. USA* **104**, 642–647. (doi:10.1073/pnas.0610082104)
- Torralla, A. & Oliva, A. 2003 Statistics of natural image categories. *Network* **14**, 391–412. (doi:10.1088/0954-898X/14/3/302)
- Wagner, A. D. & Shannon, J. 2005 Parietal lobe contributions to episodic memory retrieval. *Trends Cogn. Sci.* **9**, 445–453. (doi:10.1016/j.tics.2005.07.001)
- Wang, K. & Jiang, T. 2008 Spontaneous activity associated with primary visual cortex: a resting-state fMRI study. *Cereb. Cortex* **18**, 697–704. (doi:10.1093/cercor/bhm105)
- Weissman, D. H., Roberts, K. C., Visscher, K. M. & Woldorff, M. G. 2006 The neural bases of momentary lapses in attention. *Nat. Neurosci.* **9**, 971–978. (doi:10.1038/nn1727)
- Williams, J. M., Ellis, N. C., Tyers, C., Healy, H., Rose, G. & MacLeod, A. K. 1996 The specificity of autobiographical memory and imageability of the future. *Mem. Cognit.* **24**, 116–125.
- Willis, J. & Todorov, A. 2006 First impressions: making up your mind after a 100 ms exposure to a face. *Psychol. Sci.* **17**, 592–598. (doi:10.1111/j.1467-9280.2006.01750.x)