Evaluating the Evidence for Expectation Suppression in the Visual System

Daniel Feuerriegel^a, Rufin Vogels^b, Gyula Kovács^c

a. Melbourne School of Psychological Sciences, The University of Melbourne,
Melbourne, Australia

b. Laboratorium voor Neuro- en Psychofysiologie, Department of Neurosciences, KU
Leuven, Leuven, Belgium

c. Institute of Psychology, Friedrich Schiller University Jena, Jena, Germany

Corresponding Author: Daniel Feuerriegel

Email: dfeuerriegel@unimelb.edu.au

Postal Address: Daniel Feuerriegel, Melbourne School of Psychological Sciences,
Redmond Barry Building, The University of Melbourne, Parkville, Victoria 3010,
Australia

Abstract

Reports of expectation suppression have shaped the development of influential predictive coding-based theories of visual perception. However recent work has highlighted confounding factors that may mimic or inflate expectation suppression effects. In this review, we describe four confounds that are prevalent across experiments that tested for expectation suppression: effects of surprise, attention, stimulus repetition and adaptation, and stimulus novelty. With these confounds in mind we then critically review the evidence for expectation suppression across probabilistic cueing, statistical learning, oddball, action-outcome learning and apparent motion designs. We found evidence for expectation suppression within a specific subset of statistical learning designs that involved weeks of sequence learning prior to neural activity measurement. Across other experimental contexts, whereby stimulus appearance probabilities were learned within one or two testing sessions, there was inconsistent evidence for genuine expectation suppression. We discuss how an absence of expectation suppression could inform models of predictive processing, repetition suppression and perceptual decision-making. We also provide suggestions for designing experiments that may better test for expectation suppression in future work.

Keywords: Expectation Suppression, Repetition Suppression, Attention, Prediction, Vision, fMRI, electrophysiology

1. Introduction

Over the last two decades, several innovative theories of visual perception have been developed based on the notion of predictive coding (e.g., Rao & Ballard, 1999; Friston, 2005; Summerfield & de Lange, 2014). These theories are based on assumptions that sensory systems are continuously attempting to infer the underlying causes of incoming sensory signals, and that discrepancies between an organism's internal models and reality are reflected in patterns of neural activity. The central tenets of predictive coding-based theories are yet to be verified empirically, and many observed phenomena that are cited in support of predictive coding can also be explained using other theoretical frameworks (Heeger, 2017; Walsh et al. 2020; Cao, 2020). However, there is one phenomenon that is frequently presented as evidence for predictive coding over other accounts of perception, which is termed expectation suppression (Summerfield et al., 2008; Todorovic & de Lange 2012; Richter & de Lange, 2019; Walsh et al., 2020).

Expectation suppression (ES) is defined here as a reduction in a measure of neural activity (e.g., single-cell or multiunit firing rates, local field potential (LFP) amplitudes or bandlimited LFP power, MEG/EEG component amplitudes or fMRI BOLD signals) following the presentation of a stimulus that is subjectively expected to appear (i.e. a predicted stimulus), compared to when the same stimulus is presented but is neither expected nor surprising (i.e. a neutral comparison condition; see Egner et al., 2010; Arnal & Giraud, 2012). This term was coined in reference to the concept of repetition suppression (RS), which describes reductions in measures of neural activity with stimulus repetition (Desimone, 1996). Importantly, the notion of 'suppression' as used within many predictive coding frameworks (e.g., Egner et al., 2010; Summerfield & de Lange, 2014) implies that activity of stimulus-selective neurons is suppressed or inhibited (e.g., via input from other neurons) when an observer's expectations are fulfilled, compared to when there are no biased expectations in favour of seeing a specific stimulus. This is congruent with the more general hypothesis that neural response magnitudes evoked by a stimulus will scale inversely with the degree of expectation in favour of that stimulus (discussed in Walsh et al., 2020).

As is evident in the definition above, the concept of ES is operationalised as a difference measure between an expected stimulus condition and a neutral comparison condition. In line with previous studies (e.g., Egner et al., 2010; Amado et al., 2016; Kaposvari et al., 2018), we define expected stimulus conditions as instances whereby a single stimulus is subjectively or objectively more likely to appear than any other stimulus that might be presented, and where that expected stimulus is presented to an observer. We define neutral stimulus conditions as instances whereby two or more different stimuli each appear with approximately equal probability (i.e. there are no strong or systematic expectancy biases that favour one stimulus over another). We additionally define surprising (i.e. mispredicted) stimulus conditions as instances where an observer expects to see one specific stimulus with high probability, but this expectation is violated and a different stimulus appears instead. Importantly, the neutral stimulus condition is defined here in *relation* to the expected and surprising stimulus conditions (i.e. neutral is defined as neither expected nor surprising, e.g., instances where there are no strong expectations to see one particular stimulus). In addition, we define expected, neutral and surprising stimulus conditions as *qualitatively* different, in line with how these conditions have been operationalised in the empirical literature reviewed below. These definitions are also constrained to tightly-controlled experimental contexts. Please note that, in such ‘neutral’ conditions, many aspects of the experiment (e.g., the types of stimuli, trial structure, stimulus timing) can be predicted with much higher certainty than in naturalistic environments that are common in everyday life. We also note that the term ‘neutral’ does not necessarily imply that expectations are perfectly balanced across the range of possible stimuli in every experimental trial. For example, subtle biases in expectations based on stimulus exposure history (commonly known as sequential effects) are pervasive across many experiments (e.g., Falmagne, 1965; Luce, 1986; Bode et al., 2012). Neutral conditions constitute sets of trials that should not induce expectations to see a particular stimulus with high (e.g., 75%) probability.

Ideally, the expected and neutral conditions should be designed so that the factor of stimulus expectation is the only difference that may plausibly lead to differences in neural response measures across these two conditions. Any co-occurring

differences with respect to other factors may potentially inflate or mimic effects of ES. In fact, this is also a central concern in relation to how other response modulations, such as RS, are measured (reviewed in Feuerriegel, 2016).

Recently, the designs of many past experiments that tested for ES have been called into question. In the vast majority of experiments reporting ES, the expected stimulus conditions were not compared against a neutral condition. Instead, they were compared against a surprising stimulus condition (e.g., Summerfield et al., 2008; Todorovic & de Lange, 2012; Grotheer & Kovács, 2015; for a review see Kovacs and Vogels, 2014). This is important because there is evidence of neural response modulations that are specifically associated with expectation *violations* (i.e. surprise-related effects) which are dissociable from the [expected – neutral] condition differences that comprise ES (e.g., Amado et al., 2016; Kaposvari et al., 2018; reviewed in section 2.1). By contrast, several theoretical frameworks explicitly describe ES as reflecting a suppression of responses arising from fulfilled expectations, rather than effects that are specifically related to expectation violations (e.g., Summerfield & Egner, 2009; Summerfield & de Lange 2014; Grotheer & Kovács, 2016; de Lange et al., 2018; Walsh et al., 2020). It is therefore critical to determine the extent of evidence for ES that is not accounted for by other conceptually distinct effects.

In this review we first describe the confounding effects of surprise-related neural response modulations. We also describe three other frequently occurring confounding factors that are important to consider when evaluating experiments that have tested for ES. We then critically review the evidence for ES across five experimental paradigms in which the subjective or objective probability of a specific stimulus is manipulated. These include probabilistic cueing, statistical learning, oddball, action-outcome learning and apparent motion designs. We have included studies in humans and animal models that have recorded fMRI BOLD signals, EEG/MEG signals, LFPs and single and multiunit firing rates. We then summarise the extent of evidence for ES across these paradigms and explore the inferences that are warranted by this body of evidence. We conclude by providing suggestions for designing experiments that may be more appropriate for measuring ES in the visual system. This review focuses specifically on how expected and neutral conditions are

designed, and complements the broader theoretical discussion regarding how to interpret ES effects (reviewed in Walsh et al., 2020).

2. Confounding Factors in Experiments Testing for Expectation Suppression

2.1. Surprise-Related Response Modulations

When an observer expects to see one stimulus, and a different one appears instead, the neural responses evoked by that surprising stimulus are often enhanced compared to stimuli that do not violate an observer's expectations (e.g., Kimura et al., 2009; Egner et al., 2010; Amado et al., 2016; Richter & de Lange., 2019; Feuerriegel et al., 2021). In the majority of studies testing for ES, researchers have implicitly or explicitly treated suppression due to fulfilled expectations and surprise-related response enhancements as two sides of the same coin. However, ES and surprise-related responses are treated as distinct effects within a number of theoretical frameworks (e.g., Kovács & Vogels, 2014; Hsu et al., 2015, 2018; Schomaker & Meeter, 2015; Grotheer & Kovács, 2016; Press et al., 2020).

In fact, ES and surprise-related effects can also be empirically separated. A number of experiments have presented expected, neutral and surprising stimuli in order to separately test for ES and surprise effects. In these studies, surprise-related effects were operationalised as [surprising – neutral] condition differences. ES effects were operationalised as [expected – neutral] condition differences. Amado et al. (2016) presented pairs of two consecutive face stimuli separated by an inter-stimulus interval. The category of the faces presented in the trial (adult male, adult female, baby) cued the probability of stimulus repetition within that trial. Depending on the trial type, a face repetition or alternation could either be expected (75% probability), neutral (50% probability) or surprising (25% probability). For alternating face pairs BOLD signals in the Fusiform Face Area (FFA; Kanwisher et al., 1997) were larger following surprising compared to neutral faces. However, BOLD signals were also slightly larger for expected compared to neutral faces (i.e. the opposite of ES). This pattern appears to be visible across other studies (e.g., Larsson & Smith, 2012; Grotheer & Kovács, 2014; Mayrhauser et al., 2014; reviewed in Kovács & Vogels, 2014).

In addition, a number of experiments have recorded EEG and presented visual oddball sequences, in which a commonly-occurring and expected (standard) stimuli were interspersed with a smaller number of rarely-occurring and surprising (deviant) stimuli. They additionally presented sequences in which a large number of stimuli (e.g., 10 different images) could each appear with equal probability, which served as a neutral stimulus condition. They reported [deviant – equiprobable] condition differences (i.e. surprise-related effects) on stimulus-evoked event-related potentials (ERPs). Importantly, these effects were visible over later time windows than the observed [standard - equiprobable] ERP differences (e.g., Kimura et al., 2009, 2010; Amado & Kovács, 2016; File et al., 2017, see also Feuerriegel et al., 2021).

Similar dissociations have also been found when recording multiunit firing rates in macaque inferior temporal (IT) cortex. In a statistical learning design, Kaposvari et al. (2018) presented macaques with sequences of five images that always appeared in the same order, whereby each image could be expected based on observation of the preceding image (structured sequences). They also presented sequences of different images that were presented in a random order (random sequences). After multiple weeks of exposure to these sequences the macaques were then presented with the structured and random sequences, as well as an occasional sequence whereby one of the stimuli in the structured sequence was replaced with a different stimulus (i.e. a surprising deviant). Firing rates in macaque IT neurons following predictable stimuli in the structured sequences were reduced compared to those in random sequences during a brief time window of around 100-140 ms following stimulus onset (i.e. an ES effect). However, enhanced responses for surprising deviants compared to stimuli in random sequences (i.e. surprise-related effects) emerged slightly later and spanned approximately 120-380 ms post stimulus onset. Using a slightly different statistical learning design, Ramachandran et al. (2016) observed ES in macaque IT neurons without concurrent surprise-related effects.

Taken together, these results indicate that surprise-related effects can be dissociated from ES in their effects on haemodynamic and electrophysiological measures. Surprise is also associated with a broad range of effects that extend beyond the visual system (reviewed in Schomaker & Meeter, 2015; Wessel, 2017; Wessel &

Aron, 2017). These findings indicate that ES should be measured by comparing expected and neutral stimulus conditions, rather than by comparing expected and surprising conditions. Conflating ES and surprise responses does not allow for inferences regarding whether genuine suppression has occurred, or whether there was simply an absence of a surprise-related neural responses in the expected stimulus condition.

2.2. Effects of Attention

Attending to a stimulus has been robustly associated with increased responses of stimulus-selective neurons (Corbetta et al., 1990; Treue, 2003; Reynolds & Heeger, 2009; Serences & Kastner, 2014; Maunsell, 2015). Effects of attention are commonly conflated with effects of stimulus expectations across a wide range of tasks (discussed in Summerfield & Egner, 2016; Rungratsameetaweemana & Serences, 2019). Following the conventions of previous work (Summerfield & de Lange, 2014; Summerfield & Egner, 2016; Rungratsameetaweemana & Serences, 2019) we define attention as a neural response modulation based on the relevance or salience of a stimulus in relation to a task or goal. We define stimulus expectation as based on the subjective probability that a stimulus will appear.

Effects of attention may confound measures of ES if attention is rapidly directed to a stimulus that violates an observer's expectations, which, in turn, enhances stimulus-selective neural responses (Kok et al., 2016; Schomaker & Meeter, 2015; Press et al., 2020). In support of this notion, Richter et al. (2019) reported [surprising – expected] BOLD signal differences when participants attended to the critical stimuli, but not when participants were required to ignore these stimuli and concurrently perform a difficult task (for similar findings see Larsson & Smith, 2012). Visual mismatch responses (defined as [deviant – standard] or [deviant – equiprobable] stimulus differences) are also attenuated when spatial attention is directed away from the critical stimuli (File & Czigler, 2019; Smout et al., 2019, 2020). These findings indicate that some surprise-related effects could be understood as a type of orienting response, whereby sufficiently unexpected stimuli are regarded as highly salient and automatically attract attention (Itti & Baldi, 2009; Alink & Blank,

2021), however this remains a contentious issue (see Summerfield & de Lange, 2014; Rungratsameetaweemana & Serences, 2019). The capture of attention following an expectation violation can also be conceptualised as an action that is taken to continuously sample and better represent an unexpected sensory event, in order to minimise uncertainty associated with the corresponding perceptual inference (for instances of similar actions see Corcoran et al., 2020). This idea is supported by observations of increased discriminability between stimulus identities (i.e. higher decoding accuracy) for surprising compared to expected stimuli (e.g., Richter et al., 2018; Tang et al., 2018). Although many experiments have manipulated expectations relating to stimulus features that are not directly task-relevant (e.g., Summerfield et al., 2008; Amado et al., 2016; Feuerriegel et al., 2018a) it is likely that surprising events will nevertheless capture attention, particularly in situations where no other stimuli are concurrently presented on a display (discussed in Alink & Blank, 2021).

It is difficult to dissociate effects of expectation and attention (discussed in Summerfield & de Lange, 2014; Rungratsameetaweemana & Serences, 2019), and some predictive coding-based accounts propose that ES should indeed be most pronounced for attended stimuli (e.g., Auksztulewicz & Friston, 2016). These competing explanations can potentially be distinguished based on the timing of observed effects with respect to the encoding of expectation-relevant stimulus features. As specified by Alink and Blank (2021), attention-based accounts of ES effects would predict that [surprising – expected] differences (associated with an orienting response) should occur earlier in the time course of the stimulus-evoked response compared to effects of directing attention toward a task-relevant stimulus (e.g., Busse et al. 2008). By contrast, prediction-based accounts would predict that expectation effects should occur later than task-related attention effects. There is some evidence that task-related attention has an earlier influence than expectation-related stimulus-evoked response modulations (reviewed in Rungratsameetaweemana & Serences, 2019), however more systematic comparisons using time-resolved recording techniques are needed to resolve this issue.

2.3. Repetition Suppression and Adaptation

Repeated presentation of a stimulus leads to reductions in neural response measures as observed using a variety of electrophysiological and neuroimaging techniques (reviewed in Grill-Spector et al., 2006). This effect is commonly known as repetition suppression (RS) or stimulus-specific adaptation (Desimone, 1996; Ringo, 1996, Kohn, 2007). RS is one phenomenon within a broader family of neural adaptation effects. Recently developed circuit-based models of neural adaptation (Dhruv et al., 2011; Solomon & Kohn, 2014; Vogels, 2016; Whitmire & Stanley, 2016) also describe an enhancement of responses to stimuli that differ from the most recently seen image, supposed to be enacted by disinhibition of stimulus-selective excitatory neurons (e.g., Kaliukhovich & Vogels, 2016). There are also effects of adaptation that are dependent on the recent stimulation history extending beyond the last encountered stimulus (Baylis & Rolls, 1987; Miller et al., 1993; Ulanovsky et al., 2004; De Baene & Vogels, 2010; Vinken & Vogels, 2017a) and long-lag repetition effects that occur when the first and second presentation of a stimulus is separated by tens or even hundreds of intervening images (e.g., Xiang & Brown, 1998; Henson, 2016; Meyer & Rust, 2018; Rangarajan et al., 2020). Adaptation-related effects can mimic ES when the expected stimuli are a repetition of the last-seen stimulus, or have been encountered more recently than stimuli in neutral expectation conditions.

Some researchers have claimed that RS is a product of expectations that are persistently weighted toward the most recently-seen stimuli (e.g., Summerfield & de Lange, 2014; Aukstulewicz & Friston, 2016). An often-cited study in support of this position is Summerfield et al. (2008), who presented trials in which two faces appeared in succession, separated by a brief inter-stimulus interval. The second face could either be a repetition of the first face (in repetition trials) or a novel face identity (in alternation trials). They reported larger [alternating – repeated] BOLD signal differences (i.e. larger repetition effects) in the FFA in blocks of trials whereby stimulus repetitions were more frequent (and thereby expected), compared to when they were less frequent and relatively unexpected. However, the repetition-by-block-type analysis design used in their experiment and in subsequent replications (e.g., Grotheer & Kovács, 2014; Kronbichler et al., 2018) was not appropriate for detecting a modulation of RS by expectations. In blocks with larger proportions of repetition trials

the repeated face images were expected and the alternating faces were surprising. In blocks with smaller numbers of repetition trials the repeated faces were instead surprising and the alternating faces were expected. Consequently, the differences in the magnitude of [alternating – repeated] signal differences could have been inflated by separate but co-occurring effects of stimulus expectations in the high repetition probability blocks, and conversely would be counteracted by surprise-related effects in the low repetition probability blocks (discussed in Grotheer & Kovacs, 2015; Amado et al., 2016; Rostalski et al., 2020). A subsequent experiment independently manipulated stimulus repetition and expectation (Grotheer & Kovács, 2015) and identified additive rather than interactive effects of RS and expectations within the same regions of interest (ROIs). RS also appears to be reliably observed in situations where repetition probability effects on BOLD signals have not been found, for example when attention is directed away from the critical stimuli (Larsson & Smith, 2012), and when presenting certain categories of stimuli, such as certain non-face objects and false fonts (e.g., Kovács et al., 2013; Grotheer & Kovács, 2014). Other experiments have also dissociated effects of repetition and stimulus expectations using measures with higher temporal resolution, such as MEG/EEG (Todorovic & de Lange, 2012; Feuerriegel et al., 2018a; Tang et al., 2018) and single-unit and multiunit firing rates and LFPs (Kaliukhovich & Vogels, 2011, 2014; Vinken et al., 2018).

Based on this evidence, ES and RS may plausibly be enacted by different underlying mechanisms (reviewed in Grotheer & Kovács, 2016; Vogels, 2016; Walsh et al., 2020). However, it cannot yet be excluded that RS and ES are related in the sense that RS reflects suppression of responses to an expected stimulus, where there is a strong, persistent and ‘automatic’ prior of stimulus repetition. This could theoretically be implemented by recurrent activity within or between areas, but appears to be unrelated to expectations as manipulated in probabilistic cueing designs. Accordingly, any clear evidence for ES should be observed in experimental designs that do not have stimulus repetition confounds.

2.4. Stimulus Novelty and Familiarity

Novel stimuli typically evoke larger measures of neural activity compared to familiar ones, as observed when recording single unit and multiunit firing rates (Li et al., 1993; Ringo, 1996; Xiang & Brown, 1998; Freedman et al., 2006; Meyer & Rust, 2018), EEG and MEG signal amplitudes (Henson et al., 2004; Mahanova et al., 2018), and BOLD signals (Mur et al., 2010; Turk-Browne et al., 2012). Extensive familiarity with a stimulus via repeated exposure over days, weeks or months may also lead to sharpened tuning of stimulus-selective neurons in IT cortex (Freedman et al., 2006; Woloszyn & Sheinberg, 2012). This sharpening effect appears to rely on different mechanisms as compared to RS as measured in short-term adaptation designs, since no sharpening of tuning for repeated stimuli occurs in the latter designs (e.g., McMahon & Olson, 2007; De Baene & Vogels, 2010). The presentation of a novel stimulus also leads to effects extending beyond the visual system, which are conceptually similar to those associated with surprising events (reviewed in Schomaker & Meeter, 2015; Wessel, 2017). Relative stimulus novelty can be an important confound in situations where expected stimulus identities are presented often within an experiment, but neutral or surprising stimuli are presented only rarely.

Here, we note that novelty and familiarity can be distinguished from the long-lag repetition effects mentioned above. Novelty and familiarity pertain to the amount of exposure or experience with a particular stimulus (often operationalised as the number of times the stimulus has been presented to an observer). Long-lag repetition effects instead relate to the recency with which a stimulus has been presented (see Xiang & Brown, 1998). For example, if meeting an old friend that one has not seen in years, their face may be highly familiar, but not recently encountered.

3. Reviewing the Evidence for Expectation Suppression Across Paradigms

Below we critically review the evidence for ES across several experimental paradigms in which ES is theorised to occur. These include probabilistic cueing, statistical learning, oddball, roving stimulus, action-outcome learning and apparent motion designs. We have included findings from experiments in humans and non-human primates which have compared neural responses following expected visual stimuli with those following neutral stimuli. We have focused on experiments that i) are designed to manipulate participants' expectations of a particular stimulus *before* the appearance of that stimulus, and ii) do not have apparent stimulus repetition, novelty, or surprise effect confounds as described above, except where acknowledgement of these confounds is useful for interpreting the literature.

3.1. Probabilistic Cueing Designs

In probabilistic cueing designs a cue stimulus is presented in each trial before the critical stimulus (depicted in Figure 1A). Different cues signal whether a particular stimulus is more likely to appear. The subsequently-presented stimulus will either conform to, or violate, the cued expectation. In neutral stimulus trials, the cue signifies that a number of different stimuli could appear with equal probability. Here we have included both explicit variants of these designs (in which participants are notified of the cue-stimulus contingencies) and implicit variants (in which participants learn these contingencies via exposure, and may not necessarily become aware of them).

The vast majority of probabilistic cueing experiments have presented only expected and surprising stimulus conditions (e.g., Kok et al., 2012a; Richter & de Lange, 2019). These experiments will not be reviewed in detail here, as it is unclear whether the observed [surprising – expected] differences are due to ES, surprise-related responses, or a mixture of both effects. Of the small number of experiments that included neutral conditions, only a minority reported statistically significant effects of ES, and a larger number of studies have produced data that are more consistent with a null effect. In addition, the [surprising – expected] condition

differences observed across fMRI experiments do not appear to replicate when analysing electrophysiological measures of neural activity.

A number of fMRI experiments entrained participants' expectations to see specific stimulus images and did not find consistent evidence for ES. Summerfield and Koechlin (2006) cued expectations of a Gabor patch with a particular orientation (expected stimuli), and also included a condition in which one of two orientations could each appear with 50% probability (neutral stimuli). Group-averaged BOLD signals were very similar across expected and neutral conditions within several ROIs located in the visual system, and were not statistically significantly different at any ROI. St John-Saaltink et al. (2015) similarly cued expectations to see either left- or right-tilted grating stimuli, and also included a 50% probability neutral condition. They found evidence for ES within areas V1, V2 and V3, but only when participants were performing an orthogonal 1-back task on concurrently-presented letter stimuli. ES was not found during a working memory-demanding 2-back task, or in another task whereby participants judged the spatial frequency of the gratings and thus attended to the critical stimuli. In the latter two tasks, BOLD signals were very similar across expected and neutral stimulus conditions across all ROIs. A lack of ES was also observed in area MT+ when entraining participants' expectations to see either expanding or contracting motion stimuli (Rahnev et al., 2011).

Other fMRI studies have entrained expectations for stimulus categories rather than specific exemplars. These have yielded tentative but inconsistent evidence for ES. Egner et al. (2010) entrained participants to expect either a face or a house stimulus, or either category with 50% probability (neutral condition). They reported a qualitative pattern of BOLD signals in the FFA that followed a surprise > neutral > expected pattern for faces, however the contrast between expected and neutral conditions was not statistically tested, and their small sample (n=16) may have prevented them from finding statistically significant ES. A similar pattern was evident for houses within the Parahippocampal Place Area (PPA), although this contrast was also not explicitly tested. Dunovan & Wheeler (2018) used a similar face/house category cueing design. BOLD signal magnitudes were very similar across expected and neutral stimulus conditions, and the difference was not statistically tested. Davis and Hasson (2018)

presented sequences in which successive stimulus categories (faces, houses, tools and flowers) and locations were either predictable (with 75% validity) or not predictable. They reported reduced averaged BOLD signals for sequences with predictable categories or locations across several areas of visual cortex.

Figure 1. Examples of probabilistic cueing and statistical learning designs. A) In probabilistic cueing designs (e.g., Egner et al., 2010), a cue stimulus signifies that a specific stimulus is more likely to appear (diamond cue), or that two or more stimuli are equally likely to appear (circle cue). B) In statistical learning designs (e.g., Mahanova et al., 2018), participants are presented with structured sequences, whereby each image predicts the image that will appear next, and random sequences, whereby each image can follow each other image with equal probability. Purple outlines denote expected stimuli, orange outlines denote surprising stimuli, and green outlines denote neutral stimuli. Example stimulus images are taken from Rossion and Pourtois (2004).

A distinct set of fMRI studies have investigated the relationship between expectation and repetition effects. In these experiments, pairs of repeated or alternating images were presented in each trial, separated by a brief inter-stimulus interval. The probability of stimulus repetition within a trial was manipulated in a block-wise fashion (e.g., Summerfield et al., 2008; Grotheer & Kovács, 2014). Most experiments only presented expected and surprising stimulus conditions (e.g., Summerfield et al., 2008; Grotheer & Kovács, 2014, 2015; Ewbank et al., 2016; Uzterath

et al., 2017). Amado et al. (2016) also included a neutral condition and reported larger BOLD signals in the FFA for expected compared to neutral alternating stimuli (i.e. the opposite of ES), and highly similar BOLD signals across expected and neutral repeated stimuli. They also found substantially larger BOLD signals evoked by surprising compared to neutral stimuli, indicating that the [surprising – expected] condition differences observed in previous work may be due to surprise-related effects rather than ES. Pajani et al. (2017) similarly presented pairs of repeated or alternating stimuli and manipulated participants' ability to predict the identity of the alternating stimulus. They reported suppressed BOLD signals in the FFA in trials whereby a specific unrepeated face image could be expected with 75% probability, compared to when the face image was randomly selected from a large stimulus set. However, there was a relative novelty confound in their design; the predictable faces were presented frequently throughout the experiment, whereas each unpredictable face image was presented only once.

Notably, the [surprise – expected] condition differences observed in fMRI experiments do not appear to replicate when using electrophysiological measures. Neither ES nor surprise-related effects were observed across multiple experiments that recorded single-unit and multiunit firing rates and LFPs in macaque IT neurons, despite robust findings of RS (Kaliukhovich & Vogels, 2011, 2014; Vinken et al., 2018). In human electrophysiological studies, effects of stimulus expectations have not exhibited clear sources in visual cortex (Summerfield et al., 2011; Feuerriegel et al., 2018a; Kok et al., 2017; Hall et al., 2018; Rungratsameetaweemana et al., 2018), despite reports of expectation-related BOLD signal modulations in very similar experiments (e.g., Kok et al., 2012a; Richter & de Lange, 2019). Tang et al. (2018) did observe smaller visual N₁ ERP component peak amplitudes evoked by expected compared to surprising Gabor patch orientations. This peak amplitude difference may be due to the fact that there were four times as many trials in the expected compared to surprising stimulus conditions, and peak amplitudes tend to be artificially inflated in conditions with smaller numbers of epochs (e.g., Thomas et al., 2004). Of these studies, only Feuerriegel et al. (2018a) included a neutral stimulus condition, however there was a

block order confound that affected the stimulus-evoked ERPs, making results from this study difficult to interpret.

One notable exception to this pattern of null effects is Bell et al. (2016), who presented face and fruit stimuli to macaques and varied the presentation probability of each stimulus category across consecutive sets of trials. They used a hierarchical Bayesian modelling approach to estimate the subjective appearance probabilities of face and fruit stimuli in each trial, resulting in expected ($>66\%$ probability), neutral ($>33\%$ and $<66\%$ probability) and surprising ($<33\%$ probability) conditions. They reported lower population firing rates in macaque IT neurons for expected compared to neutral faces between approximately 150-300 ms post stimulus onset, along with a large and sustained difference between surprising and neutral conditions between approximately 150-500 ms. These differences were not found for fruit stimuli, or for either stimulus category when they were partially occluded by visual noise.

There are two potential confounds which may have produced the ES observed in their experiment. First of all, the neutral condition included situations in which the estimated subjective probability of stimulus appearance was less than 50%, meaning that a small subset of surprise-related responses may also have contributed to population firing rate increases in the neutral condition. In addition, their design had a RS-related confound; trials with high model-derived subjective probabilities were more likely to be preceded by the same stimulus category in recently-completed trials. They attempted to account for this using a regression model, however this model did not properly account for effects of recent stimulation history (as demonstrated by Vinken & Vogels, 2017b). This may be why effects of stimulus expectation were observed in their design, but not in very similar experiments that controlled for repetition effects (Kaliukhovich & Vogels, 2011, 2014; Vinken et al., 2018).

3.2. Statistical Learning Designs

In statistical learning designs (Figure 1B), sequences of two or more stimuli are presented to an observer, who learns the transition probabilities between successive stimuli via repeated exposure (e.g., Fiser & Aslin, 2002; Turk-Browne et al., 2005;

Sherman et al., 2020). For example, within a training session stimulus B may follow stimulus A with high probability within a sequence (structured sequences in Figure 1B). An observer consequently learns to associate the two stimuli and expects stimulus B to appear after seeing stimulus A. In subsequent testing sessions, different instances of stimuli are presented which either conform to, or violate, these expectations. Some variants of these designs also include neutral conditions, in which the presentation of a certain stimulus signifies that two or more images could each appear with equal probability (e.g., Richter et al., 2018). Other variants also include neutral conditions comprising sequences of images that are presented in a random order, so that image-specific expectations cannot be formed (the random sequences in Figure 1B). These randomised, neutral stimulus sequences are typically also presented during training sessions to equate familiarity across structured and random sequences.

In contrast to findings from probabilistic cueing designs, a number of studies using statistical learning designs have reported effects of learned expectations on the activity of macaque IT neurons. All of these experiments involved extensive training over several weeks prior to the testing sessions. Multiple studies reported suppressed firing rates for the previously associated (and thereby expected) stimuli compared to surprising stimuli (Meyer & Olson, 2011; Meyer et al., 2014; Schwiedrzik & Freiwald, 2017), however these studies did not include a neutral stimulus condition. Ramachandran et al. (2017) included additional neutral, pseudorandom stimulus sequences, in which any of six different stimuli could follow a particular stimulus image with equal probability. They reported suppressed firing rates for expected compared to neutral stimuli. Interestingly, they did not report response increases compared to the neutral condition when the macaques' expectations were occasionally violated in the structured sequences (i.e., there was little evidence of surprise-related effects). Another study by Ramachandran et al. (2016) included a slightly different neutral stimulus condition, in which one of two different stimuli could each follow a given stimulus image with 50% probability. In both sequence types they occasionally presented surprising stimuli which did not conform to the learned transition probabilities. They did not directly compare firing rates following expected and neutral stimuli, however the plotted time-courses of firing rates were remarkably

similar across these conditions. Kaposvari et al. (2018) also included a neutral condition in which the image presentation order was randomised within a sequence. They reported suppressed multiunit firing rates for expected compared to neutral stimuli between 100-140 ms post stimulus onset, although this was only statistically significant in one of the two macaques that were tested. This suppression effect was also much smaller in magnitude than expectation violation (i.e. surprise) effects in their design. Similar suppression effects were observed in macaque extrastriate cortex in an equivalent design when measuring fMRI signals in structured (predictable) compared to random (neutral) sequences (Vergnienx & Vogels, 2020). Taken together, these studies provide evidence for ES across electrophysiological and fMRI measures.

Neuroimaging studies of human participants have not yielded evidence for ES in statistical learning designs. Turk-Browne et al. (2009) presented short sequences of 12 stimuli in pre-learned, structured sequences and in randomly-ordered sequences. Unlike the macaque studies above, they reported larger BOLD signals in bilateral lateral occipital complex (LOC) for predictable compared to random sequences (i.e. the opposite of ES). Another study using a very similar design (Otsuka & Saiki, 2017) did not report statistically significant differences between predictable and random sequences in the PPA either (as also reported in Turk-Browne et al., 2010), however they did observe smaller BOLD signals for predictable sequences in left posterior cingulate cortex. Richter et al. (2018) included sequences of two stimuli, whereby the first stimulus either perfectly predicted the second one (expected condition), or cued one of two images to appear with equal probability (neutral condition). They did not find BOLD signal differences in V1 or LOC, and Bayes factors indicated preferential evidence for the null hypothesis across these ROIs.

To our knowledge, there is only one human MEG experiment that included expected and neutral stimulus conditions. Mahanova et al. (2018) presented structured and randomly-ordered sequences and did not find signal amplitude differences between these conditions. However, stimuli in both of these conditions evoked smaller responses compared to a third sequence type in which pre-existing learned sequence-based expectations were violated (i.e. a surprising stimulus condition).

Overall, there is evidence for ES from non-human primate studies that involved multiple weeks of sequence learning (Ramachandran et al., 2017; Kaposvari et al., 2018; Vergniew & Vogels, 2020). However, ES was not reliably observed when measuring BOLD or MEG signal amplitudes in humans when sequences were learned over shorter time scales. It is possible that ES as observed in statistical learning designs may be specific to situations whereby extensive training has been undertaken, which may be associated with changes in synaptic connections occurring over extended time scales throughout an organism's life (e.g., Yotsumoto et al., 2008). This could potentially explain why similar effects on invasive electrophysiological measures are not observed when statistical regularities are learned over shorter time scales, such as within or across experimental blocks (e.g., Kaliukhovich & Vogels, 2011, 2014; Vinken et al., 2018). Similar distinctions between short- and longer-term learning effects have also been observed in relation to motor action-linked sensory predictions (reviewed in Dogge et al., 2019).

One additional consideration that is particular to statistical learning designs is that, with repeated paired presentation of two images, neurons which initially fire robustly to only one image will begin responding to both images. This phenomenon is known as pair coding (e.g., Miyashita, 1988; Erickson & Desimone, 1999). As noted by Ramachandran et al., (2017), a possible consequence of pair coding is that with repeated stimulation of the same neurons there may occur some degree of cross-stimulus adaptation (e.g., Sawamura et al., 2006). The timing and duration of the suppression effects in Kaposvari et al. (2018) and Ramachandran et al. (2017) are broadly consistent with effects of cross-adaptation in macaque IT neurons observed in Sawamura et al. (2006). However, this cross-adaptation effect would be bidirectional and occur regardless of the stimulus presentation order (e.g., for stimulus pairs AB and BA). Suppression was only found in Kaposvari et al. (2018) when the presentation order of paired stimuli was consistent across training and test phases, indicating that cross-adaptation is unlikely to fully account for the observed ES effects.

3.3 Oddball and Roving Stimulus Designs

In oddball designs the same stimulus is presented in high and low probability contexts (Figure 2A). In the high probability context, a critical stimulus A is presented often within a sequence as the ‘standard’ stimulus, which is interspersed with a rare ‘deviant’ stimulus B (e.g., AABAAAABAA). In the low probability context, the same stimulus image is instead presented as a rare deviant (e.g., BBBBABBBBBAB). Differences in visual stimulus-evoked ERPs that are related to expectations can be observed from approximately 150-300 ms post stimulus onset (Czigler et al., 2004; Kimura et al., 2009; Stefanics et al., 2011), which are commonly known as visual mismatch responses or the visual mismatch negativity (vMMN; reviewed in Stefanics et al., 2014; Kremlacek et al., 2016). Despite the large literature on visual mismatch effects, there is little evidence for ES effects that cannot be accounted for by confounding factors. The vast majority of oddball and roving stimulus designs are not well-suited to testing for ES and are more appropriate for specifically detecting effects of expectation violation.

In visual oddball designs, the expected standard stimulus is almost always a repetition of the preceding stimulus, whereas the surprising deviants are seldom repeated within the sequence, resulting in an inherent stimulus repetition confound. In order to isolate effects of surprise that are not due to stimulus repetition effects, researchers have also included additional sequences in which many different stimuli (including the deviant and standard stimulus images) are randomly interspersed. In these so-called equiprobable control sequences (e.g., Jacobsen & Schroger, 2001; Amado & Kovács, 2016) each stimulus image is presented with the same probability as the deviant stimuli in the classical oddball sequences. Consequently, expectations cannot be reliably formed for any particular stimulus image in these sequences. By comparing responses to deviants and equiprobable stimuli, researchers have reported evidence of surprise effects on ERPs (Czigler et al., 2002; Astikainen, et al., 2008; Kimura et al., 2009; Amado & Kovács, 2016). However, comparisons between expected and equiprobable conditions remain confounded by stimulus repetition effects. In addition, the same [deviant – equiprobable] differences were not found when measuring firing rates and LFPs from macaque IT neurons (Kaliukhovich & Vogels, 2014).

Recently, Feuerriegel et al. (2018b) developed a design which controls for immediate stimulus repetition effects. They identified differences in ERPs evoked by surprising and expected conditions that were consistent with earlier observations of visual mismatch responses. In a follow-up experiment (Feuerriegel et al., 2021) they presented expected, neutral and surprising stimuli to isolate effects of ES and surprise. They reported clear evidence for surprise responses which occurred when an observer's expectations were violated. However, they did not find any evidence for ES when comparing expected and neutral stimuli, with Bayes factors indicating evidence favouring the null hypothesis. A similar pattern of results could be observed in the group-averaged ERP waveforms in an earlier study by Stefanics et al. (2011) that similarly accounted for repetition effects within an oddball-type design. However, they did not statistically compare ERP amplitudes across expected and neutral conditions.

A number of experiments have instead used roving stimulus designs (e.g., Stefanics et al., 2018; Smout et al., 2019), in which trains of consecutively repeated stimuli are presented to an observer (e.g., AAAAABBBBBBCCCC) and expectations are assumed to be weighted toward those stimuli which have been recently presented. These designs also have an inherent stimulus repetition confound; the expected stimuli are almost always an immediate repetition of the previous stimulus. For this reason, it is difficult to determine evidence of ES based on findings from these experiments.

Figure 2. Examples of oddball, action-outcome learning and apparent motion designs.

A) In classical oddball sequences a commonly-appearing standard stimulus is interspersed with a rarely-appearing deviant. In equiprobable sequences the same stimuli are interspersed with a large number of different images, which each appear with equal probability. B) In action-outcome learning designs a motor action (such as a button press) is paired with the appearance of a specific stimulus (expected condition), or is paired with multiple stimuli which could each appear with equal probability (neutral condition). C) In apparent motion designs the appearance of a stimulus at a particular location is either expected based on the immediate history of apparent motion (in linear motion sequences) or is preceded by a randomly-generated sequence of stimulus locations (in random sequences). Purple outlines denote expected stimuli, orange outlines denote surprising stimuli, and green outlines denote neutral stimuli. Example stimulus images are taken from Rossion and Pourtois (2004).

3.4. Action-Outcome Learning Designs

In action-outcome learning designs, participants' motor actions (e.g., button presses) are consistently associated with a specific visual event, for example the appearance of a face or house image (e.g., Blakemore et al., 1999; Bednark et al., 2015; depicted in Figure 2B). These experiments are designed to emulate the causal relationships between our motor actions and their sensory consequences that are ubiquitous in

everyday life. For example, we quickly learn that flicking a light switch will turn on a light, which makes the contents of a room visible. In neutral conditions, the participant's motor action is paired with multiple sensory events that can each occur with equal probability.

Across two fMRI experiments there was not consistent evidence for ES. Yon et al. (2018) presented a virtual hand avatar to participants and manipulated whether this avatar moved congruently with the participant's own finger movements. They found larger BOLD signals in incongruent conditions, whereby the visual feedback did not match the participant's motor actions, in bilateral occipital cortex, left occipital cortex and right occipitotemporal cortex. However, this was only found for voxels that were tuned away from the visually-presented finger movement, and not for voxels that responded preferentially to that stimulus. This pattern indicates a sharpening of neural population responses, rather than a general suppression of neural activity that is often inferred from observations of [expected - surprising] differences using univariate measures. This pattern is also not entirely consistent with surprise effects in incongruent trials that are associated with a capture of attention, as this would be expected to increase BOLD signals for voxels tuned to the stimuli. Hindy et al. (2019) trained participants to expect the appearance of specific fractal stimuli following left and right button presses, and also included a condition in which one of two different images could appear regardless of which button was pressed (i.e. a neutral condition). They did not observe differences in BOLD signals across expected and neutral conditions in early visual cortex or the hippocampus.

3.5. Apparent Motion Designs

In apparent motion designs, expectations are entrained by the presentation of a stimulus that moves in discrete steps along a predictable motion trajectory (Figure 2C), or by alternately presenting stimuli in two different locations to entrain a perception of rapid back-and-forth motion. In addition to testing for effects of ES, studies using these designs have uncovered evidence of response latency shifts that are consistent with predictive motion extrapolation in the visual system (e.g., Berry et al., 1999; Jancke et al., 2004; Blom et al., 2020; reviewed in Hogendoorn & Burkitt, 2019).

Two fMRI studies (Alink et al., 2010; Shen et al., 2020) provide some very tentative evidence for ES, although these effects can also be potentially explained by a surprise-related capture of attention (Alink et al., 2010, discussed in Alink & Blank, 2021) or by another confound that is not related to expectations about the critical stimulus (Shen et al., 2020). Effects indicative of ES were not reported when measuring ERPs.

Alink et al. (2010) induced apparent motion by alternatively flashing horizontal bars in the top and bottom right hand corners of a display. They then presented an additional target bar that appeared just above the bar in the lower right hand corner. This target bar could be either presented at a time that is congruent or incongruent with the perceived trajectory of apparent motion. They reported that BOLD signals in V1 were larger when the timing of the target bar's appearance violated expectations established by the apparent motion trajectory. However, it is unclear whether the effects seen in this design reflect attentional capture that occurs when motion-linked expectations are violated (i.e. a surprise response) or suppression that may occur when the bar conforms to the expected trajectory (discussed in Alink & Blank, 2021).

Shen et al. (2020) presented Gabor patches in the top and bottom right corners of a visual field to induce apparent motion. They then measured BOLD signals relative to the onset of a third Gabor that was presented in a location congruent with the perceived motion trajectory (i.e. an expected stimulus location). Responses to the third Gabor in same location were also measured when the preceding two Gabors flickered simultaneously at both top right and bottom right locations, which did not induce a perception of apparent motion (i.e. a neutral location). They reported lower BOLD signal amplitudes in the apparent motion condition in V2, however this reduction was already apparent in both V1 and V2 several seconds before the third Gabor was presented. This indicates that the observed suppression was likely a consequence of BOLD signal differences related to alternating or simultaneous presentation of the preceding stimuli, which makes their findings difficult to interpret.

Experiments that recorded ERPs did not find evidence congruent with ES. Across two experiments, Pollux et al. (2011) presented a black bar at different locations on a display. They presented sequences following predictable linear motion

trajectories and randomly-scrambled permutations of these sequences. They did not find visual P1-N1 peak-to-peak ERP amplitude differences across these conditions in either experiment, which was consistent with their earlier study using the same design (Pollux & Guo, 2009). Doherty et al. (2005) also recorded EEG and presented sequences in which a red circle moved from left to right across a display, either following a linear motion trajectory or a random trajectory. In addition, the consistency of the interstimulus interval was either fixed or random in different conditions, allowing a dissociation of spatial and temporal predictability. They reported larger P1 component amplitudes in conditions with a predictable linear motion trajectory compared to unpredictable trajectories, which is consistent with effects of directed spatial attention on visual evoked ERPs (e.g., Mangun, 1995; Eimer, 1998) but is not consistent with ES.

4. Summarising the Extent of Evidence for Expectation Suppression

We have critically reviewed the evidence for ES in the visual system across five widely used paradigms, each of which involve predictive cueing of visual stimulus categories, exemplars or images. There was evidence for ES across a subset of experiments that recorded neural activity in macaque IT cortex, which involved multiple weeks of sequence learning prior to neural activity measurement (Ramachandran et al., 2017; Kaposvari et al., 2018; Vergniew & Vogels, 2020). Across other designs, whereby stimulus appearance probabilities were learned within one or two testing sessions, evidence for ES was inconsistent. Only a minority of fMRI studies found evidence for ES; a larger number reported either null effects or larger BOLD signals for expected compared to neutral stimuli. In addition, there was scant evidence of ES as measured using electrophysiological recordings in these designs.

If ES is instead interpreted as the absence of a surprise-related response that occurs when expectations are violated (rather than a suppression or inhibition of stimulus-evoked responses) then there appears to be substantial evidence for such an effect. In this case, the term ‘expectation suppression’ is arguably a misnomer, and [surprising – expected] differences in neural activity measures should be relabelled in terms of surprise-related responses. It is still unclear to what extent the reported

surprise-related responses reflect the relative contributions of prediction error signalling (Smout et al., 2019), capture of attention that occurs when an observer's expectations have been violated (Larsson & Smith, 2012; Richter et al., 2018; discussed in Alink & Blank 2021), and increased pupil dilation following surprising events (Richter & de Lange, 2019). Attention-based accounts of surprise responses predict that stimulus-specific multivariate neural representations will be enhanced for surprising compared to expected stimuli, which has been observed in most (e.g., Jiang et al., 2013; Richter et al., 2018; Tang et al., 2018) but not all (e.g., Kok et al., 2012b; Yon et al., 2018) previous studies that tested for expectation effects on decoding performance.

We note that our review has only covered experiments that manipulated expectations relating to visual stimuli, and that there is a large literature on audition and somatosensory perception that lies outside the scope of this paper. We recommend that findings within these sensory modalities be critically reviewed in a similar way. We also caution that our review explicitly focused on changes in neural measures *within* the visual system, and that expectation-related suppression has sometimes been observed in frontal and parietal areas that are implicated in processes such as perceptual decision-making (e.g., Rahnev et al., 2011). Our review does not encompass such decision-related brain areas because contemporary models of expectation suppression have focused on neural activity within sensory systems.

Here, we also recognise that an absence (or scarcity) of evidence does not necessarily imply evidence of absence. Consequently, we believe that it is critical to develop experiments that better isolate and quantify effects of ES and control for confounding factors. Below we explore the theoretical implications of an absence of ES in the visual system to underscore why designing adequate tests for this effect is critically important. We conclude our review by providing guidelines on how to better design experiments that aim to test for ES.

5. Implications for Models of Predictive Processing

An absence of ES would appear to be problematic for a number of predictive processing theories that specifically describe the suppression or inhibition of evoked responses for expected visual stimuli (e.g., Bays & Wolpert, 2007; Summerfield & Egner, 2009; Summerfield & de Lange 2014; Grotheer & Kovács, 2016; de Lange et al., 2018; Richter et al., 2018). We will not engage with the particulars of each theoretical framework here. However, it appears that these theories might be better framed in terms of surprise-related responses that occur when expectations are violated by unexpected sensory events, rather than suppression or ‘silencing’ of expected sensory representations (e.g., as described in Alink et al., 2010; Kok et al., 2012a; Richter & de Lange, 2019). We are also careful to distinguish between the predictive coding-based theories mentioned above and those that are grounded in active inference (e.g., Parr et al., 2019; Corcoran et al., 2020; Hohwy, 2020). The latter are not process models and do not explicitly rely on observations of ES.

The apparent absence of ES does, however, appear to be consistent with a recently-developed opposing process model of predictive processing and perception (Press et al., 2020; for a similar computational model see Heeger, 2017). This model describes predictive mechanisms that bias perceptual inferences toward those sensory events which are expected, where this occurs prior to the expected event. This may be enacted, for example, via pre-activation of neural representations of expected stimuli (as reported in Kok et al., 2017; Blom et al., 2020; Robinson et al., 2020; see also Albright, 2012). This model also describes neural response enhancements following detection of expectation violations, which can be detected later in the time course of the stimulus-evoked response. Importantly, this model specifies that surprise responses occur specifically in situations where an observer’s expectations are violated and the sensory event is sufficiently unexpected. Accordingly, this model is not explicitly built upon the concept of ES as defined in our review.

Here, we note that the model proposed by Press et al. (2020) defines surprisal as a continuous variable that is quantified as Kullback-Liebler divergence (KLD; see also Itti & Baldi, 2009), in contrast to the categorical neutral/surprising conditions described in this review. It is unclear whether surprise responses are elicited once KLD reaches a fixed threshold to become sufficiently unexpected (where this threshold may

vary across contexts and across individuals), or whether increases in KLD render surprise responses as more gradually likely to occur. In the latter case, increases in the prevalence of surprise responses may be easily confused with increases in the magnitude of surprise responses in trial-averaged measures of neural activity.

6. Implications for Separable Prediction and Prediction Error Units

Some variants of predictive coding models posit that there are two separable populations of neurons in the visual system, which are concentrated in different cortical layers: one population that encodes predictions (i.e. representations of the causes of sensory inputs) and another that encodes prediction errors (e.g., Friston, 2010; Wacongne et al., 2012; Bastos et al., 2012; reviewed in Walsh et al., 2020). One possibility that is congruent with the findings reviewed here is that the contributions of prediction and prediction-error units to BOLD signals are roughly equivalent, so that any reductions in BOLD signals due to fulfilled expectations would be counteracted by the increased activity of prediction/representation units in expected stimulus conditions. For example, neural activity that is related to preparation for a specific stimulus image, such as pre-activation of expected stimulus representations that can occur even before the stimulus is scheduled to appear (e.g., Kok et al., 2017; Blom et al., 2020), might increase BOLD signal magnitudes in trials where a specific stimulus is expected. When assessing the observed null results from studies using time-resolved electrophysiological recording methods, this would also imply that prediction and prediction error units would have almost identical time-courses of activity with respect to stimulus onset, and similar contributions to EEG measures summated at the scalp. However, this notion is incongruent with model-derived estimates that prediction error units contribute twice as much to the BOLD signal compared to prediction/representation units (Egner et al., 2010).

Predictive coding theories differ in their specification of prediction and prediction error units (discussed in Walsh et al., 2020; Cao, 2020). When testing specific accounts that describe different anatomical or laminar distributions of each unit (e.g., Feldman & Friston, 2010; Bastos et al, 2012) decoding methods could be useful to assess whether multivariate patterns of neural activity can distinguish

between expected and neutral conditions. This approach could be used on data from invasive electrophysiological recordings (such as LFPs) to detect subtle differences in the relative activity of prediction and prediction error units. Laminar fMRI may also be useful for identifying effects on BOLD signals that are dominant in different layers of cortex (Muckli et al., 2015; Lawrence et al., 2019).

7. Implications for Theoretical Accounts of Repetition Suppression

Several predictive coding-based models have been proposed to account for RS (e.g., Summerfield et al., 2008; Summerfield & de Lange, 2014; Aukztulewicz & Friston, 2016). In these models, an observer's expectations are persistently weighted toward those stimuli which have been encountered in the immediate past, resulting in ES when stimuli are repeated. If neural responses are not actually suppressed when an observer's expectations are fulfilled, then this indicates that RS cannot be easily explained as a product of ES within these frameworks.

An absence of ES instead provides preferential support for an alternative class of predictive coding models, which specify changes in synaptic connectivity that are predominantly dependent on recent stimulation history (e.g., Garrido et al., 2009). These findings also provide preferential support for a circuit-based models of RS, which describe different effects on response profiles of both excitatory and inhibitory neurons following stimulus repetition without reference to ES or other modulatory mechanisms that are specified in predictive coding accounts (e.g., Dhruv, et al., 2011; Solomon & Kohn, 2014; Vogels, 2016; Kaliukhovich & Vogels, 2016; Whitmire & Stanley, 2016; Natan et al., 2015, 2017). The latter class of models can also explicitly accommodate associated effects of inherited adaptation (Dhruv et al., 2011; Kohn, 2007; Larsson et al., 2016) and secondary effects on the activity of inhibitory interneurons (e.g., Wissig & Kohn, 2012; Patterson et al., 2013; Kaliukhovich & Vogels, 2016).

Of course, this does not mean that RS does not facilitate some predictive function within the brain. Attenuating sensory responses to recently-seen stimuli can preferentially enhance visual evoked responses to events that are novel or surprising,

which could allow an organism to differentiate between commonplace and unusual events (Ulanovsky et al., 2003; Vogels, 2016). Even RS that occurs via local circuit mechanisms and inherited adaptation can presumably be incorporated into many predictive coding models without issue (Walsh et al., 2020).

8. Expectations and Perceptual Decision-Making

Several studies reviewed here did report expectation-related effects on choice proportions and response times in perceptual decision-making tasks (e.g., Rahnev et al., 2011; Vinken et al., 2018) and such effects are well-established in the literature (e.g., Mulder et al., 2012; van Ravenzwaaij et al., 2012; Yon et al. 2020). ES has been proposed to underlie improved decision-making performance for expected stimuli via sharpening of population-level neural responses (e.g., Kok et al., 2012b; de Lange et al., 2018), whereby activity of stimulus-selective neurons tuned *away* from the expected stimulus is reduced, rather than activity of neurons that are maximally responsive to that stimulus. This is the opposite of that proposed by suppression-based predictive coding models mentioned above, which describe larger reductions in activity for neurons that are tuned toward the stimulus. The role of expectations in biasing or influencing decision-relevant *sensory* representations has also been recently called into question (Bang & Rahnev, 2017; Rungratsameetaweemana et al., 2018). Indeed, several studies using multivariate decoding methods have found that neural representations of stimuli are better discriminable when stimuli are surprising rather than when they are expected (e.g., Jiang et al., 2013; Richter et al., 2018; Tang et al., 2018). So, why do we typically observe performance advantages (such as faster response times) when an observer's expectations are fulfilled?

This performance advantage may instead be enacted via distributed effects that occur both before and after the sensory encoding of stimulus attributes. When the presentation of a specific stimulus is linked to a specific motor action, expectations to see that stimulus also result in a build-up of preparatory activity in motor cortex which precedes stimulus onset (e.g., de Lange et al., 2013; Kelly et al., 2020). There is also evidence that neural representations of expected sensory events can be pre-activated just prior to the arrival of afferent sensory input (e.g., Kok et al., 2017; Blom

et al., 2020) which may comprise an early source of decision-relevant sensory evidence that biases perceptual decisions in favour of expected events (Summerfield & de Lange, 2014; Yon et al., 2018; Feuerriegel et al., 2020). This stimulus-selective pre-activation could also plausibly lead to observations of sharpened multivariate representations for expected stimuli in fMRI BOLD measures that capture contributions of both pre- and peri-stimulus activity (e.g., Kok et al., 2012b; Yon et al., 2018). Effects of expectations are also observed across frontal and parietal areas (e.g., Rahnev et al., 2011; Cavanagh & Frank, 2014; Wessel, 2017) which likely reflect adjustments to decision-making processes that occur downstream of sensory encoding stages. Similar processes may also underlie the phenomenon of repetition priming, which has been shown to consistently correlate with neural activity measures in frontal areas but not with the extent of RS in the visual system (McMahon & Olson, 2007; Horner, 2012; Wig, 2012; Schacter et al., 2012). Given the evidence for this range of sensory-, decision- and motor-related processes that optimise decision-making in line with an observer's expectations, it is unclear whether the concept of ES is necessary to explain expectation-related effects on decision-making performance.

9. Guidelines for Designing Experiments to Test for Expectation Suppression

Isolating effects of ES is a formidable task and requires great care to account for associated confounds relating to attention, surprise-related responses, stimulus novelty and repetition/adaptation effects. Here, we propose some suggestions for designing experiments which may better isolate and test for ES effects in the visual system.

One potential reason why ES was seldom reported is because the sizes of these effects are much smaller than those of surprise responses, meaning that they may be difficult to detect using sample sizes typical of most human neuroimaging experiments. Where feasible, large samples and high numbers of trials per participant are advised to isolate and quantify ES effects. As ES is theorised to occur to some extent in each individual (e.g., Friston, 2005), the use of small-*N* designs, in which very large amounts of data are collected from each individual, may be an efficient way to achieve sufficient precision (Smith & Little, 2018). Precise measures in single

participants may also be useful for verifying whether [surprising – expected] and [expected – neutral] differences are actually exhibited by most or all individuals.

In addition, the expectation-related effects in fMRI studies did not replicate when using electrophysiological measures of neural activity. Experiments testing for ES should employ a range of recording techniques under similar or identical stimulation conditions. Time-resolved electrophysiological measures may be especially useful for determining the latency of expectation-related effects with respect to the encoding of stimulus features, which is useful for dissociating attention-based and prediction-based accounts of ES (discussed in Alink & Blank, 2021).

As emphasised throughout this review, the appropriate design of expected and neutral stimulus conditions is also critical for minimising bias in the measurement of ES. The following suggestions may be useful to consider when designing experiments that minimise confounds while measuring ES.

As evident from the work reviewed above, distinguishing neutral from expected and surprising stimulus conditions can be challenging, and may depend on assumptions regarding how participants' expectations are entrained and updated throughout the experiment. These assumptions should be made explicit, and should be mathematically formalised wherever possible (e.g., Bell et al., 2016). In some cases, calculating quantitative measures of expectancy or surprise may be preferred over using qualitative (i.e., categorical) condition labels. Parametric manipulations of expectation-related variables, such as stimulus appearance probability (e.g., Feuerriegel et al., 2021) or transition probability may allow closer investigation of how such variables map onto neural response measures than when using expected/neutral/surprising categories. Kullback-Liebler divergence may prove a particularly useful metric for quantifying surprise, as it requires calculation of the precision or certainty associated with both prior expectations and posterior beliefs (Itti & Baldi, 2009; Press et al., 2020). Estimating an observer's degree of certainty in their predictions, for example using hierarchical learning models (e.g., Lawson et al., 2021) may be particularly important when rates of learning and expectation updating might vary across conditions or over the course of an experiment. For example, the prior information that is used to derive expectations is known to differ across

conditions with relatively stable and volatile cue-outcome contingencies (e.g., Behrens et al., 2007; Lawson et al., 2021). Here we caution that, given the observed prevalence of surprise effects and the lack of [expected - neutral] differences in studies using categorical condition labels, the mapping of model-derived quantities to neural response magnitudes is likely to be highly non-linear. Consequently, application of linear regression models may lead to misleading results and incorrect inferences. In addition, it is likely that some types of expectations are resistant to modification when an observer is presented with contradictory evidence (for examples see Yon et al., 2019).

Potential repetition suppression and adaptation effects (including long-lag repetition effects) could be identified in experiment designs by considering the recency with which the same or similar stimuli were encountered in previous trials, and whether this differs across expected and neutral stimulus conditions. This includes, for example, ensuring that stimuli are not consistently preceded by highly similar stimuli in the previous trial within a given condition (e.g., a Gabor oriented at 90 degrees preceded by a Gabor oriented at 75 degrees). Probabilistic cueing designs, in which presentation probabilities are cued in each trial (e.g., Todorovic & de Lange, 2012), may be more appropriate than other designs that entrain expectations by manipulating the relative presentation frequency of each stimulus (e.g., Bastos et al., 2020). This is because the latter class of designs often suffers from delayed repetition or relative novelty confounds. Effects of recent stimulation history should also be explicitly modelled in probabilistic cueing designs in order to avoid incidental biases.

To account for effects related to stimulus novelty and familiarity, the number of presentations of each stimulus image should be balanced across conditions throughout the experiment. Familiarisation sessions conducted prior to neural activity measurement may also be useful to minimise gradual changes in novelty/familiarity throughout a recording session (e.g., Mahanova et al., 2018; Kaposvari et al., 2018; Rostalski et al., 2020).

We also caution that, even in designs where relevant confounds are minimised or controlled, the interpretation of ES effects that are measured as univariate [expected - neutral] differences may not be straightforward. Overall reductions in

neural response measures could be produced by either ‘cancellation’ or ‘sharpening’ effects, which describe reductions in activity of stimulus-selective neurons that are tuned toward a presented stimulus (cancellation) or tuned away from that stimulus (sharpening, discussed in de Lange et al., 2018; Press et al., 2020). Additional analyses that can distinguish between these two broad classes of neurons, for example measures of fMRI voxel selectivity as inferred using multivariate techniques (e.g., Kok et al., 2012b; Yon et al., 2018; Richter & de Lange, 2019), could be used to better characterise observed univariate ES effects.

10. Heterogeneity in Neutral Stimulus Conditions Across Experiments

Across the studies reviewed here, it appears that the exact design of the neutral condition may be potentially relevant to whether or not ES is observed. The majority of experiments compared an expected stimulus condition, in which a specific image or stimulus attribute could be predicted to appear with more than 50% probability, with a neutral condition in which two stimuli could each appear with 50% probability (e.g., Summerfield & Koechlin, 2006; Egnér et al., 2010; Rahnev et al. 2011; St John-Saaltink et al., 2015; Amado et al., 2016; Hindy et al., 2020). The majority of these studies did not report statistically significant ES effects. It is possible that, in these experiments, the difference in appearance probabilities was too small to drive detectable ES effects. For example, this might occur when the neutral stimulus appears with 50% probability and the expected stimulus appears with 75% probability. This issue is difficult to resolve because most previous studies have used expected/neutral/surprising category labels. Consequently, researchers have not explicitly described the linear or non-linear functions that relate differences in stimulus appearance probability to differences in neural response magnitudes. If this mapping is in fact non-linear across the range of probabilities used in the experiments reviewed above, then this could have potentially prevented detection of categorical [expected – neutral] differences in previous work. To help resolve this issue, models of ES should specify this mapping in advance, or propose multiple, mathematically-explicit mapping schemes that can be tested against neuroimaging data. This will also help to explicate assumptions regarding the relative contributions of representation and prediction error neurons to the measured signals

(discussed in section 6), as well as assumptions regarding how observers form, evaluate and update their expectations over time.

A smaller number of experiments included neutral conditions in which a larger number of stimuli (e.g., between 5-20 stimuli) could each appear with equal probability. Interestingly, a larger proportion of these did find statistically significant [expected – neutral] differences indicative of ES when using fMRI (Pajani et al., 2017; Davis & Hasson, 2018; Vergnieux & Vogels, 2020) and electrophysiological recordings (Ramachandran et al., 2017; Kaposvari et al., 2018). However, three of these studies also involved macaques undergoing multiple weeks of sequence learning prior to neural activity measurement (Ramachandran et al., 2017; Kaposvari et al., 2018; Vergnieux & Vogels, 2020) and one of the two studies of human participants involved a potential novelty confound, which may have mimicked the reported ES effects (Pajani et al., 2017).

Interestingly, neuroimaging and electrophysiological measures of neural activity do appear to differ across these two types of neutral conditions. Rostalski et al. (2020) presented pairs of face images within a trial, which could either be repeated or alternating face identities. In one condition, upon seeing the first face, only two faces (i.e. the repeated face or a specific alternating face identity) could each appear with 50% probability as the second stimulus. In another condition, the probability of image repetition was held constant at 50%, but one of five different alternating face identities could each appear with 10% probability. In the former condition, BOLD signals evoked by alternating face pairs were smaller than the latter condition, in which a larger range of face identities could potentially appear. Similar effects were found when measuring ERPs (Feuerriegel et al., 2018a).

However, it is still unclear how these effects relate to those of expectations based on transition probabilities, and more work is needed to resolve this issue. Importantly, in Rostalski et al. (2020) the *relative* likelihood between the repeated face and each alternating face identity differed across conditions. In the latter condition with five possible alternating faces, upon seeing the first face in the trial a repetition of that face was five times more likely to appear than any other individual face identity that could appear as the second stimulus. If expectations in situations of >2 possible

outcomes are computed based on *relative* likelihood of appearance between each pair of outcomes (similar to decision confidence computations, e.g., Li & Ma, 2020) then the relatively more likely repeated face image might be categorised as an expected stimulus, and each alternating face image would be classed as a relatively surprising stimulus. This may have led to surprise-related BOLD signal increases in this condition, compared to the condition where only one specific alternating face could appear. Alternatively, it may be that there is a non-linear function relating *absolute* stimulus appearance probability to neural response magnitudes, whereby large differences are found between 50% and 10% appearance probability conditions, but not between e.g., the 75% and 50% probability conditions used in many studies reviewed above.

11. Conclusion

In this review we have described four important confounds that are prevalent in studies of ES, and we have reviewed the existing literature with these confounds in mind. We report that there is inconsistent evidence for ES when this effect is defined as a suppression of neural activity for expected compared to neutral stimulus conditions, although ES was consistently found in a subset of statistical learning designs, involving long-term training periods. As the notion of ES is central to several highly-influential theoretical frameworks, there is a need for experiments that are appropriately designed to measure and test for this effect.

Acknowledgements

This work was supported by a University of Melbourne Decision Science Fellowship awarded to D.F. G.K was supported by a Deutsche Forschungsgemeinschaft Grant [grant number KO 3918/5-1]. R.V was supported by Fonds voor Wetenschappelijk Onderzoek (FWO) Vlaanderen (Go88418N) and KU Leuven [grant C14/17/109]. Funding sources had no role in the design, writing or interpretation of this review. We are grateful to Andrew Corcoran for providing feedback on an earlier draft of this manuscript. Declarations of interest: none.

References

- Albright, T. D. (2012). On the perception of probable things: neural substrates of associative memory, imagery, and perception. *Neuron*, 74(2), 227-245.
- Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus predictability reduces responses in primary visual cortex. *Journal of Neuroscience*, 30(8), 2960-2966.
- Alink, A., & Blank, H. (2021). Can expectation suppression be explained by reduced attention to predictable stimuli? *Neuroimage*, 117824.
- Amado, C., Hermann, P., Kovács, P., Grotheer, M., Vidnyánszky, Z., Kovács, G. (2016). The contribution of surprise to the prediction based modulation of fMRI responses. *Neuropsychologia*, 84, 105-112.
- Amado, C., & Kovács, G. (2016). Does surprise enhancement or repetition suppression explain visual mismatch negativity? *European Journal of Neuroscience*, 43(12), 1590-1600.
- Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, 16(7), 390-398.
- Astikainen, P., Lillstrang, E., & Ruusuvirta, T. (2008). Visual mismatch negativity for changes in orientation - A sensory memory-dependent response. *European Journal of Neuroscience*, 28(11), 2319-2324.
- Auksztulewicz, R., Friston, K. (2016). Repetition suppression and its contextual determinants in predictive coding. *Cortex*, 80, 125-140.
- Bang, J. W., & Rahnev, D. (2017). Stimulus expectation alters decision criterion but not sensory signal in perceptual decision making. *Scientific Reports*, 7(1), 1-12.
- Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, 76(4), 695-711.
- Bastos, A. M., Lundqvist, M., Waite, A. S., Kopell, N., & Miller, E. K. (2020). Layer and rhythm specificity for predictive routing. *Proceedings of the National Academy of Sciences*, 117(49), 31459-31469.

- Baylis, G. C., & Rolls, E. T. (1987). Responses of neurons in the inferior temporal cortex in short term and serial recognition memory tasks. *Experimental Brain Research*, 65(3), 614-622.
- Bays, P. M. & Wolpert, D. M. (2007). Computational principles of sensorimotor control that minimize uncertainty and variability. *Journal of Physiology*, 578, 387-396.
- Bednark, J. G., Poonian, S. K., Palghat, K., McFadyen, J., & Cunnington, R. (2015). Identity-specific predictions and implicit measures of agency. *Psychology of Consciousness: Theory, Research, and Practice*, 2(3), 253.
- Bell, A. H., Summerfield, C., Morin, E. L., Malecek, N. J., & Ungerleider, L. G. (2016). Encoding of stimulus probability in macaque inferior temporal cortex. *Current Biology*, 26(17), 2280-2290.
- Berry, M. J., Brivanlou, I.H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature* 398, 334-338.
- Blakemore, S. J., Frith, C. D., & Wolpert, D. M. (1999). Spatio-temporal prediction modulates the perception of self-produced stimuli. *Journal of Cognitive Neuroscience*, 11, 551-559.
- Blom, T., Feuerriegel, D., Johnson, P., Bode, S., & Hogendoorn, H. (2020). Predictions drive neural representations of visual events ahead of incoming sensory information. *Proceedings of the National Academy of Sciences of the U.S.A.*, 117(13), 7510-7515.
- Bode, S., Sewell, D. K., Lilburn, S., Forte, J. D., Smith, P. L., & Stahl, J. (2012). Predicting perceptual decision biases from early brain activity. *Journal of Neuroscience*, 32(36), 12488-12498.
- Busse, L., Katzner, S., & Treue, S. (2008). Temporal dynamics of neuronal modulation during exogenous and endogenous shifts of visual attention in macaque area MT. *Proceedings of the National Academy of Sciences*, 105(42), 16380-16385.

- Cao, R. (2020). New labels for old ideas: Predictive processing and the interpretation of neural signals. *Review of Philosophy and Psychology*. doi 10.1007/s13164-020-00481-x
- Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, 18(8), 414-421.
- Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., & Petersen, S. E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science*, 248, 1556e1559.
- Corcoran, A. W., Pezzulo, G., & Hohwy, J. (2020). From allostatic agents to counterfactual cognisers: Active inference, biological regulation, and the origins of cognition. *Biology & Philosophy*, 35(3), 1-45.
- Czigler, I., Balazs, L., & Pato, L. G. (2004). Visual change detection: Event-related potentials are dependent on stimulus location in humans. *Neuroscience Letters*, 364(3), 149-153.
- Czigler, I., Balazs, L., & Winkler, I. (2002). Memory-based detection of task-irrelevant visual changes. *Psychophysiology*, 39(6), 869-873.
- Davis, B., & Hasson, U. (2018). Predictability of what or where reduces brain activity, but a bottleneck occurs when both are predictable. *NeuroImage*, 167, 224-236.
- De Baene, W., & Vogels, R. (2010). Effects of adaptation on the stimulus selectivity of macaque inferior temporal spiking activity and local field potentials. *Cerebral Cortex*, 20(9), 2145-2165.
- de Lange, F. P., Rahnev, D. A., Donner, T. H., & Lau, H. (2013). Prestimulus oscillatory activity over motor cortex reflects perceptual expectations. *Journal of Neuroscience*, 33(4), 1400-1410.
- de Lange, F. P., Heilbron, M., & Kok, P. (2018). How do expectations shape perception? *Trends in Cognitive Sciences*, 22(9), 764-779.
- Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences of the United States of America*, 93(24), 13494-13499.

- Dhruv, N. T., Tailby, C., Sokol, S. H., & Lennie, P. (2011). Multiple adaptable mechanisms early in the primate visual pathway. *Journal of Neuroscience*, 31, 15016–15025.
- Dogge, M., Custers, R., & Aarts, H. (2019). Moving forward: On the limits of motor-based forward models. *Trends in Cognitive Sciences*, 23(9), 743-753.
- Doherty, J. R., Rao, A., Mesulam, M. M., & Nobre, A. C. (2005). Synergistic effect of combined temporal and spatial expectations on visual attention. *Journal of Neuroscience*, 25(36), 8259-8266.
- Dunovan, K., & Wheeler, M. E. (2018). Computational and neural signatures of pre and post-sensory expectation bias in inferior temporal cortex. *Scientific Reports*, 8(1), 1-13.
- Egner, T., Monti, J. M., & Summerfield, C. (2010). Expectation and surprise determine neural population responses in the ventral visual stream. *Journal of Neuroscience*, 30(49), 16601-16608.
- Eimer, M. (1998). Mechanisms of visuospatial attention: Evidence from event-related brain potentials. *Visual Cognition*, 5, 257–286.
- Erickson, C. A., & Desimone, R. (1999). Responses of macaque perirhinal neurons during and after visual stimulus association learning. *Journal of Neuroscience*, 19(23), 10404 –10416.
- Ewbank, M. P., von Dem Hagen, E. A., Powell, T. E., Henson, R. N., & Calder, A. J. (2016). The effect of perceptual expectation on repetition suppression to faces is not modulated by variation in autistic traits. *Cortex*, 80, 51-60.
- Falmagne, J. C. (1965). Stochastic models for choice reaction time with applications to experimental results. *Journal of Mathematical Psychology*, 2(1), 77-124.
- Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, 4, 1–23.
- Feuerriegel, D. (2016). Selecting appropriate designs and comparison conditions in repetition paradigms. *Cortex*, 80, 196-205.

- Feuerriegel, D., Churches, O., Coussens, S., & Keage, H. A. D. (2018a). Evidence for spatiotemporally distinct effects of image repetition and perceptual expectations as measured by event-related potentials. *Neuroimage*, 169, 94–105.
- Feuerriegel, D., Keage, H. A. D., Rossion, B., & Quek, G. L. (2018b). Immediate stimulus repetition abolishes stimulus expectation and surprise effects in fast periodic visual oddball designs. *Biological Psychology*, 138, 110–125.
- Feuerriegel, D., Yook, J., Quek, G. L., Hogendoorn, H., & Bode, S. (2021). Visual mismatch responses index surprise signalling but not expectation suppression. *Cortex*, 134, 16–29.
- Feuerriegel, D., Blom, T., & Hogendoorn, H. (2020). Predictive activation of sensory representations as a source of evidence in perceptual decision-making. *Cortex*. doi 10.1016/j.cortex.2020.12.008
- File, D., File, B., Bodnár, F., Sulykos, I., Kecskés-Kovács, K., & Czigler, I. (2017). Visual mismatch negativity (vMMN) for low-and high-level deviances: A control study. *Attention, Perception, & Psychophysics*, 79(7), 2153–2170.
- File, D., & Czigler, I. (2019). Automatic detection of violations of statistical regularities in the periphery is affected by the focus of spatial attention: A visual mismatch negativity study. *European Journal of Neuroscience*, 49(10), 1348–1356.
- Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 458–467.
- Freedman, D., J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2006). Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. *Cerebral Cortex*, 16, 1631–1644.
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 360(1456), 815–836.
- Friston, K.J. 2010. The free-energy principle: A unified brain theory? *Nat. Rev. Neurosci.* 11: 127–138.

- Garrido, M. I., Kilner, J. M., Kiebel, S. J., Stephan, K. E., Baldeweg, T., & Friston, K. J. (2009). Repetition suppression and plasticity in the human brain. *Neuroimage*, 48(1), 269-279.
- Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, 10(1), 14-23.
- Grotheer, M., & Kovács, G. (2014). Repetition probability effects depend on prior experiences. *Journal of Neuroscience*, 34, 6640-6646.
- Grotheer, M., & Kovács, G. (2015). The relationship between stimulus repetitions and fulfilled expectations. *Neuropsychologia*, 67, 175-182.
- Grotheer, M., & Kovács, G. (2016). Can predictive coding explain repetition suppression? *Cortex*, 80, 113-124.
- Hall, M. G., Mattingley, J. B., & Dux, P. E. (2018). Electrophysiological correlates of incidentally learned expectations in human vision. *Journal of Neurophysiology*, 119(4), 1461-1470.
- Heeger, D. J. (2017). Theory of cortical function. *Proceedings of the National Academy of Sciences*, 114(8), 1773-1782.
- Henson, R. N., Rylands, A., Ross, E., Vuilleumier, P., & Rugg, M. D. (2004). The effect of repetition lag on electrophysiological and haemodynamic correlates of visual object priming. *Neuroimage*, 21(4), 1674-1689.
- Henson, R. N. (2016). Repetition suppression to faces in the fusiform face area: A personal and dynamic journey. *Cortex*, 80, 174-184.
- Hindy, N. C., Avery, E. W., & Turk-Browne, N. B. (2019). Hippocampal-neocortical interactions sharpen over time for predictive actions. *Nature Communications*, 10(1), 1-13.
- Hogendoorn, H., & Burkitt, A. N. (2019). Predictive coding with neural transmission delays: A real-time temporal alignment hypothesis. *Eneuro*, 6(2).
- Horner, A. (2012). Focusing on the frontal cortex. *Cognitive Neuroscience*, 3(3-4), 246-247.

- Hsu, Y. F., Le Bars, S., Hämäläinen, J. A., & Waszak, F. (2015). Distinctive representation of mispredicted and unpredicted prediction errors in human electroencephalography. *Journal of Neuroscience*, 35(43), 14653-14660.
- Hsu, Y. F., Hämäläinen, J. A., & Waszak, F. (2018). The processing of mispredicted and unpredicted sensory inputs interact differently with attention. *Neuropsychologia*, 111, 85-91.
- Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, 49(10), 1295-1306.
- Jacobsen, T., & Schroger, E. (2001). Is there pre-attentive memory-based comparison of pitch? *Psychophysiology*, 38(4), 723-727.
- Jancke, D., Erlhagen, W., Schöner, G., & Dinse, H. R. (2004). Shorter latencies formotion trajectories than for flashes in population responses of cat primary visual cortex. *Journal of Physiology*, 556, 971-982.
- Jiang, J., Summerfield, C., & Egner, T. (2013). Attention sharpens the distinction between expected and unexpected percepts in the visual brain. *Journal of Neuroscience*, 33(47), 18438-18447.
- Kaliukhovich, D. A., & Vogels, R. (2011). Stimulus repetition probability does not affect repetition suppression in macaque inferior temporal cortex. *Cerebral Cortex*, 21, 1547-1558.
- Kaliukhovich, D. A., Vogels, R. (2014). Neurons in macaque inferior temporal cortex show no surprise response to deviants in visual oddball sequences. *Journal of Neuroscience*, 34, 12801-12815.
- Kaliukhovich, D. A., & Vogels, R. (2016). Divisive normalization predicts adaptation-induced response changes in macaque inferior temporal cortex. *Journal of Neuroscience*, 36(22), 6116-6128.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17(11), 4302-4311.

- Kaposvari, P., Kumar, S., & Vogels, R. (2018). Statistical learning signals in macaque inferior temporal cortex. *Cerebral Cortex*, 28(1), 250-266.
- Kelly, S. P., Corbett, E. A., & O'Connell, R. G. (2020). Neurocomputational mechanisms of prior-informed perceptual decision-making in humans. *Nature Human Behaviour*, 1-15.
- Kimura, M., Katayama, J., Ohira, H., & Schroger, E. (2009). Visual mismatch negativity: New evidence from the equiprobable paradigm. *Psychophysiology*, 46(2), 402-409.
- Kimura, M., Schröger, E., Czigler, I., & Ohira, H. (2010). Human visual system automatically encodes sequential regularities of discrete events. *Journal of Cognitive Neuroscience*, 22(6), 1124-1139.
- Kok, P., Rahnev, D., Jehee, J. F., Lau, H. C., & De Lange, F. P. (2012a). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex*, 22(9), 2197-2206.
- Kok, P., Jehee, J. F., & De Lange, F. P. (2012b). Less is more: Expectation sharpens representations in the primary visual cortex. *Neuron*, 75(2), 265-270.
- Kok, P., Van Lieshout, L. L., & De Lange, F. P. (2016). Local expectation violations result in global activity gain in primary visual cortex. *Scientific Reports*, 6, 37706.
- Kok, P., Mostert, P., & de Lange, F. P. (2017). Prior expectations induce prestimulus sensory templates. *Proceedings of the National Academy of Sciences of the U.S.A.*, 114(39), 10473-10478.
- Kohn, A. (2007). Visual adaptation: Physiology, mechanisms, and functional benefits. *Journal of Neurophysiology*, 97(5), 3155-3164.
- Kovács, G., Kaiser, D., Kaliukhovich, D. A., Vidnyánszky, Z., & Vogels, R. (2013). Repetition probability does not affect fMRI repetition suppression for objects. *Journal of Neuroscience*, 33(23), 9805-9812.
- Kovács, G., & Vogels, R. (2014). When does repetition suppression depend on repetition probability? *Frontiers in Human Neuroscience*, 8, 685.

- Kremlacek, J., Kreegipuu, K., Tales, A., Astikainen, P., Poldver, N., Naatanen, R., & Stefanics, G. (2016). Visual mismatch negativity (vMMN): A review and meta-analysis of studies in psychiatric and neurological disorders. *Cortex*, 80, 76-112.
- Kronbichler, L., Said-Yürekli, S., Kronbichler, M. (2018). Perceptual expectations of object stimuli modulate repetition suppression in a delayed repetition design. *Scientific Reports*, 8(1), 1-8.
- Larsson, J., & Smith, A. T. (2012). fMRI repetition suppression: Neuronal adaptation or stimulus expectation? *Cerebral Cortex*, 22, 567-576.
- Larsson, J., Solomon, S. G., & Kohn, A. (2016). fMRI adaptation revisited. *Cortex*, 80, 154-160.
- Lawrence, S. J., Formisano, E., Muckli, L., & de Lange, F. P. (2019). Laminar fMRI: Applications for cognitive neuroscience. *Neuroimage*, 197, 785-791.
- Lawson, R. P., Bisby, J., Nord, C. L., Burgess, N., & Rees, G. (2021). The computational, pharmacological, and physiological determinants of sensory learning under uncertainty. *Current Biology*, 31(1), 163-172.
- Li, H. H., & Ma, W. J. (2020). Confidence reports in decision-making with multiple alternatives violate the Bayesian confidence hypothesis. *Nature Communications*, 11(1), 1-11.
- Li, L., Miller, E. K., & Desimone, R. (1993). The representation of stimulus familiarity in anterior inferior temporal cortex. *Journal of Neurophysiology*, 69(6), 1918-1929.
- Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York, NY: Oxford University Press.
- Manahova, M. E., Mostert, P., Kok, P., Schoffelen, J. M., & De Lange, F. P. (2018). Stimulus familiarity and expectation jointly modulate neural activity in the visual ventral stream. *Journal of Cognitive Neuroscience*, 30(9), 1366-1377.
- Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology*, 32, 4-18.

Evaluating the Evidence for Expectation Suppression

- Maunsell, J. H. (2015). Neuronal mechanisms of visual attention. *Annual Review of Vision Science*, 1, 373-391.
- McMahon, D. B. T., & Olson, C. R. (2007). Repetition suppression in monkey inferotemporal cortex: Relation to behavioral priming. *Journal of Neurophysiology*, 97, 3532-3543.
- Meyer, T., Olson, C.R., 2011. Statistical learning of visual transitions in monkey inferotemporal cortex. *Proceedings of the National Academy of Sciences of the U.S.A.*, 108, 19401-19406.
- Meyer, T., Walker, C., Cho, R. Y., & Olson, C. R. (2014). Image familiarization sharpens response dynamics of neurons in inferotemporal cortex. *Nature Neuroscience*, 17(10), 1388-1394.
- Meyer, T., & Rust, N. C. (2018). Single-exposure visual memory judgments are reflected in inferotemporal cortex. *eLife*, 7:e32259
- Miller, E. K., Li, L., & Desimone, R. (1993). Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *Journal of Neuroscience*, 13(4), 1460-1478.
- Miyashita, Y. (1988). Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature*, 335, 817-820.
- Muckli, L., De Martino, F., Vizioli, L., Petro, L. S., Smith, F. W., Ugurbil, K., ... & Yacoub, E. (2015). Contextual feedback to superficial layers of V1. *Current Biology*, 25(20), 2690-2695.
- Mulder, M. J., Wagenmakers, E. J., Ratcliff, R., Boekel, W., & Forstmann, B. U. (2012). Bias in the brain: A diffusion model analysis of prior probability and potential payoff. *Journal of Neuroscience*, 32(7), 2335-2343.
- Mur, M., Ruff, D. A., Bodurka, J., Bandettini, P. A., & Kriegeskorte, N. (2010). Face-identity change activation outside the face system: "Release from adaptation" may not always indicate neuronal selectivity. *Cerebral Cortex*, 20, 2027-2042.

- Natan, R. G., Briguglio, J. J., Mwilambwe-Tshilobo, L., Jones, S. I., Aizenberg, M., Goldberg, E. M., & Geffen, M. N. (2015). Complementary control of sensory adaptation by two types of cortical interneurons. *Elife*, 4, e09868.
- Natan, R. G., Rao, W., & Geffen, M. N. (2017). Cortical interneurons differentially shape frequency tuning following adaptation. *Cell Reports*, 21(4), 878-890.
- Otsuka, S., and Saiki, J. (2017). Neural correlates of implicit knowledge about statistical regularities. *Experimental Brain Research*, 235, 3573-3583.
- Pajani, A., Kouider, S., Roux, P., & de Gardelle, V. (2017). Unsuppressable repetition suppression and exemplar-specific expectation suppression in the fusiform face area. *Scientific Reports*, 7(160).
- Parr, T., Corcoran, A. W., Friston, K. J., & Hohwy, J. (2019). Perceptual awareness and active inference. *Neuroscience of Consciousness*, 2019(1). doi 10.1093/nc/niz012
- Patterson, C. A., Wissig, S. C., & Kohn, A. (2013). Distinct effects of brief and prolonged adaptation on orientation tuning in primary visual cortex. *Journal of Neuroscience*, 33(2), 532-543.
- Pollux, P. M., & Guo, K. (2009). Event-related potential correlates of spatiotemporal regularities in vision. *NeuroReport*, 20(5), 525-530.
- Pollux, P. M. J., Hall, S., Roebuck, H., & Guo, K. (2011). Event-related potential correlates of the interaction between attention and spatiotemporal context regularity in vision. *Neuroscience*, 190, 258-269.
- Press, C., Kok, P., & Yon, D. (2020). The perceptual prediction paradox. *Trends in Cognitive Sciences*, 24(1), 13-24.
- Rahnev, D., Lau, H., & De Lange, F. P. (2011). Prior expectation modulates the interaction between sensory and prefrontal regions in the human brain. *Journal of Neuroscience*, 31(29), 10741-10748.
- Ramachandran, S., Meyer, T., & Olson, C. R. (2016). Prediction suppression in monkey inferotemporal cortex depends on the conditional probability between images. *Journal of Neurophysiology*, 115(1), 355-362.

- Ramachandran, S., Meyer, T., & Olson, C. R. (2017). Prediction suppression and surprise enhancement in monkey inferotemporal cortex. *Journal of Neurophysiology*, 118(1), 374-382.
- Rangarajan, V., Jacques, C., Knight, R., Weiner, K., & Grill-Spector, K. (2020). Diverse temporal dynamics of repetition suppression revealed by intracranial recordings in human ventral temporal cortex. *Cerebral Cortex*. doi 10.1093/cercor/bhaa173
- Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2, 79-87.
- van Ravenzwaaij, D., Mulder, M., Tuerlinckx, F., & Wagenmakers, E. J. (2012). Do the dynamics of prior information depend on task context? An analysis of optimal performance and an empirical test. *Frontiers in Psychology*, 3, 132.
- Reynolds, J. H., & Heeger, D. J. (2009). The normalization model of attention. *Neuron*, 61(2), 168-185.
- Richter, D., Ekman, M., & de Lange, F. P. (2018). Suppressed sensory response to predictable object stimuli throughout the ventral visual stream. *Journal of Neuroscience*, 38(34), 7452-7461.
- Richter, D., & De Lange, F. P. (2019). Statistical learning attenuates visual activity only for attended stimuli. *eLife*, 8, e47869.
- Ringo, J. L. (1996). Stimulus specific adaptation in inferior temporal and medial temporal cortex of the monkey. *Behavioural Brain Research*, 76(1-2), 191-197.
- Robinson, A. K., Grootswagers, T., Shatek, S. M., Gerboni, J., Holcombe, A., & Carlson, T. A. (2020). Overlapping neural representations for the position of visible and imagined objects. *ArXiv*, arXiv:2010.09932.
- Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, 33, 217-236.

- Rostalski, S. M., Amado, C., Kovács, G., & Feuerriegel, D. (2020). Measures of repetition suppression in the Fusiform Face Area are inflated by co-occurring effects of statistically learned visual associations. *Cortex*.
- Rungratsameetaweemana, N., Itthipuripat, S., Salazar, A., & Serences, J. T. (2018). Expectations do not alter early sensory processing during perceptual decision-making. *Journal of Neuroscience*, 38, 5632–5648.
- Rungratsameetaweemana, N., & Serences, J. T. (2019). Dissociating the impact of attention and expectation on early sensory processing. *Current Opinion in Psychology*, 29, 181-186.
- Sawamura, H., Orban, G. A., & Vogels, R. (2006). Selectivity of neuronal adaptation does not match response selectivity: A single-cell study of the fMRI adaptation paradigm. *Neuron*, 49, 307-318.
- Schacter, D. L., Wig, G. S., & Stevens, W. D. (2012). Reductions in cortical activity during priming. *Current Opinion in Neurobiology*, 17, 171-176.
- Schomaker, J., & Meeter, M. (2015). Short- and long-lasting consequences of novelty, deviance and surprise on brain and cognition. *Neuroscience and Biobehavioural Reviews*, 55, 268-279.
- Schwiedrzik, C. M., & Freiwald, W. A. (2017). High-level prediction signals in a low-level area of the macaque face-processing hierarchy. *Neuron*, 96(1), 89-97.
- Serences, J. T., & Kastner, S. (2014). A multi-level account of selective attention. *The Oxford Handbook of Attention*, 76–104.
- Shen, L., Han, B., & de Lange, F. P. (2020). Apparent motion induces activity suppression in early visual cortex and impairs visual detection. *Journal of Neuroscience*, 40(28), 5471-5479.
- Smith, P. L., & Little, D. R. (2018). Small is beautiful: In defense of the small-N design. *Psychonomic Bulletin & Review*, 25(6), 2083-2101.
- Smout, C. A., Tang, M. F., Garrido, M. I., & Mattingley, J. B. (2019). Attention promotes the neural encoding of prediction errors. *PLoS Biology*, 17(7), e3000368

Evaluating the Evidence for Expectation Suppression

- Smout, C. A., Garrido, M. I., & Mattingley, J. B. (2020). Global effects of feature-based attention depend on surprise. *NeuroImage*, 116785.
- Solomon, S. G., & Kohn, A. (2014). Moving sensory adaptation beyond suppressive effects in single neurons. *Current Biology*, 24(20), R1012-R1022.
- St. John-Saaltink, E., Utzerath, C., Kok, P., Lau, H. C., & De Lange, F. P. (2015). Expectation suppression in early visual cortex depends on task set. *PLoS One*, 10(6), e0131172.
- Stefanics, G., Kimura, M., & Czigler, I. (2011). Visual mismatch negativity reveals automatic detection of sequential regularity violation. *Frontiers in Human Neuroscience*, 5, 46.
- Stefanics, G., Astikainen, P., & Czigler, I. (2014). Visual mismatch negativity (vMMN): A prediction error signal in the visual modality. *Frontiers in Human Neuroscience*, 8, 1074.
- Stefanics, G., Heinzle, J., Horvath, A.A., & Stephan, K.E. (2018). Visual mismatch negativity and predictive coding: A computational single-trial ERP study. *The Journal of Neuroscience*, 38(16), 4020-4030.
- Summerfield, C., & Koechlin, E. (2008). A neural representation of prior information during perceptual inference. *Neuron*, 59(2), 336-347.
- Summerfield, C., Trittschuh, E.H., Monti, J.M., Mesulam, M.-M., Egner, T., 2008. Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience*, 11, 1004-1006.
- Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, 13(9), 403-409.
- Summerfield, C., & Egner, T. (2016). Feature-based attention and feature-based expectation. *Trends in Cognitive Sciences*, 20(6), 401-404.
- Summerfield, C., Wyart, V., Johnen, V. M., & de Gardelle, V. (2011). Human scalp electroencephalography reveals that repetition suppression varies with expectation. *Frontiers in Human Neuroscience*, 5, 1-13.

- Summerfield, C., & de Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. *Nature Reviews Neuroscience*, 15, 745-756.
- Tang, M. F., Smout, C. A., Arabzadeh, E., & Mattingley, J. B. (2018). Prediction error and repetition suppression have distinct effects on neural representations of visual information. *eLife*, 7, e33123.
- Thomas, D. G., Grice, J. W., Najm-Briscoe, R. G., & Williams Miller, J. (2004). The influence of unequal numbers of trials on comparisons of average event-related potentials. *Developmental Neuropsychology*, 26(3), 753-774.
- Todorovic, A., de Lange, F. P. (2012). Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *Journal of Neuroscience*, 32, 13389-13395.
- Treue, S. (2003). Visual attention: The where, what, how and why of saliency. *Current Opinion in Neurobiology*, 13(4), 428-432.
- Turk-Browne, N. B., Jungé, J. A., & Scholl, B.J. (2005) The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, 134, 552-564.
- Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience*, 21, 1934-45.
- Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning. *Journal of Neuroscience*, 30(33), 11177-11187.
- Turk-Browne N. B., Simon M. G., & Sederberg P. B. (2012) Scene representations in parahippocampal cortex depend on temporal context. *Journal of Neuroscience*, 32, 7202-7207
- Ulanovsky, N., Las, L., & Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nature Neuroscience*, 6(4), 391-398.

- Ulanovsky, N., Las, L., Farkas, D., & Nelken, I. (2004). Multiple time scales of adaptation in auditory cortex neurons. *Journal of Neuroscience*, 24(46), 10440-10453.
- Uzterath, C., St John-Saaltink, E., Buitelaar, J., & de Lange, F.P. (2017). Repetition suppression to objects is modulated by stimulus-specific expectations. *Scientific Reports*, 7, 8781.
- Vergnieux, V., & Vogels, R. (2020). Statistical Learning Signals for Complex Visual Images in Macaque Early Visual Cortex. *Frontiers in Neuroscience*.
- Vinken, K., Vogels, R., & de Beeck, H. O. (2017a). Recent visual experience shapes visual processing in rats through stimulus-specific adaptation and response enhancement. *Current Biology*, 27(6), 914-919.
- Vinken, K., & Vogels, R. (2017b). Adaptation can explain evidence for encoding of probabilistic information in macaque inferior temporal cortex. *Current Biology*, 27, R1210-R1211.
- Vinken, K., Op de Beeck, H., & Vogels, R. (2018). Face repetition probability does not affect repetition suppression in macaque inferotemporal cortex. *Journal of Neuroscience*, 38(34), 7492-7504.
- Vogels, R. (2016). Sources of adaptation of inferior temporal cortical responses. *Cortex*, 80, 185-195.
- Wacongne, C., Changeux, J. P., & Dehaene, S. (2012). A neuronal model of predictive coding accounting for the mismatch negativity. *Journal of Neuroscience*, 32(11), 3665-3678.
- Walsh, K. S., McGovern, D. P., Clark, A., & O'Connell, R. G. (2020). Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Annals of the New York Academy of Sciences*, 1464(1), 242.
- Wessel, J. R. (2017). An adaptive orienting theory of error processing. *Psychophysiology*, 55, 1-21.
- Wessel, J. R., & Aron, A. R. (2017). On the globality of motor suppression: Unexpected events and their influence on behavior and cognition. *Neuron*, 93(2), 259-280.

Evaluating the Evidence for Expectation Suppression

- Whitmire, C. J., & Stanley, G. B. (2016). Rapid sensory adaptation redux: A circuit perspective. *Neuron*, 92(2), 298-315.
- Wig, G. (2012). Repetition suppression and repetition priming are processing outcomes. *Cognitive Neuroscience*, 3(3-4), 247-248.
- Wissig, S. C., & Kohn, A. (2012). The influence of surround suppression on adaptation effects in primary visual cortex. *Journal of Neurophysiology*, 107(12), 3370-3384.
- Woloszyn L & Sheinberg DL (2012). Effects of long-term visual experience on responses of distinct classes of single units in inferior temporal cortex, *Neuron*, 74(1), 193-205.
- Xiang, J., & Brown, M. W. (1998). Differential neuronal encoding of novelty, familiarity and recency in regions of the anterior temporal lobe. *Neuropharmacology*, 37, 657-676.
- Yon, D., Gilbert, S. J., de Lange, F. P., & Press, C. (2018). Action sharpens sensory representations of expected outcomes. *Nature Communications*, 9(1), 1-8.
- Yon, D., Zainzinger, V., de Lange, F. P., Eimer, M., & Press, C. (2020). Action biases perceptual decisions toward expected outcomes. *Journal of Experimental Psychology: General*.
- Yotsumoto, Y., Watanabe, T., & Sasaki, Y. (2008). Different dynamics of performance and brain activation in the time course of perceptual learning. *Neuron*, 57(6), 827-833.