



# Novelty Manipulations, Memory Performance, and Predictive Coding: the Role of Unexpectedness

Richárd Reichardt<sup>1,2\*</sup>, Bertalan Polner<sup>1</sup> and Péter Simor<sup>2,3,4</sup>

<sup>1</sup>Department of Cognitive Science, Budapest University of Technology and Economics, Budapest, Hungary, <sup>2</sup>Institute of Psychology, Eötvös Lóránd University, Budapest, Hungary, <sup>3</sup>Institute of Behavioural Sciences, Semmelweis University, Budapest, Hungary, <sup>4</sup>UR2NF, Neuropsychology and Functional Neuroimaging Research Unit, CRCN—Center for Research in Cognition and Neurosciences and UNI—ULB Neurosciences Institute, Université Libre de Bruxelles (ULB), Brussels, Belgium

## OPEN ACCESS

### Edited by:

Bradley Russell Buchsbaum,  
Baycrest Hospital, Canada

### Reviewed by:

Nico Bunzeck,  
University of Lübeck, Germany  
Avi Mendelsohn,  
University of Haifa, Israel

### \*Correspondence:

Richárd Reichardt  
rreichardt@cogsci.bme.hu

### Specialty section:

This article was submitted to  
Cognitive Neuroscience, a section of  
the journal *Frontiers in Human  
Neuroscience*

**Received:** 08 January 2020

**Accepted:** 06 April 2020

**Published:** 29 April 2020

### Citation:

Reichardt R, Polner B and Simor P  
(2020) Novelty Manipulations,  
Memory Performance, and Predictive  
Coding: the Role of Unexpectedness.  
*Front. Hum. Neurosci.* 14:152.  
doi: 10.3389/fnhum.2020.00152

Novelty is central to the study of memory, but the wide range of experimental manipulations aimed to reveal its effects on learning produced inconsistent results. The novelty/encoding hypothesis suggests that novel information undergoes enhanced encoding and thus leads to benefits in memory, especially in recognition performance; however, recent studies cast doubts on this assumption. On the other hand, data from animal studies provided evidence on the robust effects of novelty manipulations on the neurophysiological correlates of memory processes. Conceptualizations and operationalizations of novelty are remarkably variable and were categorized into different subtypes, such as stimulus, context, associative or spatial novelty. Here, we summarize previous findings about the effects of novelty on memory and suggest that predictive coding theories provide a framework that could shed light on the differential influence of novelty manipulations on memory performance. In line with predictive coding theories, we emphasize the role of unexpectedness as a crucial property mediating the behavioral and neural effects of novelty manipulations.

**Keywords:** memory, novelty, surprise, expectation, predictive coding

## INTRODUCTION

Any notion to be explored by the scientific community needs to be defined unambiguously and novelty is not an exception. The so-called common sense or memory-based definition states that novelty is any aspect of a percept that is not already contained in the memory systems of the observer (Berlyne, 1960; Barto et al., 2013). Accordingly, any novelty experienced by the observer requires novel representations to be generated for that stimulus or event. The novelty/encoding hypothesis postulates that this newly created representation is more easily reactivated later, yielding enhanced recognition on the behavioral level (Tulving and Kroll, 1995).

The novelty/encoding hypothesis was partly based on the early findings of positron emission tomography (PET) studies showing that novel stimuli elicited activations in the hippocampal formation, the medial dorsal nucleus of the thalamus and the anterior and inferior parts of the cingulate cortex (Tulving et al., 1994, 1996). In a behavioral

study inspired by these findings, participants had to learn lists containing words, some of which were familiarized before the experiment, while others were novel to the participants. In line with the novelty/encoding hypothesis, recognition rates for novel words were significantly higher than those for familiar words (Tulving and Kroll, 1995).

Nevertheless, the interpretation of the data was soon challenged—it was pointed out that hit rates for novel vs. familiar stimuli may have differed because the false alarm rate for familiar words was relatively higher, suggesting that the results may have stemmed from an interference effect within the familiar stimulus set (Dobbins et al., 1998). This interference-based explanation of the original findings, however, failed to completely refute the novelty/encoding hypothesis, as other studies still supported a memory-enhancing effect of novelty. For instance, a repetition-based study showed that the probability of recognition for the novel (and rare) stimuli was linked to the activation in prefrontal and temporal regions during encoding (Kirchhoff et al., 2000), while another study evidenced the novelty effect in a verbal recognition memory test (Kormi-Nouri et al., 2005).

Still, more recent comprehensive reviews tend to conclude that the novelty/encoding hypothesis is not unequivocally supported by empirical studies (Poppenk et al., 2010; Schomaker and Meeter, 2015). However, a parallel line of animal studies produced more consistent results. These inquiries revealed a clear association between novelty manipulations and electrophysiological and molecular markers of memory processes, such as dopaminergic modulation of long-term potentiation (LTP) in the hippocampus (Lisman and Grace, 2005). In these studies, the experimental animals were usually placed in a novel environment, while electrophysiological or neurochemical variables were registered (Floresco et al., 2001; Legault and Wise, 2001).

The results suggest that novelty is detected by the hippocampus and through its connections to the ventral tegmental area, the detection of novelty can elicit dopamine release in the hippocampus, facilitating LTP at the activated synapses (Lisman and Grace, 2005; Shohamy and Adcock, 2010). This idea can be viewed as a neurobiological formulation of the novelty/encoding hypothesis. Thus, the literature on the memory effects of novelty is rather conflicting—on the one hand, human studies produced controversial results, on the other hand, animal studies provided support for the novelty effect. The large variability of novelty manipulations urged the field to define their distinct subtypes.

## THE CATEGORIZATION OF NOVELTY MANIPULATIONS

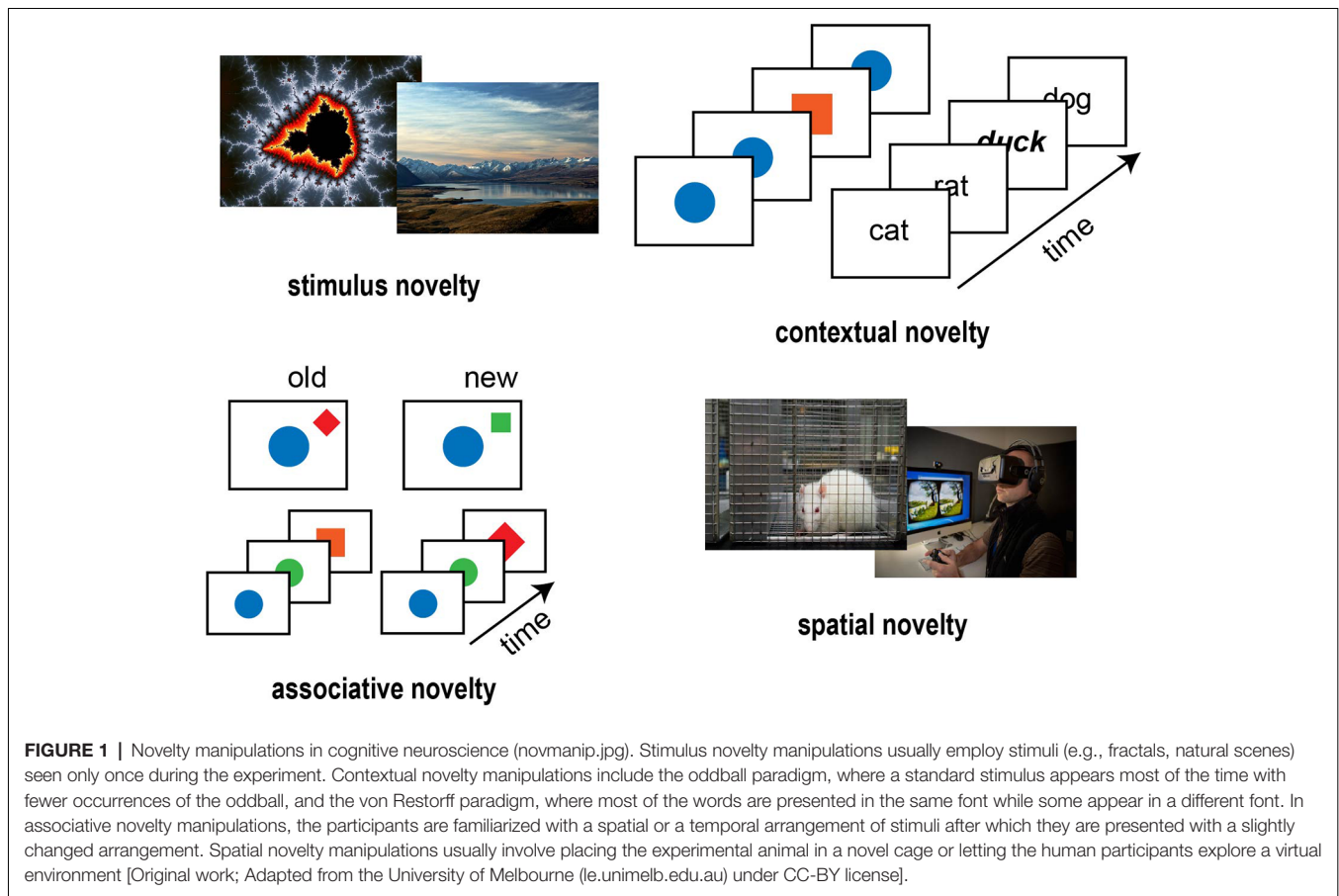
The first attempts at the categorization of novelty were based on the temporal aspect (Berlyne, 1960; Barto et al., 2013). Berlyne distinguished complete novelty, where the organism has never encountered the stimulus before. He also differentiated short-term and long-term novelty, where the stimulus has been encountered before, but not in the last few minutes or the last few days, respectively (Berlyne, 1960). The temporal aspect of novelty is usually not addressed in experimental studies; thus,

we can reasonably suppose that most manipulations are aimed to produce complete novelty.

Berlyne also suggested that the amount of novelty in each percept is an important aspect of novelty. He differentiated absolute novelty, which is supposedly a feature that has never been encountered before, while relative novelty is a novel arrangement of familiar features. Berlyne argued that an experienced observer cannot be presented with absolute novelty, as any percept can be related to previous experiences in a meaningful way, as to be only relatively novel. His classic example is that of seeing the tallest man for the first time: this experience may be considered novel, but the man probably only differs significantly in height from other people one has seen (Berlyne, 1960). As absolute novelty may only be experienced by younger children, we can conclude that the current literature studying adult humans and non-human animals is preoccupied with relative novelty.

Although Berlyne's ideas are certainly of value for the field, the categories he proposed cannot be used to differentiate between the breadth of novelty manipulations appearing in research. Most authors usually distinguish stimulus novelty and contextual novelty (Ranganath and Rainer, 2003), but some also consider associative novelty or even spatial novelty as distinct categories (Nyberg, 2005; Barto et al., 2013; Schomaker and Meeter, 2015; **Figure 1**). First, stimulus novelty manipulations, probably recognized by most as the archetypal cases of novelty, intend to use stimuli that have never been encountered by the observer before (Barto et al., 2013; Schomaker and Meeter, 2015). A range of different stimuli has been applied in paradigms, where stimulus novelty was manipulated, for example, fractals, simple shapes, natural scenes, and words. Importantly, it is methodically challenging to test memory with these paradigms, since memory performance must be compared across novel vs. well-known stimuli, an inherently biased comparison. This obstacle has been overcome by some paradigms, however, the experimental results obtained with these can be considered mixed on the existence of a clear, stimulus-level memory effect of novelty (Poppenk et al., 2010; Schomaker and Meeter, 2015; Reggev et al., 2017). For example, the classical novelty effect has been shown to result from a source discrimination problem relating to the familiar stimuli (Dobbins et al., 1998; however, see Kormi-Nouri et al., 2005; and Wittmann et al., 2007; for examples of stimulus-level memory effects of novel stimuli).

Second, contextual novelty is defined as a stimulus or an event unexpectedly arising in a given context (Ranganath and Rainer, 2003; Nyberg, 2005). The studies listed under the contextual novelty category mostly utilize some form of the oddball task, where a standard stimulus is presented frequently with the occasional appearance of an oddball stimulus (Ranganath and Rainer, 2003; Barto et al., 2013; Schomaker and Meeter, 2015). Another similar paradigm that falls under the contextual novelty category is the von Restorff paradigm (Nyberg, 2005)—in this paradigm, the subjects learn a list of words, with some words differing in font type or font size (von Restorff, 1933; Karis et al., 1984; Geraci and Manzano, 2010; Schomaker and Meeter, 2015). Contrasting with stimulus novelty, these studies usually show that contextual novelty produces robust benefits on memory



performance (Polich, 2007; Barto et al., 2013; Schomaker and Meeter, 2015).

Third, associative novelty is defined as a novel arrangement of familiar stimuli (Nyberg, 2005; Kumaran and Maguire, 2007). Paradigms utilizing associative novelty usually present a spatial or temporal configuration of distinct stimuli in a learning phase and assess the recognition responses of the participants in a test phase, where the arrangements are sometimes shuffled to elicit associative novelty. These experiments are indirectly associated with memory enhancements (Nyberg, 2005; Barto et al., 2013), as associative novelty elicits an increase in the blood-oxygen-level-dependent (BOLD) signal in the hippocampus and the dopaminergic midbrain (Düzel et al., 2003; Schott et al., 2004; Köhler et al., 2005; Kumaran and Maguire, 2006, 2007).

Finally, spatial novelty is distinguished by some authors by their robust effects on memory (Schomaker and Meeter, 2015; Schomaker, 2019). Spatial novelty is a novel spatial environment, and thus spatial novelty manipulations use more complex stimuli than the other categories—in human studies, this involves virtual reality environments (Schomaker et al., 2014b), while in animal studies the subjects are placed in novel cages for example (Jeewajee et al., 2008). These manipulations usually produce marked behavioral manifestations and several physiological correlates of memory enhancements (Schomaker and Meeter, 2015; Schomaker, 2019).

However, this categorization scheme is not without criticism. For example, it has been suggested that contextual and associative novelty manipulations produce memory effects due to unexpectedness, interpreting the reported memory enhancements problematic (Barto et al., 2013). Furthermore, this categorization does not explain the inconsistent findings concerning the effects of novelty on memory performance. Why does one sort of novelty manipulation exert a demonstrable effect on memory, whereas another does not? This question may be answered by the predictive coding framework applied to the functioning of memory systems.

## PREDICTIVE CODING AND NOVELTY

The idea that the brain is a predictive system, constantly trying to predict its sensory inputs to ease processing and thus conserve energy, is central to several similar formulations, which we refer to as predictive coding theories (Knill and Pouget, 2004; Friston, 2005, 2010; Clark, 2013). There is ample evidence that such a computational strategy might indeed be utilized in several functional units of the brain (den Ouden et al., 2012; Hyman et al., 2017; de Lange et al., 2018).

The main idea of predictive coding is that the brain constantly uses an inner model to produce predictions about the sensory information that it is likely to encounter. The predictions are

compared to the incoming information and only the difference, the prediction error is processed further, which is used to update the inner model so that next time in a similar situation the brain will be able to generate more accurate predictions. The predictive coding framework suggests that the comparisons required to generate prediction errors are made on several different levels in the processing hierarchy of the brain (Knill and Pouget, 2004; Friston, 2005, 2010). The scenario offered by predictive coding is highly relevant for memory research, as one might assume that the inner model, or at least a part of it, could be equated with memory itself and that learning can be considered as the updating mechanism driven by prediction errors (Henson and Gagnepain, 2010; van Kesteren et al., 2012). It is also important to note, however, that prediction errors are also generated on the lower levels of the hierarchy, which may serve veridical perception without influencing information processing at higher levels. This way, prediction errors at lower levels of the processing hierarchy may not necessarily have an impact on learning and memory.

If we take this idea and turn to the memory effects of novelty, however, we find an apparent contradiction. Novelty, by definition, is something that is not contained in the memory systems or the inner model of the observer. As such, it is impossible to accurately predict novelty, therefore it should always generate prediction errors and lead to learning. However, as we have discussed before, this simplistic interpretation is not unequivocally supported by experimental data (Poppenk et al., 2010; Schomaker and Meeter, 2015; Reggev et al., 2017). It could be, that some novelty manipulations employ stimuli that are not deemed important enough by the human information processing system to be encoded and consolidated in long-term memory.

First, we stress that the aspects of our inner models that we can access consciously (which is usually referred to as explicit memory), can only be scarcely detailed (Hobson and Friston, 2012). On the one hand, this assumption is supported by anecdotal evidence from the phenomenology of imagination and lucid dreaming. These processes are thought to be internally generated (using the inner model) and lacking in detail. A thought experiment supports this idea in the case of imagination—when we try to imagine a tiger, we can quickly conclude that minor perceptual details, such as the patterns on the fur, are not stored accurately, as most of us are unable to imagine them precisely (Chater, 2019). People who frequently have lucid dreams also report that during lucid dreaming they are unable to focus on minor visual details of surfaces (Nir and Tononi, 2010), which suggests that this information is not stored in the inner model. On the other hand, theoretical considerations also suggest that the inner model cannot be infinitely detailed, as it would result in a highly unstable system, which requires constant updating, since highly detailed predictions would always produce considerable prediction errors throughout the processing hierarchy (Friston, 2010; Hobson and Friston, 2012). If the inner model is not infinitely detailed, observers may encounter perceptual novelty that does not generate prediction errors driving learning on this level of the hierarchy, simply because the details in question are not stored in those parts of the inner model. This line of thought also suggests that not

everything novel will necessarily generate a novel representation in these levels of hierarchy (i.e., in episodic and semantic memory systems). Supposedly, representations enabling veridical perception form on lower-level sensory areas of the brain (den Ouden et al., 2012; de Lange et al., 2018).

Second, we propose that unexpectedness must be of a certain degree to facilitate learning. This idea is most prominently defined in the literature on surprise. However, the use of the term surprise is also somewhat complicated. On one hand, it refers to an emotion that includes a range of physiological, experiential, behavioral and cognitive components, and is elicited by unexpected stimuli or events. On the other hand, surprise can also be defined in an information theory framework as a measure reflecting to what extent is certain sensory data unlikely under a certain belief held by an observer (Friston, 2010; O'Reilly et al., 2013; Seer et al., 2016; Faraji et al., 2018). The physiological and behavioral responses triggered by unexpected stimuli or events are collectively termed the orienting response (Sokolov, 1963; Alexander and Brown, 2019; Reizenzein et al., 2019). This includes directing the appropriate sensory organs of the observer to the source of the discrepancy, skin conductance changes, heart rate deceleration, and pupil dilation. The cognitive processes elicited by unexpectedness are termed the surprise response, according to the cognitive-evolutionary model of surprise (Reizenzein et al., 2019). The surprise response consists of the interruption of the ongoing cognitive processes and a shift of attention to the cause of the discrepancy, which may be accompanied by the conscious experience of surprise, and finally, the learning of the unexpected information. The generation of the surprise response is thought to be based on a comparison between expected and observed inputs, largely paralleling the computation of prediction errors (see above). When the outcome of this comparison exceeds a threshold, a surprise signal is generated, which drives the responses to unexpected events. The idea that a certain threshold needs to be overcome by any given event to elicit these responses is derived from the observation that not any unexpected event readily evokes them (Reizenzein et al., 2019). For example, the surprise response may be elicited by the appearance of a novel coworker, while we may completely ignore a novel item of clothing on a familiar colleague, or a novel item on his/her desk. These examples can be considered cases of stimulus novelty, where the relevance of the stimulus in question differs profoundly just as its effects on cognition. It is conceivable that if we are very interested in clothes or the item in question is the only one on an otherwise empty desk, the surprise response is evoked: motivation, the complexity of the environment or uncertainty may all influence the threshold (i.e., the degree of unexpectedness needed to evoke the surprise response). These ideas fit well with the framework of predictive coding and the suggestion that the degree of prediction errors is central to learning helps to organize the disparate results of novelty manipulations: if unexpectedness elicited by a novelty manipulation reaches a certain threshold, the surprise response is evoked and learning is enhanced.

Here, we propose that to be memorable, novelty must evoke the surprise response, and the reason for the inconsistency in the literature on the memory effects of novelty is that not all novelty manipulations can consistently elicit this phenomenon.



In the following sections, we will review the research findings supporting this view.

## STIMULUS NOVELTY: PREDICTION ERRORS MAY ONLY SERVE VERIDICAL PERCEPTION

Stimulus novelty is one of the basic categories of novelty manipulations in cognitive neuroscience. Repetition suppression and the novelty N2 are neural responses reliably elicited by stimulus novelty manipulations. Several authors approach stimulus novelty based on its neural effects, or more precisely, the difference between the effects elicited by familiar and novel stimuli (Ranganath and Rainer, 2003; Schomaker and Meeter, 2015). In this view, repetition suppression, which refers to the decreased neural response to repeated stimuli, is the complement of the stimulus novelty effect (i.e., increased neural activation in response to novel stimuli; Ranganath and Rainer, 2003; Nyberg, 2005; Barto et al., 2013). Repetition suppression is a gradual reduction in neural activity in response to repeated items that have been demonstrated in several modalities with numerous methods, ranging from single-cell recordings to functional magnetic resonance imaging (Ranganath and Rainer, 2003; Grill-Spector et al., 2006; Grotheer and Kovács, 2016). A crucial study showed that repetition suppression is modulated by expectation (Summerfield et al., 2008), a finding fitting neatly in the predictive coding scheme, but defying the classic theories of repetition suppression, which postulate that repetition suppression results from inherent properties of neural networks—not subject to higher-level computational effects.

In a predictive coding framework, the activity in the sensory cortices reflects prediction errors, and thus repetition suppression could be a reduction in prediction errors, which may reflect a more accurate prediction of a known stimulus. In this scheme it is conceivable that predictions from higher levels arrive as an inhibition to the neurons exhibiting repetition suppression, hence decreasing the prediction errors that the same input caused before (Friston, 2005; den Ouden et al., 2012; Gotts et al., 2012; Auzsztulewicz and Friston, 2016; Grotheer and Kovács, 2016). The finding of Summerfield et al. (2008) that the degree of repetition suppression depends on the probability of repetition, came to be called expectation suppression. Later studies were able to differentiate between repetition suppression and expectation suppression, showing that repetition suppression is an earlier effect which could be followed by expectation suppression (Todorovic and de Lange, 2012; Grotheer and Kovács, 2015). These effects may be epiphenomena caused by the bottom-up traveling of the prediction error, repetition suppression being the expression of a lower level computation and expectation suppression resulting from higher-level prediction errors (Grotheer and Kovács, 2016). Curiously, studies using complex visual stimuli (e. g. fractals, false fonts, natural scenes), the archetypical cases of stimulus novelty, were unable to demonstrate expectation suppression (Kaliukhovich and Vogels, 2011; Kovács et al., 2013; Grotheer and Kovács, 2014). Therefore, repetition suppression seems to

depend on the subjects' expertise with the stimuli (Grotheer and Kovács, 2016), as unfamiliar complex stimuli are harder to predict accurately. We can conclude that although stimulus novelty elicits repetition suppression reliably to enable veridical perception, it does not trigger higher level prediction errors that can elicit expectation suppression or the surprise response and enhance memory.

Stimulus novelty manipulations may also elicit an event-related potential (ERP) component, that is, the anterior N2. The anterior N2 is an ERP, which peaks over frontal regions around 250–300 ms after stimulus presentation in the visual modality and it is the second negative peak following stimulus presentation (Folstein and Van Petten, 2008). It is important to note that the anterior N2 is made up of two different components, one of which is associated with cognitive control and the other one with mismatch detection (Folstein and Van Petten, 2008) or novelty detection (Schomaker and Meeter, 2015). The exact interpretation of the anterior N2 is still debated; however, several authors have proposed that the anterior N2 is an index of novelty processing (Chong et al., 2008; Tarbi et al., 2011; Schomaker et al., 2014a; Schomaker and Meeter, 2015, 2018). The anterior N2 was first observed in oddball paradigms (Folstein and Van Petten, 2008). The simplest of these paradigms is the two-stimulus oddball task, in which the participant is presented with a frequent standard stimulus, requiring no response, and a rare target, which requires a response. Usually, the rare stimulus elicits an anterior N2 (Folstein and Van Petten, 2008), which might appear with a larger amplitude when these stimuli are trial-unique, that is, they appear in only one trial (Breton et al., 1988). Studies with the visual oddball paradigm showed that the appearance and amplitude of the N2 depends on the degree of difference between the perceptual features of frequent and rare stimuli (Courchesne et al., 1975; Thomas and Nelson, 1996; Daffner et al., 2000; Czigler and Balázs, 2005; Barkaszi et al., 2013; Ferrari et al., 2015). A recent study found that expected (vs. unexpected) novel pictures elicited a higher N2, indicating that expectations can enhance the detection of novelty (Schomaker and Meeter, 2018). Based on these findings, several authors proposed that the anterior N2 also reflects low-level prediction errors, which are necessary for veridical perception when there is a considerable change in the incoming visual information (den Ouden et al., 2012; Stefanics et al., 2014; Grotheer and Kovács, 2016).

In sum, two of the most characteristic neural phenomena elicited by stimulus novelty manipulations, the anterior N2, and repetition suppression, seem to reflect low-level prediction errors, which are thought to be necessary for veridical perception (den Ouden et al., 2012; Stefanics et al., 2014; Grotheer and Kovács, 2016). Since veridical perception presumably depends on the statistical model of the visual environment represented in specific neural networks, stimuli that can be processed by the current structure of such networks should not elicit any change on these levels (that is, no visual stimulus for an adult human). Also, the stimuli are probably not unexpected enough after several repetitions to necessitate a change in higher-order networks representing the statistics of the task at hand (i.e., a novel fractal elicits only low-level prediction errors, as its details are processed by the sensory system, however it fails to elicit

higher-order prediction errors, as the appearance of another novel fractal is already predictable). Thus, it is conceivable that these prediction errors do not necessitate a lasting update in the inner model, stimulus novelty manipulations do not inherently elicit the surprise response and thus have no inherent effects on memory formation.

On the other hand, several studies seem to suggest that novel stimuli elicit neurophysiological responses associated with memory processes (Bunzeck and Düzel, 2006; Wittmann et al., 2007). Since these studies are mostly built on a neurobiological theory of learning based on animal studies using spatial novelty manipulations, they will be discussed in-depth in a later section of this review, but it is important to note here, that the responses ascribed to novelty in this line of research have been suggested to result from the coinciding unexpectedness (Barto et al., 2013; Schultz, 2016).

## THE SURPRISE RESPONSE IS CONSISTENTLY ELICITED BY OTHER NOVELTY MANIPULATIONS

In the next sections, we will examine contextual, associative and spatial novelty manipulations. The studies suggest that these novelty manipulations consistently elicit the surprise response and produce memory-boosting effects (Ranganath and Rainer, 2003; Nyberg, 2005; Schomaker and Meeter, 2015).

### Contextual Novelty

Studies utilizing contextual novelty manipulations probably elicit strong expectations about the presented stimuli, since the standards are very frequent and thus highly probable to appear (Folstein and Van Petten, 2008; Barto et al., 2013; Schomaker and Meeter, 2015). It seems that in these tasks the appearance of the oddballs elicits surprise, as self-reports and the appearance of the surprise response (or orienting response) suggest (Reisenzein et al., 2019).

The P3 ERP component is also evoked in these oddball paradigms (Polich, 2007; Nieuwenhuis et al., 2011) and since the orienting response and the P3 ERP are elicited by highly similar conditions, it has been suggested that the P3 is the neural correlate of the orienting response (Nieuwenhuis et al., 2011). The P3 component can be divided into subcomponents, of which the P3b (or the classical P3 or P300) is generated in the oddball tasks. It has a parieto-central scalp distribution and occurs with 300–400 ms latency after stimulus presentation (Friedman et al., 2001; Polich, 2007; Folstein and Van Petten, 2008; Nieuwenhuis et al., 2011; Schomaker and Meeter, 2015). The amplitude of the P3b is sensitive to the probability of occurrence, the degree of deviance from the standards, the motivational significance of the stimulus and also the degree of attention paid to it (Ranganath and Rainer, 2003; Nieuwenhuis et al., 2011), while the P3b is not sensitive to the physical attributes of the eliciting stimulus (Folstein and Van Petten, 2008; Nieuwenhuis et al., 2011). The relative difference to the standards seems to be more important than the objective or absolute complexity of the stimuli, suggesting that expectancy and not novelty is the central variable determining this response.

Also, the P3b tends to habituate as more and more novel stimuli are presented (Nieuwenhuis et al., 2011) and it is conceivable that, as expectation of novelty increases, the elicited prediction errors become insufficient to elicit a surprise response. The von Restorff paradigm could be considered an oddball task with an attached memory test, therefore its effects probably depend on the unexpectedness of the different typesetting, and indeed, deviant words in this task also elicit the P3b (Karis et al., 1984). To sum up, these findings suggest that the P3b is not elicited by novelty *per se*, but it is related to the unexpectedness and presumably the surprise response in studies using contextual novelty manipulations (Schomaker and Meeter, 2015; Reisenzein et al., 2019).

The association between surprise, the orienting response, and the P3 ERP component is close (Nieuwenhuis et al., 2011; Barto et al., 2013; Reisenzein et al., 2019), and thus the neural basis of the P3 is relevant to the neural basis of the surprise response. Some authors proposed that a hippocampal computation is the origin of the P3 (Jeewajee et al., 2008), while others point to parts of the cerebral cortex as the source (Polich, 2007). A concise review of the relevant literature claims that the locus coeruleus (LC) has a major role in generating the P3 (Nieuwenhuis et al., 2011), and the release of norepinephrine in the cortex may be the proximal cause of the P3. The latency of the LC response (~150–200 ms) is in line with the latency of the P3, considering the slow conduction velocity of LC fibers (Nieuwenhuis et al., 2005, 2011). Noradrenaline release is assumed to enhance the signal-to-noise ratio in the affected brain areas, which might be a preparative step to enhance the further analysis of the surprising stimulus, corresponding to the attentional shift in the cognitive-evolutionary model of surprise (Reisenzein et al., 2019). Interestingly, a theory on the function of the P3 suggests that it is a neural correlate of cognitive inhibition (Polich, 2007), which corresponds to the interruption proposed by the cognitive-evolutionary model of surprise. On the other hand, recent studies showed that the LC influences hippocampal memory encoding by releasing dopamine (Kempadoo et al., 2016; Wagatsuma et al., 2018). It seems that the LC elicits the P3 through its effects on the cortex (Nieuwenhuis et al., 2011), while it also alters hippocampal function (Kempadoo et al., 2016; Wagatsuma et al., 2018). Therefore, the hippocampal response could be a correlate, and not the cause of the P3 (Jeewajee et al., 2008). These experiments show that several contextual novelty manipulations can elicit the surprise response and thus can be expected to have memory effects, consistent with the results (Ranganath and Rainer, 2003; Kafkas and Montaldi, 2018).

### Associative and Spatial Novelty

Several studies showed increased hippocampal and midbrain (probably dopaminergic) BOLD activity in response to associative novelty (Düzel et al., 2003; Schott et al., 2004; Köhler et al., 2005; Kumaran and Maguire, 2006, 2007). The memory effects reported in these studies were explained by the influential hippocampal-VTA loop model (Lisman and Grace, 2005). This model organized hippocampal and midbrain responses to novel stimuli into a model, explaining the entry of information into memory.

The main idea is that the CA1 region of the hippocampus compares incoming to stored information and thus computes a novelty signal, in less than 100 ms after stimulus presentation (Ruusuvirta et al., 1995; Brankazk et al., 1996). This signal is transferred to the ventral tegmental area through the downward arc of the hippocampal-VTA loop, which involves the subiculum (Legault and Wise, 2001; Floresco et al., 2003), the nucleus accumbens and the ventral pallidum (Floresco et al., 2001). The ventral tegmental area innervates the hippocampus (Gasbarri et al., 1994a,b) and releases dopamine, which enhances LTP in the hippocampus (Shohamy and Adcock, 2010; Lisman et al., 2011). It is important to note that the studies serving as the basis for this model are mostly animal studies, in which novelty is usually achieved by placing the animals in a novel environment (Floresco et al., 2001; Legault and Wise, 2001).

However, the midbrain dopaminergic responses reported in several studies using associative and stimulus novelty manipulations (Schott et al., 2004, 2006; Bunzeck and Düzel, 2006; Bunzeck et al., 2007; Wittmann et al., 2007, 2008; Krebs et al., 2009) may have been caused by the unexpectedness and not by novelty *per se* (Barto et al., 2013; Schultz, 2016). The reasoning behind this statement builds up from three different clues. First, a study showed that meaningless novel stimuli do not reliably elicit midbrain activations (Stoppel et al., 2009). Second, the study showed directly that dopaminergic midbrain neurons change their firing patterns in response to novel stimuli used light flashes, which did not change during the experiment (Ljungberg et al., 1992), thus they cannot be considered to evoke stimulus novelty. Several authors suggested that it was the unpredictability of the flashes that elicited dopaminergic activity, and novelty is not enough in itself to elicit this response (Barto et al., 2013; Schultz, 2016). Third, it has been shown that hippocampal BOLD responses to novel stimuli quickly habituate as more and more novel items are presented (Murty et al., 2013). These findings suggest that both the hippocampal and the VTA responses are elicited by the unexpectedness and not by novelty.

In an animal study thought to utilize associative novelty, rats have been trained in a maze with simple visual stimuli outside the maze as landmarks (Jenkins et al., 2004). The position of the well-known extra-maze cues was changed on the critical trial and thus familiar stimuli appeared in a novel spatial arrangement, which activated the hippocampus, as indicated by increased c-Fos expression in this region. In another study, it has been shown that the hippocampal theta rhythm is sharply reduced when an animal encounters novelty in its environment, but this is more prominent for unexpected changes in a familiar environment than to a novel environment (Jeewajee et al., 2008). These results seem to suggest that the hippocampal match-mismatch computations compare expectations to actual sensory input, therefore it is also a response to unexpectedness. Since the response pattern of the hippocampus in these studies overlaps with the elicitation requirements of the P3, it has been hypothesized that the reduction in hippocampal theta generates the P3 (Jeewajee et al., 2008). As we discussed, the P3 ERP component is associated with the orienting response, it may be a concomitant

of the orienting response (Nieuwenhuis et al., 2011) and as such, a neural correlate of surprise (Barto et al., 2013; Reizenzein et al., 2019).

Spatial novelty as an entirely distinct category is proposed by Schomaker and Meeter, based on the fact that it has a well-described, long-lasting effect on memory (Schomaker and Meeter, 2015). For example, a study showed that viewing novel pictures depicting indoor or outdoor scenes before learning words improves the recollection and the recall of these words (Fenker et al., 2008), while another demonstrated that the exploration of a novel virtual environment before learning produces an increase in the recall of these words (Schomaker et al., 2014b). The studies that established the distinctive memory effect of spatial novelty examined LTP in the hippocampus of rats exploring a novel spatial environment (Straube et al., 2003; Davis et al., 2004; Uzakov et al., 2005) or the BOLD response in the hippocampus of humans performing memory tests involving spatial stimuli (Fenker et al., 2008; Schomaker et al., 2014b). The neural correlates of spatial novelty manipulations are similar to what we have already seen associated with contextual and associative novelty manipulations, thus, it seems reasonable to suggest that the driving force behind them may also be the surprise signal. Space is a central perceptual dimension which is always highly relevant to an organism capable of motion, thus, it is highly probable that unexpected changes in this dimension, like those elicited during the exploration of a novel (let it be real or virtual) environment, generate the surprise signal and drive learning. However, several studies show that these enhancements can be detected even after the exploration of novelty (Fenker et al., 2008; Schomaker et al., 2014b). A possible explanation is that during the exploration of a novel environment surprise signals are generated with a higher frequency and this results in a prolonged enhancement of learning. This would be in line with the “penumbra hypothesis,” which states that increased dopamine availability, which may be a concomitant of the surprise response, benefits memory encoding in general and not just the encoding of the stimulus that elicited its release (Lisman et al., 2011).

In sum, it seems that in associative and spatial novelty manipulations novelty inherently coincides with high unexpectedness. Since the neurophysiological responses associated with these manipulations have been shown to appear in response to unexpected but familiar stimuli, habituate or even fail to appear when supposedly only lower-level prediction errors are elicited by the appearance of a stimulus, the surprise response may provide a more parsimonious explanation for them than a novelty effect. The neurophysiological correlates found in these studies are suggestive of the surprise response—hippocampal activation and the P3 ERP component have been shown to habituate to novelty (Nieuwenhuis et al., 2011; Murty et al., 2013), suggesting that unexpectedness is the relevant stimulus dimension for these responses, as the stimuli are always novel, but the predictability of the occurrence of a novel stimulus increases during the task. It is important to note that the confounding of novelty manipulations by surprise may only be problematic



in human studies, while in animal studies novelty and surprise usually coincide because the experimental manipulations are less subtle in this field. Using the same terminology in experimental designs aimed at human participants may lead to confusing results on the memory effects of novelty (Poppenk et al., 2010; Barto et al., 2013; Schomaker and Meeter, 2015; Reggev et al., 2017), as in these experiments the gradual change in unexpectedness may be more impactful for memory formation.

## CONCLUSION

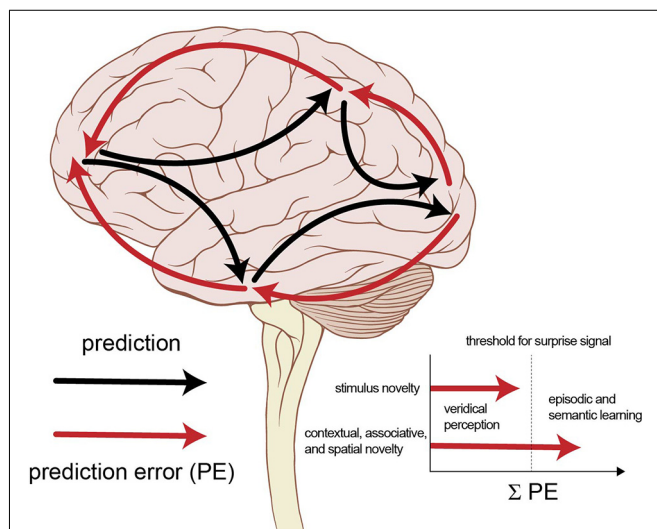
Here, we propose that the memory-enhancing effects of various novelty manipulations used in cognitive neuroscience mostly depend on the degree of unexpectedness (Figure 2). We suggest that inconsistencies of the reported memory effects are in part due to the different degree of unexpectedness induced by the manipulations. Some stimulus novelty manipulations may only elicit perceptual prediction errors in the lower levels of the processing hierarchy, which is required for veridical perception, but the prediction errors generated not always reach the threshold to elicit the surprise response and thus, do not lead to enhanced learning. Contextual, associative and spatial novelty manipulations, however, are often unexpected in addition to being novel and usually elicit greater prediction errors, which result in the updating of the inner model, that is, learning. Spatial

novelty could be a special case as it might involve a greater amount of novel information than the other types of novelty, we considered, or due to a special memory effect of spatial information. More work emphasizing the difference between unexpectedness and novelty is needed to organize previous findings meaningfully and to uncover the intricacies of the neurocognitive processes involved in the processing and learning of novel and surprising stimuli or events.

Considering the degree of unexpectedness when it comes to generating a hypothesis about the probable memory effects of a novelty manipulation would be more helpful than judging its similarity to some of the flagship paradigms in a given category. Several recent studies already hinted at this conclusion (Greve et al., 2017; Pine et al., 2018). However, the exact degree of unexpectedness depends on the inner model of the observer, to which the experimenter does not have direct access. Therefore, future studies ought to use an easily quantifiable experimental method, where the unexpectedness of every experimental event can be estimated as a function of the previous events of the task, or design experimental tasks where unexpectedness and novelty are orthogonally varied, to show that unexpectedness drives learning. The degree of unexpectedness is likely to be represented in the brain, but it does not have a single, unambiguous neural correlate (Bach and Dolan, 2012), which further complicates its assessment. Future studies should aim to clarify the neural correlates of surprise, in hopes of providing the means to assess the ability of different experimental manipulations to elicit the surprise response and judge the source of the memory effects more precisely.

Finally, we would like to address a conspicuous omission: this article did not touch upon the effects other variables may have on memory for novel stimuli or events, e.g., top-down control, motivation, stimulus material, retention time, testing method. The motivation of the learner or even her familiarity with the material may have major effects on learning (van Kesteren et al., 2012; Miendlarzewska et al., 2016). These variables may be integrated into the present account as eliciting their effects by changing the threshold for the elicitation of the surprise response, which is modulating the strength of the prediction error required for learning. However, it is important to note that this account of surprise elicited learning only details the starting step of the process: that unexpected stimuli may direct attention and prepare the cognitive system for memory encoding (Lee et al., 2006; Reizenzein et al., 2019). Presumably, depending on how the information to be encoded relates to the information contained in the memory system, its encoding, consolidation, and even retrieval can also be affected by motivation and the novelty of the information itself (van Kesteren et al., 2012; Miendlarzewska et al., 2016).

Another interesting question pertains to the effects of stimulus types and other methodological differences which could also affect memory performance. Could these effects also be integrated into the predictive coding framework? Stimulus material effects may come to be explained through the strength of the prediction errors they elicit, as more complex stimuli may elicit greater prediction errors. Retention time and testing method effects, however, may rely on other neurocognitive



**FIGURE 2 |** Novelty manipulations and the surprise signal (novelty Surprise.jpg). The brain constantly generates prediction errors as the expectations generated by its inner model are compared to its actual inputs. These prediction errors are summed to judge the necessity of updating the inner model. If the sum of the prediction errors is high enough, it results in the surprise response—the focusing of attention and enhanced learning. Stimulus novelty manipulations usually fail to elicit this response, while it is frequently evoked by contextual, associative and spatial novelty manipulations [Modified figure; Patrick J. Lynch, medical illustrator ([https://commons.wikimedia.org/wiki/File:Brain\\_human\\_lateral\\_view.svg](https://commons.wikimedia.org/wiki/File:Brain_human_lateral_view.svg)), “Brain stem normal human”, added features by R. Reichardt, <https://creativecommons.org/licenses/by/2.5/legalcode>].



processes that occur after initial encoding, although it is admittedly hard to distinguish between these effects with classical paradigms (Cohen et al., 2014). These factors should also be systematically investigated concerning the degree of unexpectedness to draw a complete picture of the memory effects of novelty.

## AUTHOR CONTRIBUTIONS

RR, BP and PS contributed equally to this work.

## REFERENCES

- Alexander, W. H., and Brown, J. W. (2019). The role of the anterior cingulate cortex in prediction error and signaling surprise. *Top. Cogn. Sci.* 11, 119–135. doi: 10.1111/tops.12307
- Auksztulewicz, R., and Friston, K. (2016). Repetition suppression and its contextual determinants in predictive coding. *Cortex* 80, 125–140. doi: 10.1016/j.cortex.2015.11.024
- Bach, D. R., and Dolan, R. J. (2012). Knowing how much you don't know: a neural organization of uncertainty estimates. *Nat. Rev. Neurosci.* 13, 572–586. doi: 10.1038/nrn3289
- Barkaszi, I., Czigler, I., and Balázs, L. (2013). Stimulus complexity effects on the event-related potentials to task-irrelevant stimuli. *Biol. Psychol.* 94, 82–89. doi: 10.1016/j.biopsycho.2013.05.007
- Barto, A., Mirolli, M., and Baldassarre, G. (2013). Novelty or surprise? *Front. Psychol.* 4:907. doi: 10.3389/fpsyg.2013.00907
- Berlyne, D. E. (1960). *Conflict, Arousal, and Curiosity*. New York, NY: McGraw-Hill Book Company.
- Brankazk, J., Seidenbecher, T., and Müller-Gärtner, H. W. (1996). Task-relevant late positive component in rats: is it related to hippocampal theta rhythm? *Hippocampus* 6, 475–482. doi: 10.1002/(sici)1098-1063(1996)6:5<475::aid-hipo1>3.0.co;2-i
- Breton, F., Ritter, W., Simson, R., and Vaughan, H. G. Jr. (1988). The N2 component elicited by stimulus matches and multiple targets. *Biol. Psychol.* 27, 23–44. doi: 10.1016/0301-0511(88)90003-8
- Bunzeck, N., and Düzel, E. (2006). Absolute coding of stimulus novelty in the human substantia nigra/VTA. *Neuron* 51, 369–379. doi: 10.1016/j.neuron.2006.06.021
- Bunzeck, N., Schütze, H., Stallforth, S., Kaufmann, J., Düzel, S., Heinze, H. J., et al. (2007). Mesolimbic novelty processing in older adults. *Cereb. Cortex* 17, 2940–2948. doi: 10.1093/cercor/bhm020
- Chater, N. (2019). *The Mind is Flat*. London: Penguin.
- Chong, H., Riis, J. L., McGinnis, S. M., Williams, D. M., Holcomb, P. J., and Daffner, K. R. (2008). To ignore or explore: top-down modulation of novelty processing. *J. Cogn. Neurosci.* 20, 120–134. doi: 10.1162/jocn.2008.20003
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 36, 181–204. doi: 10.1017/s0140525x12000477
- Cohen, N., Pell, L., Edelson, M. G., Ben-Yakov, A., Pine, A., and Dudai, Y. (2014). Peri-encoding predictors of memory encoding and consolidation. *Neurosci. Biobehav. Rev.* 50, 128–142. doi: 10.1016/j.neubiorev.2014.11.002
- Courchesne, E., Hillyard, S. A., and Galambos, R. (1975). Stimulus novelty, task relevance and the visual evoked potential in man. *Electroencephalogr. Clin. Neurophysiol.* 39, 131–143. doi: 10.1016/0013-4694(75)90003-6
- Czigler, I., and Balázs, L. (2005). Age-related effects of novel visual stimuli in a letter-matching task: an event-related potential study. *Biol. Psychol.* 69, 229–242. doi: 10.1016/j.biopsycho.2004.06.006
- Daffner, K. R., Mesulam, M. M., Scinto, L. F. M., Calvo, V., Faust, R., and Holcomb, P. J. (2000). An electrophysiological index of stimulus unfamiliarity. *Psychophysiology* 37, 737–747. doi: 10.1111/1469-8986.3760737
- Davis, C. D., Jones, F. L., and Derrick, B. E. (2004). Novel environments enhance the induction and maintenance of long-term potentiation in the dentate gyrus. *J. Neurosci.* 24, 6497–6506. doi: 10.1523/JNEUROSCI.4970-03.2004

## FUNDING

BP was supported by the BME-Biotechnology FIKP grant of EMMI (BME FIKP-BIO), and by the National Research, Development and Innovation Office (NKFI/OTKA K 128599). PS was supported by the Hungarian Scientific Research Fund (NKFI FK 128100) of the National Research, Development and Innovation Office. This work was completed in the ELTE Institutional Excellence Program (783 3/2018/FEKUTSRAT) supported by the Hungarian Ministry of Human Capacities.

- de Lange, F. P., Heilbron, M., and Kok, P. (2018). How do expectations shape perception? *Trends Cogn. Sci.* 22, 764–779. doi: 10.1016/j.tics.2018.06.002
- den Ouden, H. E. M., Kok, P., and de Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Front. Psychol.* 3:548. doi: 10.3389/fpsyg.2012.00548
- Dobbins, I. G., Kroll, N. E. A., Yonelinas, A. P., and Liu, Q. (1998). Distinctiveness in recognition and free recall: the role of recollection in the rejection of the familiar. *J. Mem. Lang.* 38, 381–400. doi: 10.1006/jmla.1997.2554
- Düzel, E., Habib, R., Rotte, M., Guderian, S., Tulving, E., and Heinze, H.-J. (2003). Human hippocampal and parahippocampal activity during visual associative recognition memory for spatial and nonspatial stimulus configurations. *J. Neurosci.* 23, 9439–9444. doi: 10.1523/JNEUROSCI.23-28-09439.2003
- Faraji, M., Preuschoff, K., and Gerstner, W. (2018). Balancing new against old information: the role of puzzlement surprise in learning. *Neural Comput.* 30, 34–83. doi: 10.1162/neco\_a\_01025
- Fenker, D. B., Frey, J. U., Schuetz, H., Heipertz, D., Heinze, H., and Düzel, E. (2008). Novel scenes improve recollection and recall of words. *J. Cogn. Neurosci.* 20, 1250–1265. doi: 10.1162/jocn.2008.20086
- Ferrari, V., Bradley, M. M., Codispoti, M., and Lang, P. J. (2015). Massed and distributed repetition of natural scenes: brain potentials and oscillatory activity. *Psychophysiology* 52, 865–872. doi: 10.1111/psyp.12424
- Floresco, S. B., Todd, C. L., and Grace, A. A. (2001). Glutamatergic afferents from the hippocampus to the nucleus accumbens regulate activity of ventral tegmental area dopamine neurons. *J. Neurosci.* 21, 4915–4922. doi: 10.1523/JNEUROSCI.21-13-04915.2001
- Floresco, S. B., West, A. R., Ash, B., Moorel, H., and Grace, A. A. (2003). Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nat. Neurosci.* 6, 968–973. doi: 10.1038/nn1103
- Folstein, J. R., and Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology* 45, 152–170. doi: 10.1111/j.1469-8986.2007.00602.x
- Friedman, D., Cycowicz, Y. M., and Gaeta, H. (2001). The novelty P3: an event-related brain potential (ERP) sign of the brainandapos; evaluation of novelty. *Neurosci. Biobehav. Rev.* 25, 355–373. doi: 10.1016/s0149-7634(01)00019-7
- Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 815–836. doi: 10.1098/rstb.2005.1622
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787
- Gasbarri, A., Packard, M. G., Campana, E., and Pacitti, C. (1994a). Anterograde and retrograde tracing of projections from the ventral tegmental area to the hippocampal formation in the rat. *Brain Res. Bull.* 33, 445–452. doi: 10.1016/0361-9230(94)90288-7
- Gasbarri, A., Verney, C., Innocenzi, R., Campana, E., and Pacitti, C. (1994b). Mesolimbic dopaminergic neurons innervating the hippocampal formation in the rat: a combined retrograde tracing and immunohistochemical study. *Brain Res.* 668, 71–79. doi: 10.1016/0006-8993(94)90512-6
- Geraci, L., and Manzano, I. (2010). Distinctive items are salient during encoding: delayed judgements of learning predict the isolation effect. *Q. J. Exp. Psychol.* 63, 50–64. doi: 10.1080/17470210902790161
- Gotts, S. J., Chow, C. C., Martin, A., Gotts, S. J., Chow, C. C., and Martin, A. (2012). Repetition priming and repetition suppression: a case for enhanced efficiency through neural synchronization. *Cogn. Neurosci.* 8928, 227–259. doi: 10.1080/17588928.2012.670617

- Greve, A., Cooper, E., Kaula, A., Anderson, M. C., and Henson, R. (2017). Does prediction error drive one-shot declarative learning? *J. Mem. Lang.* 94, 149–165. doi: 10.1016/j.jml.2016.11.001
- Grill-Spector, K., Henson, R., and Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci.* 10, 14–23. doi: 10.1016/j.tics.2005.11.006
- Grotheer, M., and Kovács, G. (2014). Repetition probability effects depend on prior experiences. *J. Neurosci.* 34, 6640–6646. doi: 10.1523/JNEUROSCI.5326-13.2014
- Grotheer, M., and Kovács, G. (2015). The relationship between stimulus repetitions and fulfilled expectations. *Neuropsychologia* 67, 175–182. doi: 10.1016/j.neuropsychologia.2014.12.017
- Grotheer, M., and Kovács, G. (2016). Can predictive coding explain repetition suppression? *Cortex* 80, 113–124. doi: 10.1016/j.cortex.2015.11.027
- Henson, R. N., and Gagnepain, P. (2010). Predictive, interactive multiple memory systems. *Hippocampus* 20, 1315–1326. doi: 10.1002/hipo.20857
- Hobson, J. A., and Friston, K. J. (2012). Waking and dreaming consciousness: neurobiological and functional considerations. *Prog. Neurobiol.* 98, 82–98. doi: 10.1016/j.pneurobio.2012.05.003
- Hyman, J. M., Holroyd, C. B., and Seamans, J. K. (2017). A novel neural prediction error found in anterior cingulate cortex ensembles. *Neuron* 95, 447.e3–456.e3. doi: 10.1016/j.neuron.2017.06.021
- Jeewajee, A., Lever, C., Burton, S., O’Keefe, J., and Burgess, N. (2008). Environmental novelty is signaled by reduction of the hippocampal theta frequency. *Hippocampus* 18, 340–348. doi: 10.1002/hipo.20394
- Jenkins, T. A., Amin, E., Pearce, J. M., Brown, M. W., and Aggleton, J. P. (2004). Novel spatial arrangements of familiar visual stimuli promote activity in the rat hippocampal formation but not the parahippocampal cortices: a c-fos expression study. *Neuroscience* 124, 43–52. doi: 10.1016/j.neuroscience.2003.11.024
- Kafkas, A., and Montaldi, D. (2018). How do memory systems detect and respond to novelty? *Neurosci. Lett.* 680, 60–68. doi: 10.1016/j.neulet.2018.01.053
- Kaliukhovich, D. A., and Vogels, R. (2011). Stimulus repetition probability does not affect repetition suppression in macaque inferior temporal cortex. *Cereb. Cortex* 21, 1547–1558. doi: 10.1093/cercor/bhq207
- Karis, D., Fabiani, M., and Donchin, E. (1984). “P300” and memory: individual differences in the von Restorff effect. *Cogn. Psychol.* 16, 177–216. doi: 10.1016/0010-0285(84)90007-0
- Kempadoo, K. A., Mosharof, E. V., Choi, S. J., Sulzer, D., and Kandel, E. R. (2016). Dopamine release from the locus coeruleus to the dorsal hippocampus promotes spatial learning and memory. *Proc. Natl. Acad. Sci. U S A* 113, 14835–14840. doi: 10.1073/pnas.1616515114
- Kirchhoff, B. A., Wagner, A. D., Maril, A., and Stern, C. E. (2000). Prefrontal-temporal circuitry for episodic encoding and subsequent memory. *J. Neurosci.* 20, 6173–6180. doi: 10.1523/JNEUROSCI.20-16-06173.2000
- Knill, D. C., and Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 27, 712–719. doi: 10.1016/j.tics.2004.10.007
- Köhler, S., Danckert, S., Gati, J. S., and Menon, R. S. (2005). Novelty responses to relational and non-relational information in the hippocampus and the parahippocampal region: a comparison based on event-related fMRI. *Hippocampus* 15, 763–774. doi: 10.1002/hipo.20098
- Kormi-Nouri, R., Nilsson, L. G., and Ohta, N. (2005). The novelty effect: support for the novelty-encoding hypothesis. *Scand. J. Psychol.* 46, 133–143. doi: 10.1111/j.1467-9450.2005.00443.x
- Kovács, G., Kaiser, D., Kaliukhovich, D. A., Vidnyánszky, Z., and Vogels, R. (2013). Repetition probability does not affect fMRI repetition suppression for objects. *J. Neurosci.* 33, 9805–9812. doi: 10.1523/JNEUROSCI.3423-12.2013
- Krebs, R. M., Schott, B. H., and Düzel, E. (2009). Personality traits are differentially associated with patterns of reward and novelty processing in the human substantia nigra/ventral tegmental area. *Biol. Psychiatry* 65, 103–110. doi: 10.1016/j.biopsych.2008.08.019
- Kumaran, D., and Maguire, E. A. (2006). An unexpected sequence of events: mismatch detection in the human hippocampus. *PLoS Biol.* 4:e424. doi: 10.1371/journal.pbio.0040424
- Kumaran, D., and Maguire, E. A. (2007). Match mismatch processes underlie human hippocampal responses to associative novelty. *J. Neurosci.* 27, 8517–8524. doi: 10.1523/JNEUROSCI.1677-07.2007
- Lee, H. J., Youn, J. M., O, M. J., Gallagher, M., and Holland, P. C. (2006). Role of substantia nigra-amygdala connections in surprise-induced enhancement of attention. *J. Neurosci.* 26, 6077–6081. doi: 10.1523/JNEUROSCI.1316-06.2006
- Legault, M., and Wise, R. A. (2001). Novelty-evoked elevations of nucleus accumbens dopamine: dependence on impulse flow from the ventral subiculum and glutamatergic neurotransmission in the ventral tegmental area. *Eur. J. Neurosci.* 13, 819–828. doi: 10.1046/j.0953-816x.2000.01448.x
- Lisman, J. E., and Grace, A. A. (2005). The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* 46, 703–713. doi: 10.1016/j.neuron.2005.05.002
- Lisman, J., Grace, A. A., and Duzel, E. (2011). A neoHebbian framework for episodic memory: role of dopamine-dependent late LTP. *Trends Neurosci.* 34, 536–547. doi: 10.1016/j.tins.2011.07.006
- Ljungberg, T., Apicella, P., and Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol.* 67, 145–163. doi: 10.1152/jn.1992.67.1.145
- Miendlarzewska, E. A., Bavelier, D., and Schwartz, S. (2016). Influence of reward motivation on human declarative memory. *Neurosci. Biobehav. Rev.* 61, 156–176. doi: 10.1016/j.neubiorev.2015.11.015
- Murty, V. P., Ballard, I. C., MacDuffie, K. E., Krebs, R. M., and Adcock, R. A. (2013). Hippocampal networks habituate as novelty accumulates. *Learn. Mem.* 20, 229–235. doi: 10.1101/lm.029728.112
- Nieuwenhuis, S., Aston-Jones, G., and Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychol. Bull.* 131, 510–532. doi: 10.1037/0033-2909.131.4.510
- Nieuwenhuis, S., De Geus, E. J., and Aston-Jones, G. (2011). The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology* 48, 162–175. doi: 10.1111/j.1469-8986.2010.01057.x
- Nir, Y., and Tononi, G. (2010). Dreaming and the brain: from phenomenology to neurophysiology. *Trends Cogn. Sci.* 14, 88–100. doi: 10.1016/j.tics.2009.12.001
- Nyberg, L. (2005). Any novelty in hippocampal formation and memory? *Curr. Opin. Neurol.* 18, 424–428. doi: 10.1097/01.wco.0000168080.99730.1c
- O’Reilly, J. X., Schüffelgen, U., Cuell, S. F., Behrens, T. E. J., Mars, R. B., and Rushworth, M. F. S. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proc. Natl. Acad. Sci. U S A* 110, E3660–E3669. doi: 10.1073/pnas.1305373110
- Pine, A., Sadeh, N., Ben-Yakov, A., Dudai, Y., and Mendelsohn, A. (2018). Knowledge acquisition is governed by striatal prediction errors. *Nat. Commun.* 9:1673. doi: 10.1038/s41467-018-03992-5
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019
- Poppenk, J., Köhler, S., and Moscovitch, M. (2010). Revisiting the novelty effect: when familiarity, not novelty, enhances memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 1321–1330. doi: 10.1037/a0019900
- Ranganath, C., and Rainer, G. (2003). Cognitive neuroscience: neural mechanisms for detecting and remembering novel events. *Nat. Rev. Neurosci.* 4, 193–202. doi: 10.1038/nrn1052
- Reggev, N., Shari, R., and Maril, A. (2017). Distinctiveness Benefits Novelty (and Not Familiarity), but Only Up to a Limit: the Prior Knowledge Perspective. *Cogn. Sci.* 42, 103–128. doi: 10.1111/cogs.12498
- Reisenzein, R., Horstmann, G., and Schützwohl, A. (2019). The cognitive-evolutionary model of surprise: a review of the evidence. *Top. Cogn. Sci.* 11, 50–74. doi: 10.1111/tops.12292
- Ruusuvirta, T., Korhonen, T., Penttonen, M., Arikoski, J., and Kivirikko, K. (1995). Behavioral and hippocampal evoked responses in an auditory oddball situation when an unconditioned stimulus is paired with deviant tones in the cat: experiment II. *Int. J. Psychophysiol.* 20, 41–47. doi: 10.1016/0167-8760(95)00025-n
- Schomaker, J. (2019). Unexplored territory: neurobiology of learning and memory. *Neurobiol. Learn. Mem.* 161, 46–50. doi: 10.1016/j.nlm.2019.03.005
- Schomaker, J., and Meeter, M. (2015). Short- and long-lasting consequences of novelty, deviance and surprise on brain and cognition. *Neurosci. Biobehav. Rev.* 55, 268–279. doi: 10.1016/j.neubiorev.2015.05.002
- Schomaker, J., and Meeter, M. (2018). Predicting the unknown: novelty processing depends on expectations. *Brain Res.* 1694, 140–148. doi: 10.1016/j.brainres.2018.05.008

- Schomaker, J., Roos, R., and Meeter, M. (2014a). Expecting the unexpected: the effects of deviance on novelty processing. *Behav. Neurosci.* 128, 146–160. doi: 10.1037/a0036289
- Schomaker, J., van Bronkhorst, M. L. and Meeter, M. (2014b). Exploring a novel environment improves motivation and promotes recall of words. *Front. Psychol.* 5:918. doi: 10.3389/fpsyg.2014.00918
- Schott, B. H., Sellner, D. B., Lauer, C. J., Habib, R., Frey, J. U., Guderian, S., et al. (2004). Activation of midbrain structures by associative novelty and the formation of explicit memory in humans. *Learn. Mem.* 11, 383–387. doi: 10.1101/lm.75004
- Schott, B. H., Seidenbecher, C. I., Fenker, D. B., Lauer, C. J., Bunzeck, N., Bernstein, H. G., et al. (2006). The dopaminergic midbrain participates in human episodic memory formation: evidence from genetic imaging. *J. Neurosci.* 26, 1407–1417. doi: 10.1523/JNEUROSCI.3463-05.2006
- Schultz, W. (2016). Dopamine reward prediction-error signalling: a two-component response. *Nat. Neurosci.* 17, 183–195. doi: 10.1038/nrn.2015.26
- Seer, C., Lange, F., Boos, M., Dengler, R., and Kopp, B. (2016). Prior probabilities modulate cortical surprise responses: a study of event-related potentials. *Brain Cogn.* 106, 78–89. doi: 10.1016/j.bandc.2016.04.011
- Shohamy, D., and Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends Cogn. Sci.* 14, 464–472. doi: 10.1016/j.tics.2010.08.002
- Sokolov, E. N. (1963). Higher nervous functions: the orienting reflex. *Annu. Rev. Physiol.* 25, 545–580. doi: 10.1146/annurev.ph.25.030163.002553
- Stefanics, G., Kremláček, J., and Czigler, I. (2014). Visual mismatch negativity: a predictive coding view. *Front. Hum. Neurosci.* 8:666. doi: 10.3389/fnhum.2014.00666
- Stoppel, C. M., Boehler, C. N., Strumpf, H., Heinze, H. J., Hopf, J. M., Düzel, E., et al. (2009). Neural correlates of exemplar novelty processing under different spatial attention conditions. *Hum. Brain Mapp.* 30, 3759–3771. doi: 10.1002/hbm.20804
- Straube, T., Korz, V., and Frey, J. U. (2003). Bidirectional modulation of long-term potentiation by novelty-exploration in rat dentate gyrus. *Neurosci. Lett.* 344, 5–8. doi: 10.1016/s0304-3940(03)00349-5
- Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M. M., and Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nat. Neurosci.* 11, 1004–1006. doi: 10.1038/nn.2163
- Tarbi, E. C., Sun, X., Holcomb, P. J., and Daffner, K. R. (2011). Surprise? Early visual novelty processing is not modulated by attention. *Psychophysiology* 48, 624–632. doi: 10.1111/j.1469-8986.2010.01129.x
- Thomas, K. M., and Nelson, C. A. (1996). Age-related changes in the electrophysiological response to visual stimulus novelty: a topographical approach. *Electroencephalogr. Clin. Neurophysiol.* 98, 294–308. doi: 10.1016/0013-4694(95)00280-4
- Todorovic, A., and de Lange, F. P. (2012). Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *J. Neurosci.* 32, 13389–13395. doi: 10.1523/JNEUROSCI.2227-12.2012
- Tulving, E., and Kroll, N. (1995). Novelty assessment in the brain and long-term memory encoding. *Psychon. Bull. Rev.* 2, 387–390. doi: 10.3758/bf03210977
- Tulving, E., Markowitsch, H. J., Craik, F. I. M., Habib, R., and Houle, S. (1996). Novelty and familiarity activations in PET studies of memory encoding and retrieval. *Cereb. Cortex* 6, 71–79. doi: 10.1093/cercor/6.1.71
- Tulving, E., Markowitsch, H. J., Kapur, S., Habib, R., and Houle, S. (1994). Novelty encoding networks in the human brain. *Neuroreport* 5, 2525–2528. doi: 10.1097/00001756-199412000-00030
- Uzakov, S., Frey, J. U., and Korz, V. (2005). Reinforcement of rat hippocampal LTP by holeboard training. *Learn. Mem.* 12, 165–171. doi: 10.1101/lm.89305
- van Kesteren, M. T. R., Ruiters, D. J., Fernández, G., and Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends Neurosci.* 35, 211–219. doi: 10.1016/j.tins.2012.02.001
- von Restorff, H. (1933). Über die wirkung von bereichsbildungen im spurenfeld. *Psychol. Forsch.* 18, 299–342. doi: 10.1007/bf02409636
- Wagatsuma, A., Okuyama, T., Sun, C., Smith, L. M., Abe, K., and Tonegawa, S. (2018). Locus coeruleus input to hippocampal CA3 drives single-trial learning of a novel context. *Proc. Natl. Acad. Sci. U S A* 115, E310–E316. doi: 10.1073/pnas.1714082115
- Wittmann, B. C., Bunzeck, N., Dolan, R. J., and Düzel, E. (2007). Anticipation of novelty recruits reward system and hippocampus while promoting recollection. *NeuroImage* 38, 194–202. doi: 10.1016/j.neuroimage.2007.06.038
- Wittmann, B. C., Daw, N. D., Seymour, B., and Dolan, R. J. (2008). Striatal activity underlies novelty-based choice in humans. *Neuron* 58, 967–973. doi: 10.1016/j.neuron.2008.04.027

**Conflict of Interest:** 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.

Copyright © 2020 Reichardt, Polner and Simor. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.