PREDICTION-RELATED NEURAL RESPONSE ALTERATIONS IN THE VENTRAL VISUAL STREAM

for the acquisition of the academic degree

doctor philosophiae (Dr. phil.)

presented to the council of the Faculty of Social and Behavioural Sciences

Friedrich-Schiller-University Jena

Submitted by

Sophie-Marie Rostalski

Born 06.09.1992 in Gotha

Biological Psychology and Cognitive Neurosciences

First examiner: Prof. Dr. Gyula Z. Kovács (FSU Jena)

Second examiner: Prof. Dr. Stefan R. Schweinberger (FSU Jena)

Oral examination: 11.10.2021

Table of Contents

ZUSAMMENFASSUNG

Theorien der prädiktiven Kodierung (*predictive coding*, PC; Rao & Ballard, 1999) spielen in der neurokognitiven Forschung zu Erklärung von Denk- und Wahrnehmungsprozessen in verschiedenen Domänen eine entscheidende Rolle. Das Grundprinzip dieser Theorien ist, dass die Wahrnehmung nicht nur auf der Bottom-up (aufwärts)-Verarbeitung des sensorischen Inputs, sondern auch auf Top-down (abwärts)-Vorhersagen beruht. Mit anderen Worten: Vorerfahrungen und Kontextinformationen erleichtern Wahrnehmungsprozesse und erhöhen dadurch deren Effizienz (Auksztulewicz & Friston, 2016; Friston, 2010). Ziel ist es, die eingehenden Informationen auf der Basis probabilistisch generierter Modelle bestmöglich vorherzusagen. Je besser die Vorhersagen zum sensorischen Input passen, desto kleiner ist der sogenannte Vorhersagefehler (*prediction error*; PE), der verarbeitet wird, um unser Modell für präzisere zukünftige Vorhersagen zu aktualisieren. Dieser Ansatz liefert Erklärungsmöglichkeiten für sehr unterschiedliche Phänomene, wie visuelle Mehrdeutigkeit, optische Täuschungen und neuronale Wahrnehmungsphänomene. In der vorliegenden Arbeit werden mehrere neuronale Antwortveränderungen in kortikalen visuellen Arealen behandelt, die mit Hilfe bildgebender Verfahren gemessen werden. Eines dieser Phänomene ist der zuverlässig beobachtete Effekt einer reduzierten neuronalen Antwort auf wiederholte im Vergleich zu alternierenden Reizen. Die so genannte Wiederholungsunterdrückung (*repetition suppression*; RS) wurde mit der prädiktiven Kodierung in Verbindung gebracht, da angenommen wird, dass Wiederholungen die Stimuli erwartbarer machen, was wiederum zu einem geringeren Vorhersagefehler und damit zu einer verminderten neuronalen Aktivität führt. Es ist jedoch fraglich, ob RS tatsächlich den Vorhersagefehler widerspiegelt oder das Resultat lokaler Prozesse neuronaler Populationen ist, der auch ohne Top-down-Einflüsse auftritt (Grill-Spector et al., 2006). Ein weiterer häufig untersuchter Effekt ist die reduzierte neuronale Reaktion auf erwarteten oder vorhergesagten visuellen Input - genannt Erwartungsunterdrückung (*expectation suppression*, ES). Ein beträchtlicher Teil der Forschung zu kontextabhängigen neuronalen Antworten, wie RS und ES, bezieht sich auf das visuelle System und insbesondere auf das Netzwerk, dass Gesichter verarbeitet. Gesichter in Gehirn werden schnell und effizient in einem komplexen hierarchischen Netzwerk verarbeitet. Darüber hinaus sind sie, obwohl sie die gleiche Grundstruktur haben, individuell und bieten verschiedene Aspekte, wie z.B. Bekanntheit, Identität oder Ausdruck, die in der Forschung über visuelle Verarbeitungsmechanismen, genutzt werden können.

In Studie I gingen wir der Frage nach, ob RS auf PC-Modelle zurückzuführen ist. Es hat sich gezeigt, dass Erwartungen höherer Ordnung RS unter bestimmten Bedingungen modulieren können (Summerfield et al., 2008). Nach PC-Theorien spiegelt RS den Vorhersagefehler wider, der nicht nur von der Diskrepanz zwischen Vorhersage (basierend auf Erwartungen) und sensorischem Input, sondern auch von der geschätzten Präzision der beiden (Auksztulewicz & Friston, 2016) abhängt. Daher sollte die Manipulation der Präzision des sensorischen Inputs das Ausmaß von ebenfalls RS beeinflussen. Wir haben RS für gut sichtbare (hohe Präzision) und verrauschte (niedrige Präzision) Gesichter in einem fMRI-Experiment gemessen. Wir fanden RS in verschiedenen gesichtsverarbeitenden Regionen und eine verringerte neuronale Antwort auf verrauschte Gesichter, aber keine Interaktion. Unsere Ergebnisse deuten darauf hin, dass sensorisches Rauschen für die Bestimmung des Ausmaßes von RS nicht entscheidend ist.

Viele frühere Studien untersuchten RS unter verschiedenen Wiederholungswahrscheinlichkeitsbedingungen und fanden und evaluierten daraufhin ES (Grotheer & Kovács, 2016). Die Effekte dieser Manipulationen könnten jedoch mit einer Charakteristik vermischt sein, die für RS-Experimente natürlich ist. Nachdem man den ersten Stimulus in einem Durchgang gesehen hat, kann der Inhalt des zweiten Bildes genau vorhergesagt werden wenn es wiederholt wird, aber nicht wenn es sich ändert. Dies ist bei allen Studien zu bedenken, die RS untersuchen. Um herauszufinden, welchen Einfluss die Vorhersagbarkeit von Stimuli in Messungen von RS leistet, führten wir in Studie II ein Trainingsexperiment durch, in dem die Teilnehmer spezifische Stimulusassoziationen lernten. In einem Kontext folgte auf ein gegebenes Gesicht A immer Gesicht B (AB-Assoziation), während in einem anderen Kontext eines von fünf verschiedenen Gesichtern zufällig ausgewählt wurde (AX). FMRT Daten, die nach den Trainingssitzungen erhoben wurden, zeigten eine signifikante Reduktion von RS, wenn die alternierenden Gesichter vorhergesagt werden konnten (AB). Mit anderen Worten, wenn für die Unterschiede in Stimulusvorhersagbarkeit nicht kontrolliert wird, kann das RS Messungen signifikant beeinflussen. Dies kann auch auf Messungen von ES übertragen werden und ist daher relevant für das Testen von PC-Theorien.

In der dritten Studie untersuchten wir Adaptions- und vorhersagebezogene Effekte in Bezug auf Erwartungen bestimmte Identitäten betreffend. Hier haben wir uns die Tatsache zunutze gemacht, dass für bekannte Personen bereits neuronale Repräsentationen vorhanden sind. Da angenommen wird, dass Regionen im okzipito-temporalen Kortex eine Rolle bei der Identitätsverarbeitung spielen (Duchaine & Yovel, 2015), untersuchten wir, ob identitätsspezifische Effekte in relevanten gesichtsverarbeitenden Arealen bildinvariant

sind. In kurzen Blöcken präsentierten wir verschiedene natürlich variierende Bilder von entweder verschiedenen oder der gleichen berühmten Person. Zusätzlich wurde in einer Bedingung die wiederholte Präsentation der gleichen Identität mit einem Bild einer anderen Person beendet. Dadurch verletzten wir potentielle identitätsspezifische Erwartungen und stellten die Hypothese auf, dass der Adaptationseffekt aufgelöst wird und sich ein erhöhter PE in den Ergebnissen widerspiegelt. Ein identitätsspezifischer Adaptationseffekt zeigte sich im Gyrus fusiformis, was die Vorstellung einer bildinvarianten Identitätskodierung in dieser Region unterstützt. Der unerwartete Wechsel der Identität löste jedoch kein erhöhtes Signal aus, wie von PC-Theorien vorhergesagt und von früheren elektrophysiologischen Studien gezeigt wurde (Johnston et al., 2016).

Insgesamt zeigen wir, wie wichtig die Vorhersagbarkeit des Stimulus für Studien ist, die RS verwenden, um erwartungsbezogene Effekte aufzudecken. Darüber hinaus sollte der Einfluss der Präzision des sensorischen Inputs auf Messungen von RS und ES in der zukünftigen Forschung mehr Aufmerksamkeit finden. Im Hinblick auf das Stimulusmaterial in den vorgestellten Studien - unbekannte, visuell bekannte und berühmte Gesichter - betonen wir auch die Wichtigkeit einer gründlichen Betrachtung der Charakteristika von Gesichtern in Bezug auf vorherige Erfahrungen und sensorischem Input, sowie Vorhersagbarkeit.

SUMMARY

Theories of predictive coding (PC; Rao & Ballard, 1999) have dominated neurocognitive research in explaining thought and perception processes in various domains. The basic principle is that perception relies not only on bottom-up processing of sensory input but also on top-down predictions. In other words, previous experiences and context information facilitate perceptual processes and thereby increase efficiency (Auksztulewicz & Friston, 2016; Friston, 2010). Together these information processing streams form what we experience as outside world events. The purpose is to provide the best explanation for incoming information based on probabilistically efficient models. The better predictions fit the sensory input, the smaller the so-called prediction error (PE) processed to update our model for more precise future predictions. This explanatory approach is promising, as very different phenomena, such as ambiguity, optical illusions, and neuronal phenomena of perception, can be explained. The current thesis describes several neuronal response alterations in cortical visual areas measured with neuroimaging methods. One of these phenomena is the reliably observed effect of a reduced neuronal response to repeated compared to alternating stimuli. The so-called repetition suppression (RS) effect was connected to predictive coding as repetitions make stimuli more expected, which results in a smaller prediction error and therefore attenuated neuronal activity. Still, it is questioned whether RS reflects the PE or is a local process by neuronal populations that occurs without top-down influences (Grill-Spector et al., 2006). Another often investigated effect is the reduced neuronal response to expected or predicted visual input – called expectation suppression (ES). A considerable body of research on contextual response changes, such as RS and ES, relates to the visual system and the face-processing network in particular. Faces are processed fast and efficiently. Furthermore, although they share the same basic structure, they are individual and offer different aspects that can be addressed in research on visual processing mechanisms, such as familiarity, identity, or expression.

In Study I, we asked the question of whether RS can be explained by PC models. It has been shown that higher-order expectations can modulate RS under specific conditions (Summerfield et al., 2008). According to PC theories, RS reflects the prediction error, which is contingent not only on the mismatch between prediction and sensory input but on the estimated precision of the two (Auksztulewicz & Friston, 2016). Hence, manipulating the precision of sensory input should affect RS magnitude. We measured RS to clear (high precision) and noisy (low precision) face stimuli in an fMRI experiment. We observed RS in

different face-processing regions and a reduced response to noisy faces but no interaction. Our results suggest that sensory noise is not crucial in determining RS magnitude.

Many previous studies investigated RS magnitude under different repetition probability conditions and thereby found and evaluated ES (Grotheer & Kovács, 2016). These manipulations, however, might be confounded with a characteristic that is natural to RS experiments. After seeing the first image of a trial, the content of the second image can be precisely predicted when it is repeated but not when it changes. This is especially important when probabilities of repetition and alternation trials differ. To discover the contribution of stimulus predictability to measures of RS in Study II, we conducted a training experiment in which participants learned specific stimulus associations. In one context, a given face A was always followed by face B, whereas in another context, one of five different faces was chosen randomly as a successor. Importantly, the proportion of repetition and alternation trials was equal for both contexts, and trials were presented interleaved. Neuroimaging data acquired after training sessions revealed a significant reduction in RS magnitude when alternating faces were predictable compared to unpredictable. In other words, controlling for stimulus predictability differences between repetition and alternation can significantly affect measures of RS. This relates to stimulus-specific expectation effects and is therefore relevant for testing consequences of PC.

In the third study, we investigated adaptation and prediction-related effects of expectation about famous faces. Here we have taken advantage of the fact that neuronal representations are already present for well-known identities. As regions in the occipito-temporal cortex were found to play a role in identity processing (Duchaine & Yovel, 2015), we asked whether identity-specific effects are image-invariant in relevant face processing areas. In short blocks, we presented different (ambient) images of either different or the same famous person. Additionally, in one condition, the repeated presentation of the same identity was finished with a picture of another person. Thereby, we violated potential identity-specific expectations and hypothesized a release from adaptation reflecting an enhanced PE. An identity-specific adaptation effect was evident in the fusiform gyrus, supporting the notion of image-invariant identity encoding in this region. However, the change of identity did not provoke a signal increase, as indicated by PC and previous electrophysiological studies (Johnston et al., 2016).

Overall, we demonstrate the importance of stimulus predictability for studies using RS to uncover expectancy-related effects. Furthermore, we suggest that the influence of sensory precision on measures of RS and ES needs more attention in future research. Concerning the stimulus material in the presented studies - unfamiliar, visually familiar, and famous

familiar faces - we also emphasize the importance of thoroughly considering the characteristics of faces in terms of prior belief and sensory input precision and predictability when using them for testing prediction-related effects.

1 INTRODUCTION

In Western thoughts, vision is generally understood as the most important sensory modality. Thinkers as early as Aristotle promoted vision as ‘the noblest of senses’ and therefore privileged over other senses such as hearing, smell, taste, and touch. By now we know, that there are differences in cultures, showing that a strict hierarchy of the senses cannot be a global assumption (Majid et al., 2018). Although vision might not be the most important sense in a global competition, it is almost irreplaceable when guiding us through the world. People rely on visual input when navigating through the streets, when learning or executing simple actions or engaging in conversation. If one thinks about the amount of visual information that a human being encounters every day, what our brain accomplishes at every waking second of our lives seems almost impossible. Especially in social interactions, efficiently recognizing and identifying humans is essential.

It is crucial to differentiate between visual sensation and perception, which will be the all-connecting process with which this thesis deals. Johannes Müller defined sensation as ‘the awareness of the states of sensory nerves, not of the outer world itself’ (Müller, 1838 quoted from Gross, 1999; page 62). Visual perception, in contrast, can be seen as an information-handling process (Attneave, 1954). It can be argued that the human brain could not possibly utilize all the information provided by stimulation states of visual input meeting the retina. Visual processing must be efficient to allow seamless percipience of our surroundings, which is necessary to navigate through the world. Luckily, our world is subject to certain ‘laws’ or ‘regularities.’ There are types of lawfulness that are significant for processes involved in visual perception. Sensory events are highly interdependent in space and time (Dong & Atick, 1995) facilitating our visual perception enormously and can be demonstrated with a simple example.

Imagine watching a person walking down the street while sitting in a café on the other side. Suppose this person disappears for a couple of seconds behind parking cars or some bushes and trees. In that case, one will expect the person to appear again at some point in space and time and automatically generate precise predictions about how this scene continues. We would predict that the person would enter our visual field again at a particular position, namely where, for example the vehicle, no longer obstructs our view, and calculate the time point of this, taking into account the walking speed and the size of whatever blocks our unobstructed view. We would probably also consider that the person could enter a house, but we know that this person will not magically disappear.

Naturally, a lot of information we receive is also redundant, as we can infer stimulation states from previous situations or surrounding sensations. How these fortunate circumstances facilitate our perception is explained in detail in chapter 1.2.1.

This thesis aims to contribute to the already large amount of research literature on how our brain sees and perceives. Specifically, the neural mechanisms underlying the efficient processing of complex visual scenes are still under debate. Depending on the experiments presented, it is indispensable to know the basics about the neuronal structures involved in visual processing and how they work together. Therefore, in chapters 1.1.1 and 1.1.2, I will provide an overview of the neuroanatomical structures of importance and their interconnections. Due to the nature of stimulus material used in all incorporated studies, the focus will be on the ventral visual stream, whose areas are responsible for processing shapes, objects, and faces or bodies and what is therefore also known as the 'what' pathway. Next, as faces are important, in chapter 1.1.3, I will discuss cognitive models of face perception and briefly introduce the differences in processing familiar and unfamiliar faces. Chapter 1.2 will describe the predictive coding framework and repetition suppression (1.2.2) and expectation suppression (1.2.3) as essential mechanisms investigated in the presented studies. Furthermore, I will shed light on the relationship between face processing and prediction processes in chapter 1.3. Finally, at the end of the introduction, I will describe our studies' general and specific aims (chapter 1.4).

1.1 The ventral visual stream

This chapter will introduce the cortical processing network for visual stimuli, focusing on ventral visual areas involved in shape and object processing. It will also provide some background information on the cortical architecture of the visual system and present cortical areas that are specifically dedicated to processing faces and facial information. Furthermore, I will introduce current cognitive models for face processing in chapter 1.1.3.

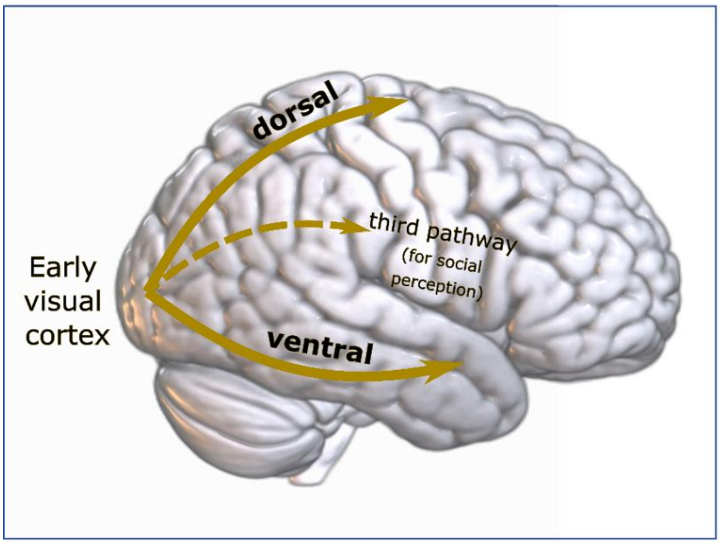


Figure 1. Schematic illustration of visual pathways.

All visual input is received at the retina and processed through the lateral geniculate nucleus (LGN) before it enters the first cortical regions in the occipital lobe, such as the primary (V1) and secondary (V2) visual cortex, as parts of the early visual cortex, and propagated forwards to occipito-temporal (ventral) and temporoparietal (dorsal) regions (Figure 1). The ventral visual stream extends over the occipitotemporal cortex and includes areas holding neural object representations (Kravitz et al., 2013). Because it is responsible for processing and identifying all kinds of objects, it is also known as the ‘what’ pathway. The dorsal visual stream is known as the ‘where’ pathway and it processes information about the location of objects and related actions. This influential model of the visual cortex was recently adapted because there is evidence for a third visual pathway, which is assumed to be specialized for social perception (Pitcher & Ungerleider, 2021).

Before explaining some influential and currently prominent models of face processing and the underlying anatomical network, the next chapter will describe the cortical architecture (1.1.1) because it is crucial for backing up the theories about face processing as well as being an essential basis for the predictive coding framework (1.2.1).

1.1.1 Cortical organization in the visual system

Our visual system's cortical architecture and organization occupied researchers for decades now. The two most extreme points of view were that the visual signal goes through the visual cortex in a strict serial way and is processed in the same manner. Another perspective is that our visual system is a highly connected network of areas, without any hierarchy in processing steps. Both hypotheses seem to be absolute and very unlikely. As with many things, the truth lies somewhere in between. Traditionally, the visual system is seen as hierarchical, multiple, parallel, and unidirectional. Here, I will review some evidence for these characterizations and will show at least one of these features to be false.

Physiological evidence for a hierarchical form of visual processing comes from features of receptive fields in different areas. Receptive fields in early visual areas are smaller compared to those in areas representing a higher-level processing stage (Freeman & Simoncelli, 2011). The two main pathways constituting the parallelism in the visual system have been already mentioned, namely: the dorsal and the ventral system (Ungerleider & Mishkin, 1982). If there is a hierarchy in both systems, the assumption of serial processing is logical. However, the dorsal and ventral stream by no means work separately from each other but are also interconnected. More detailed studies of the connections between cortical structures provided insights into the organization and flow of information. Cortical connections between visual areas reveal different cortical layer profiles (Figure 2B, see Figure 2A for an overview of cortical layers), as found in an early monkey study by Maunsell and Van Essen (1983). They found that connections from visual areas V1, V2, and V3 to the middle temporal visual area (MT) predominantly originate in supragranular layers (layer I – III) of cells. Connections from MT to V1-3 end in infragranular (layers V and VI) layers as well. The authors found the projections from area V1, V2, and V3 to be feedforward going to MT, whereas the reciprocal connections are indicative of feedback projections. This finding reveals that the visual system cannot rely on a unidirectional processing stream but is instead composed of areas that are connected in a bidirectional manner. The different profiles of cortical layers were then used to better characterize the relationships between cortical areas.

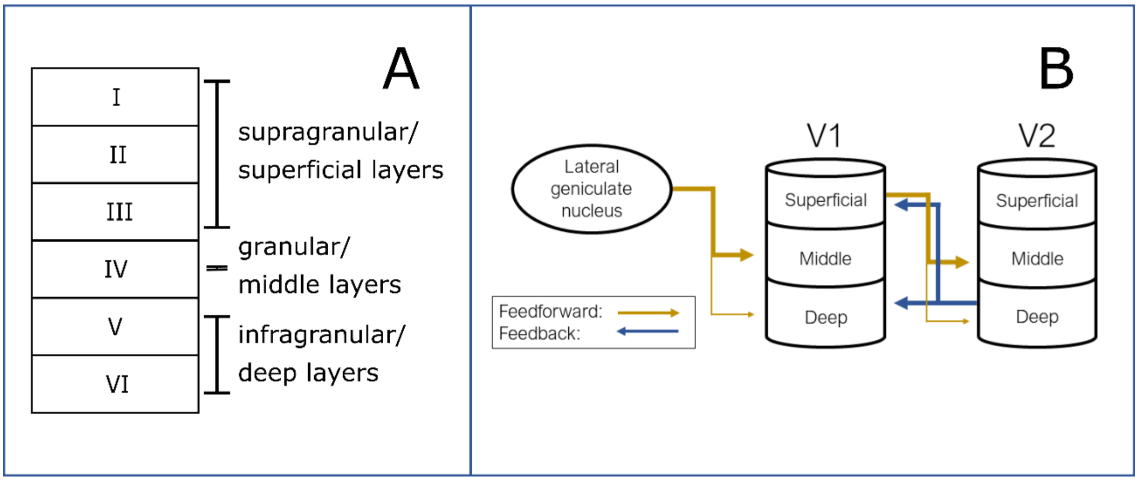


Figure 2. A Schematic illustration of cortical layers of the neocortex. B schematic illustration of cortical layer profiles of feedforward and feedback connections between human lateral geniculate nucleus (LGN) and early visual cortical areas V1 and V2. Adapted from Lawrence et al. (2019).

As early as 1991, Felleman & Van Essen (Felleman & Van Essen, 1991) reported a summary of 32 cortical areas related to vision in the macaque brain. This large number of areas dealing with visual input indicates that a serial processing stream is unlikely. Additionally, the authors report 305 connections between those areas, suggesting a highly connected network. Those pathways were characterized in terms of laminar origin and terminations, including the major types of patterns indicative of ascending, descending, or lateral connections. These laminar patterns were used to define hierarchical relationships between areas. On this basis, Felleman and Van Essen (1991) propose a visual hierarchy model consisting of ten processing levels. At this point in time, it was still an open question whether each level in the hierarchy represents a specific information processing stage that can be characterized and separated from other levels. The authors found processing streams to be distinct but intertwined, supporting the idea of parallel processing. We now know that the top-down connections rather than the bottom-up pathways play a dominant role in visual processing. For instance, the LGN was found to receive a vast amount of information from cortical areas and not from the retina (Casagrande et al., 2006). This asymmetry in extrinsic connections, and the asymmetry in laminar specificity, define the sensory brain as being hierarchically organized. Aside from the presented studies discussed above, the asymmetry between feedforward and feedback connections is also present in occipital and fusiform face-sensitive areas that will be introduced later (Chen et al., 2008).

The overwhelming number of cortico-cortical connections found by Felleman and Van Essen (1991) suggests a highly distributed hierarchical network underlying visual processing. However, the exact role of the different pathways found was still not clear. Ascending pathways were easy to explain, as it is the sensory signal that has to be transported from the senses to areas, representing higher stages in the processing hierarchy. This is also known as bottom-up processing. Top-down processing refers to the addition of information from higher areas in order to better classify sensory information. If we now assume that higher cognitive processing also means more abstract processing (Mumford, 1992), then descending connections make little sense at first glance since lower levels of processing cannot interpret this information at all. However, what they can do is process input from higher levels that have been broken down into a structure similar to that which bottom-up inputs have at that level. Mumford therefore hypothesized that higher-level cortical cells transfer template information in the weights of lower-level area synapses they are connected with via feedback connections. Furthermore, he suggests that these templates must be flexible to guarantee the recognition of an object under various conditions. In other words, presentations are invariant to specific features. Last but not least, Mumford (1992) pointed out that ascending pathways may carry what he termed residuals, the mismatch between the actual sensory input and the template from the higher-order neuronal populations. These considerations led shortly thereafter to the emergence of a theory that claims to explain the neural mechanisms underlying the functioning of the brain. Chapter 1.2.1 is dedicated to the predictive coding framework.

1.1.2 Specialized areas in the ventral visual stream

I will now introduce four main areas in the ventral visual stream, starting with the lateral occipital cortex, and then ascending in the hierarchy to the occipital and fusiform face areas and the anterior temporal lobe. All these areas are depicted in Figure 3 as spherical regions around exemplar coordinates. Especially the occipital and fusiform face areas are essential to the current thesis and were defined in all our studies. However, we did not localize all areas depicted in Figure 3. Therefore, and for illustration purposes, exemplar coordinates were chosen here. However, locations for the lateral occipital cortex and occipital and fusiform gyrus closely resemble our own data (see chapter 2.3). For the anterior temporal lobe coordinates comparable to those in Pobric et al. (2007) were chosen.

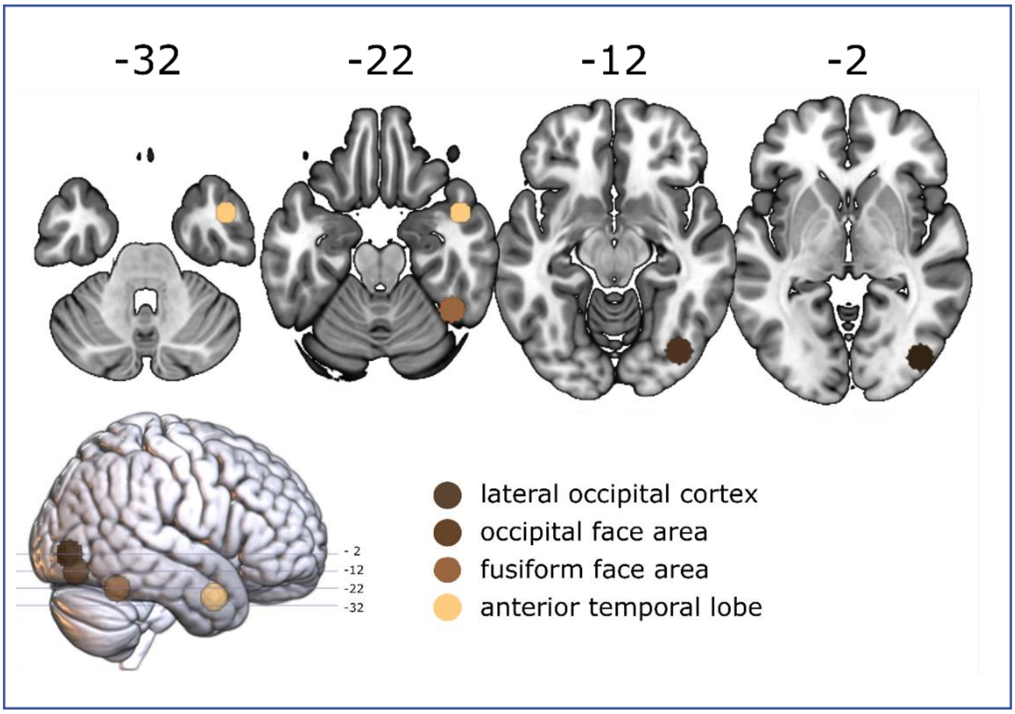


Figure 3. The ventral visual areas, related to face processing. Area locations were created using Marsbar (Brett, 2011) and plotted onto a template brain using MRIcroGL (https://www.nitrc.org/projects/mricrogl/). Each region of interest was defined a sphere around exemplar Montreal Neurological Institute coordinates: Lateral occipital cortex: 42, -80, -2. Occipital face area: 38, -76, -12. Fusiform face area: 42, -52, -22. Anterior temporal lobe: 47, 5, -27.

The *lateral occipital cortex* (LO; Grill-Spector, Kourtzi, & Kanwisher, 2001; Malach et al., 1995) is part of the lateral occipital complex (LOC) and is located in the lateral ventral regions of the occipital lobe. It belongs to the object-processing network and is selective to all visual objects (Grill-Spector et al., 2001) and specifically holds shape representations of objects (Kourtzi & Kanwisher, 2001). Thus, it does not respond to faces specifically, but in terms of faces as being objects as well. It was found to be connected bidirectionally to the occipital and fusiform face areas by a study using dynamic causal modeling (DCM) on functional magnetic resonance imaging (fMRI) data (Nagy et al., 2012). Therefore, it is assumed to play a role in face processing as well. Because faces and objects are not handled differently at this point, the LO might be responsible for the first step of discrimination between objects and faces.

The *occipital face area* (OFA; Gauthier et al., 2000) responds to faces and face parts (Pitcher et al., 2007). It gets input from LO (Nagy et al., 2012) and is assumed to play a role in earlier structural processing of face-related information (Rotshtein et al., 2005). For a long time, the OFA was considered the entry point into a network that deals specifically with facial information. In other words, the OFA detects faces and transfers this information to higher-order cortical areas. Pitcher and colleagues (2007) provided insights into the functional role of the OFA using transcranial magnet stimulation (TMS). TMS of the right OFA disrupted the discrimination of face parts, but only when the authors delivered pulses 60-100ms after stimulus onset. This result supports the assumed early-stage role the OFA plays in face processing. However, whether this area could play a role in processing higher-order facial information is under debate. For example, a lesion in the right inferior occipital cortex caused symptoms of prosopagnosia (Rossion et al., 2003). Thus, this area seems to be crucial for face identification, if only for its role in integrating information from lower and higher-level visual areas, as they showed typical activations in this study. In several other studies using transcranial magnetic stimulation, its specific function was clarified further. Pitcher and colleagues (2008) showed that TMS pulses over OFA disrupt the discrimination of facial expressions and face discrimination itself (Pitcher et al., 2009). TMS pulses over OFA also eliminated training effects in a face-matching task (Ambrus et al., 2017a, 2017b) and reduced priming effects for faces. Later, the OFA was also found to be involved in the association of semantic information to faces using TMS (Eick et al., 2020). Its close connection to areas believed to be higher in the hierarchy, such as regions in the fusiform gyrus (Gschwind et al., 2012), suggests responses that reflect feedback information integrated into the OFA.

We have long known that the *fusiform face area* (FFA; Kanwisher et al., 1997) located on the ventral surface of the temporal lobe is crucial in the processing of faces. Although this area is probably one of the most investigated areas in the ventral visual stream, its specific role in face processing is still fully clarified. As the name suggests, its irreplaceable role in identifying faces (Haxby et al., 2000) seemed inviolable for some time (Grill-Spector et al., 2004). Contradictory to the previously described OFA, it accounts for so-called holistic face processing (Zhang et al., 2012) and shows a high correlation with face-specific activity measured with EEG around 160 ms (N170). This is significantly later than OFA activity, which is related to earlier event-related potential (ERP) latencies (~ 110 ms) (Sadeh et al., 2010).

Furthermore, the FFA was found to show neuronal adaptation effects to a repeated average face image, but not to different face images (Gauthier et al., 2000). Adaptation to different images of the same identity (ID), despite image changes, would suggest an image-invariant identity representation. It is clear that the FFA does not only detect faces but processes them at an individual level. However, whether the FFA holds an image-invariant representation of individual faces is not fully clarified (Davies-Thompson et al., 2009; Kriegeskorte et al., 2007; Xu et al., 2009). For example, Davies-Thompson and colleagues (2013) could not distinguish FFA responses to blocks with different images of one versus images of various identities using adaptation techniques. Furthermore, Kriegeskorte and colleagues (2007) used multivariate pattern analyses (MVPA) on responses to different face identities and could not find distinguishable activations in the FFA. Further studies tried to examine ID representations in the fusiform gyrus using multivariate methods. For famous faces, activation patterns in the FFA were decodable (Axelrod & Yovel, 2015), suggesting an ID representation in this area. Later, blood oxygen level-dependent (BOLD) responses to unfamiliar faces could be distinguished using similar methods (Anzellotti et al., 2014; Nestor et al., 2011), although others failed to find ID representations in areas other than the anterior temporal lobe (Kriegeskorte et al., 2007). A recent study, however, could show that both FFA and OFA hold ID representations (Tsantani et al., 2021). Both areas can discriminate between identities, as revealed with multivariate pattern analyses. Whereas representational distances (computed with representational similarity analyses (RSA)) of faces can be explained mainly by low-level features such as pixel-wise dissimilarity in the OFA, among others, gender and perceived similarity explain distance measurements in the FFA. Overall, the last years' results regarding the nature of ID representations in the FFA were very diverse. By now, the large body of research, including univariate and multivariate analyses on familiarity and ID encoding in relevant regions, suggests that fusiform regions contribute to both (Kovács, 2020). Inconsistencies could be explained, for example, by the fact that feedback from higher areas plays a role, and those influences may not be evident in all studies.

Furthermore, Weiner and Grill-Spector, (2012) proposed the division of the FFA into a posterior (pFFA, but also posterior fusiform gyrus) and middle part (mFFA, but also middle fusiform gyrus). Whether these two separate regions are engaged in different aspects of face processing is still unclear. There is evidence that contrasting faces versus landscapes can best identify the posterior part of the FFA. In contrast, the more anterior part of the FFA responds more robustly to faces than objects, houses, or landscapes (Schwarz et al., 2019). The finding of differences in cellular architectures in posterior and more anterior regions of the fusiform gyrus strengthens the idea of different functional tasks (Grill-Spector et al., 2017). Besides, evidence from MEG recordings suggests different functional roles of posterior versus anterior parts of the fusiform gyrus, as the anterior FFA is activated later than the posterior part (Fan et al., 2020).

Additionally, this area was linked to processing stimuli of visual expertise (Gauthier et al., 2000). Faces are probably the most important class of visual stimuli for humans. The ability to not only identify and recognize individuals but also to extract a large amount of information from their faces offers great advantages in social situations. Still, whether areas that process faces, such as the FFA, are dedicated to faces or stimulus categories we need to process at a high level of expertise are still under debate. Higher neuronal activity to familiar compared to unfamiliar faces and symbols is consistent with the latter notion (Henson et al., 2000). Besides, it seems inefficient that our brain developed a network of areas dedicated to processing just one kind of stimulus. Thus, there is still a debate on whether our expertise for faces is based on neuronal networks and mechanisms dedicated to memorizing and recognizing faces (Kanwisher et al., 1997; Kanwisher & Yovel, 2006) or whether the same network can serve as a basis for gaining expertise in any other category, such as cars or the artificially created Greebles (Gauthier et al., 1999). The general consensus is that we are indeed face experts (McGugin et al., 2019) and that there are specialized areas for the processing and eventually recognition of faces.

The a*nterior temporal pole* (ATL) is located anterior of the FFA in the temporal lobe. It is also known as the anterior face patch and is involved in face perception as well (Harry et al., 2016). The ATL is assumed to be a high-level area in face processing, responsible for person identification (Von Der Heide et al., 2013) and providing semantic information about a person (Chiou & Lambon Ralph, 2018). A region in the anterior inferotemporal gyrus was also found to show differentiable response patterns to two individual faces (Kriegeskorte et al., 2007). Additionally, the ATL showed higher activity for famous than for non-famous faces in a PET study (Gorno-Tempini & Price, 2001). Unlike the other areas in the ventral visual stream, there is evidence for functional lateralization of the anterior temporal face patches (Rice et al., 2018; Von Der Heide et al., 2013). The ATL and earlier face responsive regions as FFA and even OFA were found to show distinct neuronal patterns to familiarized versus unfamiliar faces revealed by an MVPA study (Goesaert & Op de Beeck, 2013).

This thesis will focus on the FFA and OFA, which have been presented above. There is no question that they are areas processing faces and are involved in identification processes, although it is not fully clarified to what extent, especially for the OFA. Also, both areas can be defined reliably using fMRI. For all experimental work presented in this research project, additional sequences for localizing both areas were carried out (chapter 2.3).

Of course, many more regions are involved in processing faces or related information such as sex, age, the familiarity of the face, or identifying a person and connecting a present face with semantic details or memories. Additional regions process changeable aspects in faces, such as viewing direction or facial expressions. Some of them will be mentioned when cognitive and structural models of face perception are introduced in the next chapter (1.1.3).

1.1.3 Cognitive and structural models of face perception

This chapter will focus on models of the face perception network, which is assumed to contain areas functionally specialized in processing faces and facial information, including recognizing and identifying persons or pinpoint emotions. The concept of functional localizationism goes back to the 19th century when experiments showed specific behavioral changes in animals after stimulating certain cortical areas. But scientists concluded that localizationism is challenging to demonstrate because excited cortical regions and observed behavior do not need to be exclusively causally related (see, Gross, 1999). The interconnections to other cortical and subcortical regions (extrinsic connections) were suggested to play a mediative role (for a review, see Phillips et al., 1984). Therefore, it is not surprising that models of neural networks - such as the one responsible for processing faces – became more extensive and complex over time. In recent years, many theories on face processing in the brain have developed (Bruce & Young, 1986; Duchaine & Yovel, 2015; Gobbini & Haxby, 2006; Haxby et al., 2000). Based on the cortical architecture explained in chapter 1.1.1, one can conclude that a hierarchical and highly interconnected network of brain regions enables effortless detection, recognition, and identification of human faces.

As early as 1986, Bruce and Young (Bruce & Young, 1986) proposed a cognitive model for processing faces that included identifying faces and facial expressions. Their model contained units specialized for specific tasks working separately or interactively and serially or in parallel. Nevertheless, this model with its specialized nodes was not linked to neuroanatomical structures.

Haxby and colleagues (2000) proposed a model in which the network underlying face perception is divided into a core and an extended network related to specific cortical and subcortical structures. The core network is assumed to process invariant features of faces, such as identity and changeable aspects of faces, such as eye gaze, lip movements, or emotional expressions for fast visual analysis in three different cortical areas. The inferior occipital gyrus is assumed to be the gateway for the early processing of facial features. It includes the previously described OFA (chapter 1.1). The lateral fusiform gyrus with the face-selective FFA is responsible for invariant aspects of faces, thus identification. The superior temporal sulcus (STS) processes changeable aspects of faces such as eye gaze (Hoffman & Haxby, 2000) or head motion (Duchaine & Yovel, 2015). The extended system contains, among others, anterior temporal regions, the intraparietal sulcus, the amygdala, and insula. In a revision of this original model, Gobbini and Haxby (2007) emphasize the processing of familiar faces.

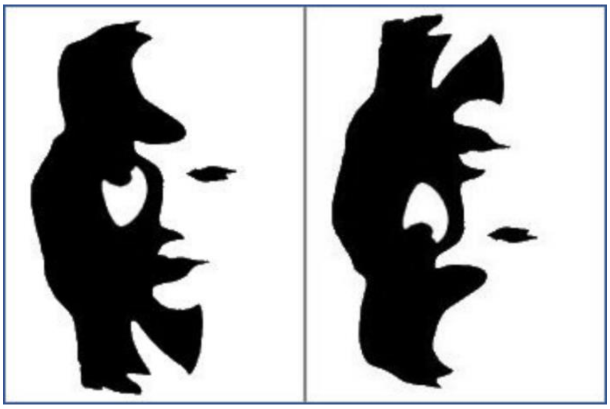


Figure 4. Examples for upright (left panel) and inverted (right panel) Mooney faces. [1]

This illustration is subject to license CC BY-SA 3.0. [https://commons.wikimedia.org/wiki/File:MooneyFaces.jpg]

---

[1] Mooney faces were used to study perceptual closure. One effortlessly perceives a face in the left image but not in the right one. The latter is only an inversion of the former, thereby shares visuospatial characteristics. The perception of a face based on very little information is an excellent example of inference processes in the visual system.

However, the finding of normal FFA activation, despite a lesion in the inferior occipital gyrus (Rossion et al., 2003), led to the conclusion that face responses may emerge independently in those two areas. This idea informed a non-hierarchical model of face perception in which face detection takes place in the FFA, followed by the fine analysis in the OFA (Gentile et al., 2017). In a recent study, the two aforementioned competing models were reviewed by examining the timing of the FFA and OFA in response to different face-like stimuli (Fan et al., 2020). MEG recordings during the presentation of famous familiar and unfamiliar faces revealed the temporal order of peak responses to face stimuli. In the right hemisphere, OFA (around 116 ms after stimulus onset) was activated before posterior FFA (pFFA; 125ms) and anterior FFA (aFFA; 150 ms), although the peak of pFFA was not significantly delayed to OFA peaking time. Regardless of familiarity of the faces, the dynamic sequence observed indicates a bottom-up hierarchical fashion of face processing. When, however, stimuli that can be perceived as faces but lack typical facial features were presented, the temporal relationship of the same areas changed. The authors used so-called Mooney faces (named after C. M. Mooney; Mooney, 1957) as shown in Figure 4 and found a significantly delayed activation of OFA (around 144 ms after stimulus onset) for Mooney faces compared to normal faces. In contrast, FFA activations were similar for normal, and Mooney faces. Thus, the dominant information flow in the absence of typical facial features can be characterized as top-down.

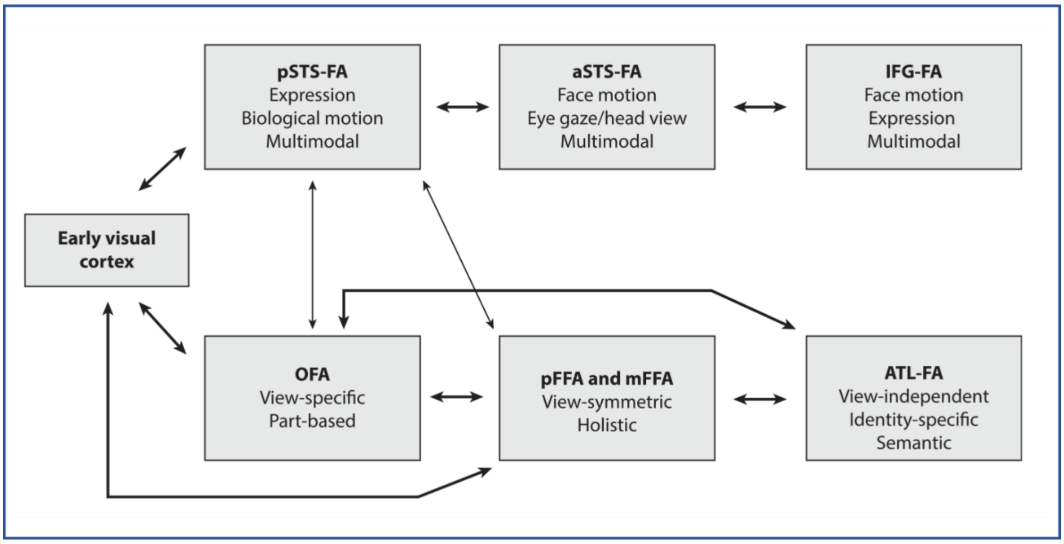


Figure 5. Revised framework for the roles and connections between face-selective areas. The ventral face-processing pathway consists of the occipital face area (OFA), the fusiform face area (FFA), and the anterior temporal lobe face area (ATL-FA), whereas the dorsal face-processing pathway comprises the posterior superior temporal sulcus face area (pSTS-FA), the anterior superior temporal sulcus face area (aSTS-FA), and the inferior frontal gyrus face area (IFG-FA). Adapted from Duchaine & Yovel (2015), permission is pending.

More recently, Duchaine and Yovel (2015) revised the model by Haxby and colleagues (2000), integrating new findings of the last years. Additional areas, such as the inferior frontal gyrus (IFG), were incorporated, and the parallel nature of the visual system with the ventral and dorsal pathways was integrated (Figure 5). The result is a version of the original two-stage model of face perception focusing on face-selective areas.

1.1.4 Familiar and unfamiliar face processing

In the studies included in this thesis, we used facial stimuli with different levels of familiarity. Therefore, in this section, I want to briefly introduce the differences in the processing of unknown/unfamiliar faces versus visually familiar and famous familiar faces. The processing of personally familiar faces will not be discussed here because they are not included in any stimulus material of the present experiments (for a review on their processing, see Ramon and Gobbini, 2018). The fact that we can effortlessly identify people we know under different and even aggravating conditions such as poor lighting conditions or in poor quality images, but experience significant difficulties in distinguishing between two images of different unknown identities (Bruce et al., 2001) shows that the process of perception differs - however, it is still not known at which point. Behavioral studies demonstrated that familiar faces are recognized faster than unfamiliar faces (Burton et al., 1999), and their processing involves different cortical and subcortical areas (for a review, see Kovács, 2020). Further, it has been shown that performance in recognition or matching tasks is negatively affected for unfamiliar faces, for example, due to changes in viewpoint (Bruce et al., 1999). There is additional evidence for quantitatively different neuronal responses to famous in comparison to visually familiar faces (Gobbini and Haxby, 2006) or unfamiliar faces (Eger et al., 2005) in the fusiform gyrus. Other studies showed differences in responses to familiar and unfamiliar faces rather in the extended system than the core face network (Gorno-Tempini & Price, 2001). A summary of neuroimaging studies investigating differences in processing familiar and unfamiliar faces has been provided by Natu and O’Toole (2011). A detailed characterization of how the neuronal representations differ could not yet be provided, as the results were very controversial. FMRI adaptation (fMR-a; Grill-Spector & Malach, 2001), a technique to infer properties of neuronal populations, has been used to investigate the neuronal representations in the FFA. In this case, neuronal adaptation results did not support image-invariant ID representations in this area (Davies-Thompson et al., 2013). This pattern was evident for both – familiar and unfamiliar faces. However, studies applying MVPA showed that fusiform activation patterns to different famous identities are indeed discriminable (Axelrod & Yovel, 2015; Tsantani et al., 2019). Recent EEG and MEG studies support the idea of less robust representations of unfamiliar compared to familiar persons (Barragan-Jason et al., 2015; Dobs et al., 2019). Results from behavioral studies on aftereffects are also in accordance with this notion (Kok et al., 2017; Ryu & Chaudhuri, 2006). In summary, there are significant differences in the neural representation of familiar and unfamiliar faces and how we process them.

At present, there is much speculation and discussion about how faces become familiar, the respective processes in different brain regions, and under which circumstances a stable identity representation develops. Learning faces can be described as the combination of increasing the tolerance of within-person variability and improving the ability to discriminate between different IDs, focusing on the former (Baker & Mondloch, 2019). In a recent review on this issue, Kovács (2020) suggests that although the process of getting to know someone is gradual, the underlying person identity network plays a role at every level of familiarity. This work has demonstrated that the FFA, although being an early region in the face-processing network, is a key area for face processing in all stages of familiarity.

In general, faces are an ideal stimulus type because they fulfill the requirement of being visually experienced stimuli, and they share a similar shape and configuration (Maurer et al., 2002). If normalized faces are superimposed on each other, the result would be blurred but still perceived as a face (Burton et al., 2005; Diamond & Carey, 1986). Still, even photographs of the same identity provide a large variability (Jenkins et al., 2011). Within one class of stimuli, different aspects can be in focus, identity being only one of them. Additionally, different degrees of face familiarity, e.g., unfamiliar versus personally familiar, allow testing of general neural phenomena at different stages of the hierarchical visual system.

Furthermore, neural correlates of processing facial stimuli can be found with electrophysiological measures and neuroimaging methods, and areas belonging to the core system of face processing can be localized reliably (see chapter 1.1.2). Extensive research literature from the last years shows that occipital and fusiform face areas encode familiarity and identity, revealed by univariate and multivariate analyses (Kovács, 2020).

Recapitulating this part of the introduction, one can say that the ventral visual stream, with its cortical architecture and functionally specialized but still highly interconnected areas, provides an excellent basis for studying neuronal mechanisms of perception. In the following, I will introduce fundamental processing mechanisms in the visual system.

1.2 Fundamental neuronal mechanisms of visual processing and their models

This chapter will describe phenomena observed in (visual) sensory processing, such as a reduced response to repeated or expected stimuli (chapters 1.2.2 and 1.2.3). The predictive coding (PC) theory is of central importance to the present work, as it provides explanations for the various neural mechanisms under investigation. Therefore, what follows first introduces a theory that for provides a shared basis for many observations of the previously mentioned phenomena and perhaps even a basis for the general brain functioning:

1.2.1 The predictive coding framework

Already Hermann von Helmholtz assumed perception to be a knowledge-driven inference process based on probabilities (Helmholtz, 1867). This view inspired a lot of research in computational (Musmann, 1979) and neuroscientific research (O'Reilly et al., 2012). Originating from image compressing algorithms in the first place, predictive coding was later transferred to the neural domain, especially the visual system (Lee & Mumford, 2003). Thus, it became an influential theory for perceptual processes of sensory information. By now, modern variations of Helmholtz's idea are among the most popular candidates for explaining how neural information is passed on. In this section, I will introduce the principles of Bayesian inference and how the neocortex can implement it through hierarchical predictive coding. Please note that predictive coding is also referred to as predictive processing and will be used interchangeably in this thesis.

1.2.1.1 The Bayes' theorem

The Bayes' theorem is a probability theory and the basis for Bayesian perceptual psychology, which combines Helmholtz's theoretical idea that the brain must infer causes of sensory inputs from those inputs themselves with statistical inferences (Helmholtz, 1867). The theorem is named after Thomas Bayes and describes the mathematical formula for calculating conditional probabilities (Bayes & Price, 1763).

The Bayes' Theorem is defined as follows:

Equation 1

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)}$$

In Equation 1, the probability of a hypothesis H based on the given data E is calculated using the ratio of the unconditional probability of the hypothesis given the data and the unconditional probability of the data alone. In other words, the posterior probability $P(H|E)$ is estimated by the prior probability $P(H)$ and the likelihood ratio $\frac{P(E|H)}{P(E)}$. Bayesian inference calculates the posterior probability of latent causes based on given data and is a fundamental principle for efficiently processing sensory input.

Predictive coding implements this computational basis in neural processing mechanisms (for a review on Bayesian inference and predictive coding, see Aitchison & Lengyel, 2017). Although there is more than one algorithm assumed to integrate with the predictive coding framework (for a review, see Spratling, 2017), I will only discuss Rao and Ballard's algorithm (Rao & Ballard, 1999) and the free energy principle (Friston, 2009; Friston & Kiebel, 2009). The two are similar and easy to combine. Furthermore, this work deals with empirical Bayes, where in contrast to standard Bayesian methods, distributions are estimated from data and then used to infer prior expectations. It can be described as a bootstrapping method with which statistical independencies in the data inform hierarchical models (Clark, 2013).

1.2.1.2 Predictive processing in cognitive neuroscience

Trying to explain extra-classical receptive field effects (Henry et al., 2013), Rao & Ballard (1999) postulated the predictive coding of natural images as the underlying mechanism of sensory processing. They trained a bidirectional hierarchical network on images of natural scenes and implemented learning algorithms that would allow it to generate an internal model of its image inputs. Further, they described predictive coding as a model for signal transmission through a hierarchical network, the "hierarchical network for predictive coding."

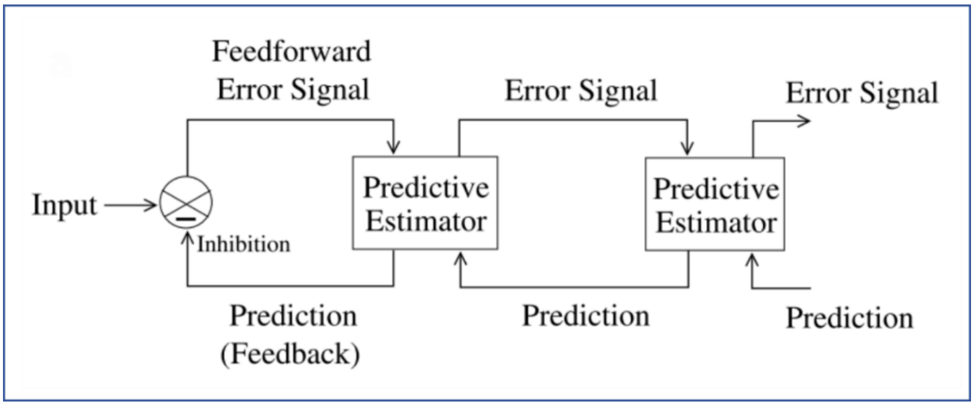


Figure 6. Schematic illustration of hierarchical systems with feedback and feedforward information flow. From Rao & Ballard (1999) with permission.

According to their model (Figure 6), this hierarchical network is composed of several levels capable of detecting residual errors. In other words, the cortical neurons estimate the difference between the visual input and a prediction of this input. For doing so, an internal model of the natural image input is essential. According to Rao and Ballard (1999, p. 80), this hierarchical internal model is learned through 'maximizing the posterior probability of generating the observed data.' A set of neuronal responses, optimized for predicting the input, from each level in the hierarchy are sent as 'predictions' to the next lower level via feedback pathways. Feedforward pathways carry information about the actual input and, more importantly, the error signal from each level (except the lowest, which is the visual stimulus itself) to the next one in the hierarchy. These residual errors carry information about the discrepancy between the prediction and input and are then used by the predictive estimators to adapt the already learned basic vectors to provide a better prediction of the sensory input for the future. In other words, higher-level knowledge is used via top-down connections to generate a construction of the sensory input. Optimally, the sensory signal will be fully 'explained away' (Clark, 2013; Friston, 2005). Since this is very unlikely, there will be some information left over, used to continuously update the internal model. From a computational point of view, one could say that the neuronal activity encodes the beliefs over states in the world. For example, signals from the primary visual cortex reflect the deviation from the expected orientation rather than the actual stimulus orientation itself. Rao & Ballard (1999) further assumed that layer II and III cortical neurons (Figure 2) are key candidates for error detection and signaling those differences. These were found to connect via feedforward to higher visual areas (Maunsell & Van Essen, 1983).

Additionally, to encode predictions and residual or prediction errors (PE), it is assumed that each level in the hierarchy hosts two functionally distinct neural subpopulations (Figure 7). Representation units reflect the expected activity based on the hypothesis about the perceptual input and sent this information to error units in the same and the next lower level. Error units signal the PE, which is the calculated difference between the predicted pattern and the pattern induced by sensory input. This information is sent to representation units in the same and the next higher level via feedforward connections. Prediction errors also depend on the precision of the prediction and the precision of the sensory data (see chapter 1.2.4), which is reflected in error units (Feldman & Friston, 2010).

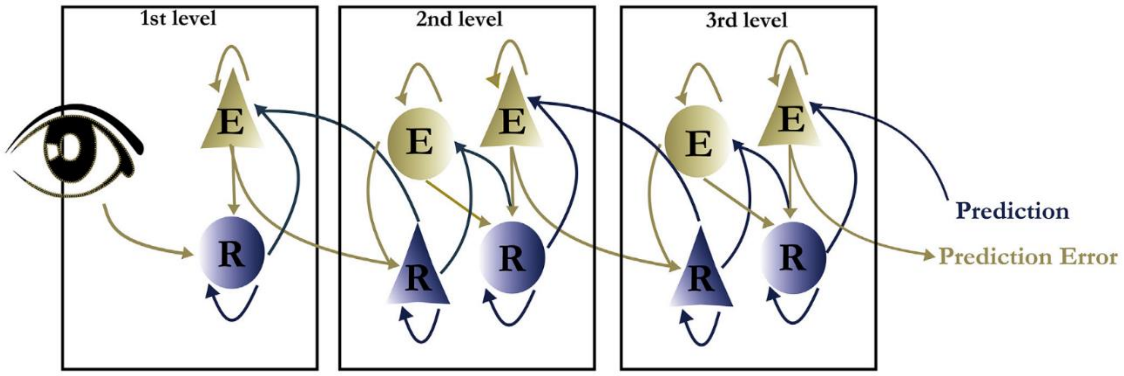


Figure 7. Illustration of the predictive coding model in a schematic presentation of the visual cortex as hierarchical (here consisting of three levels) containing feedback and feedforward loops. In order to optimize perception, deep pyramidal cells (representation units (R); blue circles and triangles) send out predictions about forthcoming perception, and superficial pyramidal cells (error units (E); gold circles and triangles) return the prediction error, the mismatch between the received predictions and the sensory input. Adapted from Grotheer & Kovács (2016), with permission.

Why PC is biologically plausible is described in terms of free energy. The free-energy principle states that any biological self-organizing system's purpose lies in minimizing free energy (Auksztulewicz & Friston, 2016; Friston et al., 2006). Besides PC, other theories of brain function are related to this principle as they all aim for optimization, and according to Karl Friston (Friston, 2010, p. 135), they can be united under one 'Helmholtzian perceptive of the brain as a generative model of the world it inhabits.' The general assumption is that only salient, novel, surprising stimuli are encoded, which is metabolically efficient. From another perspective, high PEs result in increased 'costs' in metabolic systems.

Friston's theory of neural responses (Friston, 2005) was the beginning of a new era in cognitive neuroscience. It combines the functional organization of cortical structures and their connections with the neural mechanisms that allow us to coordinate efficiently through our lives. We know by now that consequences of (potential) predictive processing exist in several modalities and at very different levels of processing. Furthermore, those phenomena can be measured with several techniques and for various features of stimuli, for example in visual perception (Hohwy et al., 2008), music (Koelsch et al., 2018), language (Henderson et al., 2016), and action perception (Ahlheim et al., 2016). Depending on the modality, domain and stimuli investigated, specialized areas will be involved.

A recent meta-analysis supported the idea of a specific but widely distributed network of cortical and subcortical structures engaged in domain-general prediction generation and violation (Siman-Tov et al., 2019). This network includes the inferior frontal gyrus (IFG), the anterior insula (AI), the middle frontal gyrus (MFG), premotor cortical (PMC) regions as well as the pre-supplementary motor area (Pre-SMA) and cortical areas in the temporoparietal junction (TPJ) and also subcortical areas like the striatum, thalamus and subthalamus, and the cerebellum. Especially frontal areas such as the IFG have been linked to prediction-related processing earlier.

1.2.1.3 How the cortical architecture enables predictive processing

I have already introduced the cortical architecture of the visual system briefly in chapter 1.1.1. Here, I want to summarize how this structure may serve as a basis for the predictive processing assumed to be the basis of our perception and general cognition.

Several scientists have addressed the information processing in the visual cortex and its relationship to Bayesian inference (Friston, 2002; Friston & Kiebel, 2009; Kersten et al., 2004; Lee & Mumford, 2003; Rao & Ballard, 1999). The way our cortex is structured is a core foundation, which is necessary to enable inferential processes. As already explained in chapter 1.1.1, the visual cortex organization can be considered both hierarchical (Felleman & Van Essen, 1991) and highly interconnected (Zeki & Shipp, 1988). These characteristics are essential because of another attribute of the human cortical architecture – the functional specialization of cortical areas. Specialization means that an area is specialized for processing particular aspects of sensory input, which becomes meaningful for functional integration (Friston, 2002). Functional specialization is based on the principle of functional segregation within the cortex. Functional segregation, in turn, signifies the grouping of neurons that share the same functional characteristics (Friston, 2005). This is a significant limitation that can be compensated by convergence and divergence of neurons. In other words, functional segregation is a constraint in the cortical structure that requires convergent and divergent cortical connections (Man et al., 2013).

Feedforward connections are assumed to carry stimulus-related information. Feedback and horizontal connections, both within and between cortical areas, form the foundation for providing contextual information (Lamme et al., 1998). Further support for predictive processing in the cortex comes from the fact that backward connections compared to forward connections are more divergent, more numerous, and transcend more levels (Zeki & Shipp, 1988). How these differences between feedforward and backward connections (extrinsic cortical connections) and other architectural principles (Bastos et al., 2012), such as functional asymmetries (Arnal & Giraud, 2012), subserve predictive processing and learning, is explained in detail in Friston (2003).

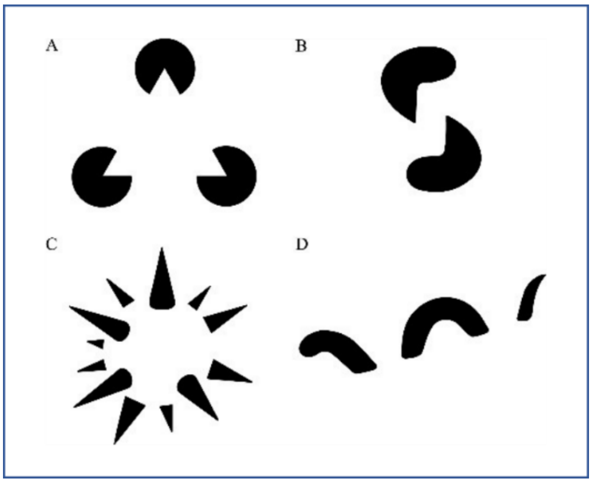


Figure 8. Examples of Kanizsa figures.[2] A: Standard Kanizsa triangle. B: Peter Tse's Volumetric Worm. C: Idesawa's Spiky Sphere. D: Peter Tse's Sea Monster.

Public Domain Infographic. [https://en.wikipedia.org/w/index.php?title=File:Reification.jpg

Recent evidence for predictive coding in the visual cortex comes from a study by Kok and colleagues (2016), who found distinct laminar activation patterns for top-down (feedback) and bottom-up (feedforward) signals. They presented participants with Kanizsa illusion figures (Figure 8) and recorded 7T laminar fMRI. Kanizsa figures allow the examination of neural responses to a shape that is not present but is induced by the surrounding shapes (e.g., the triangle in Figure 8A). Their data revealed enhanced activity in deep layers of V1 regions retinotopically matched to the induced triangle shape. This suggests that expectation units signal the presence of a triangle shape, and those expectations originate from higher-level areas and terminate in deep layers of V1 transferred via feedback connections (compare Figure 2B). In middle and superficial layers, reduced activity was found, which could represent the absence of the bottom-up input (for review of the evidence for similar mechanisms in the auditory cortex, see Heilbron & Chait, 2018).

Similarly, Muckli et al. (2015) found an enhanced signal in superficial layers of V1 that corresponded to an occluded part of a visual scene. They conclude that this enhanced signal corresponds to feedback signals from higher visual areas representing expected visual input in this subsection. Together these studies provide evidence for the dissociation of connection types, based on laminar profiles, between different areas in the hierarchy of the visual processing stream. Moreover, they support the idea of two functionally distinct neural populations that represent either predictions or prediction errors (Figure 7).

So far, predictive coding and underlying key aspects and mechanisms have been described broadly. In the following chapters, specific phenomena that have been extensively studied

---

[2] Kanizsa figures refer to visual stimuli that induce an optical illusion first described by the Italian psychologist Gaetano Kanizsa in 1955. The Kanizsa triangle (A) is known as a *subjective* contour illusion. Stimuli like those became prominent for investigating how the brain perceives visual information (see Lehar, 2003).

in the past and are related to current research are presented and related to predictive processing theories.

1.2.2 Repetition suppression

As repetition suppression is central to the presented studies, this chapter focuses on this phenomenon and its characteristics. I will introduce several theories describing possible neuronal causes for response reduction due to repetitions, comprehending the previously described predictive processing framework as well as models based on intrinsic neuronal mechanisms, and briefly review pieces of evidence for the different explanations.

Neuronal adaptation in the sensory system means the modulation of responsiveness to reoccurring or constantly presented stimuli. It is assumed to result from synaptic plasticity and is a much-investigated phenomenon. Changing the temporal context of a stimulus most likely leads to a change in the neuronal response to that stimulus. When a specific stimulus has occurred before (i.e., it is repeated), a suppressed neuronal response can generally be observed, which is why this phenomenon has been termed repetition suppression (RS; Desimone, 1996). Please note that this reduced neural activity has also been described as stimulus-specific adaptation (SSA; Sobotka & Ringo, 1994), mnemonic filtering (Miller et al., 1991), neural priming (Maccotta & Buckner, 2004), or simply as adaptation (Grill-Spector & Malach, 2001). The RS effect is one of the most extensively studied neural phenomena of the last decades, mainly in visual and auditory perception research. Especially after it was assumed to represent the neural basis of priming (Schacter & Buckner, 1998; Wiggs & Martin, 1998), and special properties of neuronal populations were determined by using fMR-a which is considered a neuroimaging manifestation of RS (Grill-Spector & Malach, 2001; but see Kar & Krekelberg, 2016; Larsson et al., 2016), several possible underlying mechanisms (chapters 1.2.2.1 and 1.2.2.2) were discussed (for a review, see Grill-Spector, Henson, & Martin, 2006).

In general, RS was found to be long-lasting already in 1993 (Li et al., 1993), as effects of repetition were evident after several hours on the neural responses. Further, RS was characterized as being stimulus-specific and surviving a large number of intervening stimuli (Li et al., 1993; Rangarajan et al., 2020). However, results from an fMRI study suggest different mechanisms underlying short- and long-interval repetition effects (Epstein et al., 2008). RS is assumed to be a largely automatic phenomenon (Kouider et al., 2009; Sayres & Grill-Spector, 2006) that is still present when observers attention was diverted (Larsson & Smith, 2012).

RS has been described as a simple mechanism representing a measure of processing efficiency that is fundamental to the processing of our daily visual input and was early associated with sensory learning in its simplest form (Groves & Thompson, 1970). It can also be described as an experience-related rapid form of plasticity. It is hypothesized that RS serves to reallocate attentional resources efficiently, as previously seen stimuli lose saliency and new or changed stimuli are emphasized (Kaliukhovich et al., 2013). However, the underlying neuronal mechanisms are still not sufficiently explained, and different mechanisms likely contribute to phenomena of response suppression.

1.2.2.1 Feedforward models of RS

The possible neural mechanisms underlying RS are manifold. Hence, theories aiming at explaining measurable neuronal effects related to repetition are versatile. In this subchapter, those relying on a primary role of bottom-up input processing and inherited neuronal mechanisms will be introduced (Figure 9).

One early explanation for RS was offered by the *Fatigue model* (Miller & Desimone, 1994), which states that all initially responding neurons respond proportionally less to a repetition of the same stimulus. In other words, all neurons responding to the first stimulus (S1) will be responsive again, but the average firing rate will be declined, as shown in Figure 9 – Fatigue model. Accordingly, firing rate adaptation (Carandini & Ferster, 1997) is assumed to cause RS effects. However, synaptic depression (reduction in synaptic efficiency) or inherited adaptation effects are alternative explanations for the proposed neuronal behavior in this model (Vogels, 2016). Recent studies on inferotemporal (IT) and MT neurons behavior show that simple neuron fatigue does not adequately explain RS effects (Fabbrini et al., 2019; Kar & Krekelberg, 2016).

The *Sharpening model* (Desimone, 1996) suggests narrower neural tuning curves and therefore sharpened representations of repeated stimulation (Figure 9 – Sharpening model). In terms of metabolic processes, the resulting sparser representations use less energy and are therefore more efficient. In contrast to fatigue models, neurons that show optimal response to the stimulus are less affected by a repeated activation and show a similar activation level. This model is related to the sharpening of tuning curves and predicts selectivity increases with repetition. Those hypotheses could not be proven in investigations of macaque IT neurons spiking activity and local field potentials after stimulus repetition (De Baene & Vogels, 2010). Also, a recent study on human fMRI responses to face stimuli repetitions also provided limited evidence for this model (Alink et al., 2018).

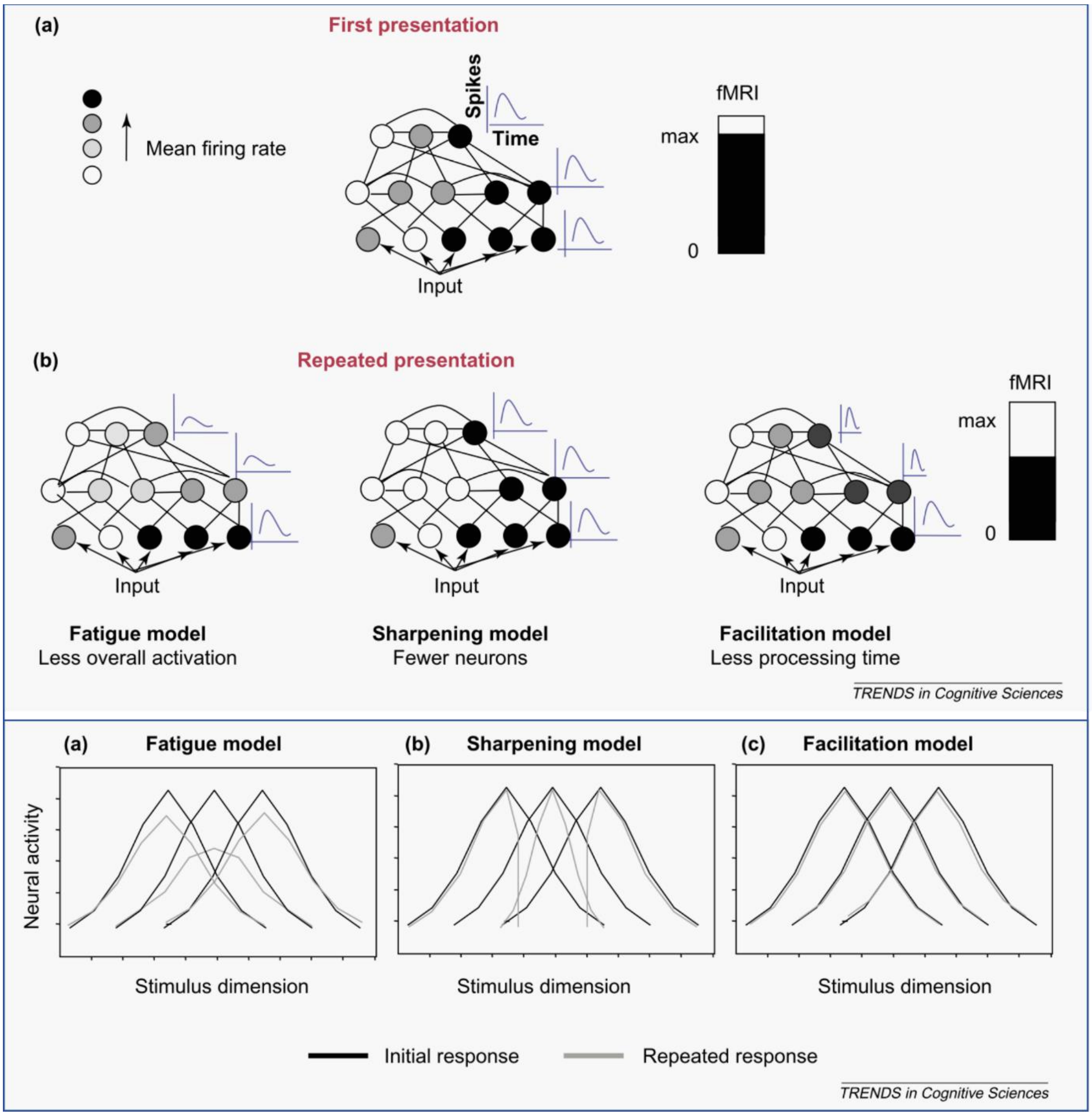


Figure 9. Firing rate of a neuronal population (upper panel) and neuronal activity (lower panel) to repeated stimulus presentations as proposed by different models. Upper panel: Mean firing rate and spiking (blue curves) of neurons are depicted for (a) first stimulus presentation and (b) repeated stimulus presentation for the different models. Lower panel: tuning curves to the first (black curves) and repeated (grey curves) stimulus presentation as predicted by the different models: (a) Fatigue model (b) Sharpening model (c) Facilitation model. Adapted from Grill-Spector et al., 2006 with permission.

The *Facilitation model* predicts faster processing for repeated stimuli (Figure 9 – Facilitation model). This could take the form of shorter latencies, more precisely faster response onset latency or earlier peak time (James & Gauthier, 2006), or shorter durations (Henson, 2016; Henson & Rugg, 2003). Evidence for this model comes from an fMRI study (Henson et al., 2002). The authors found not only a reduced magnitude in the fusiform BOLD response to a repetition of a face but a reduced latency of this response as well. Support for an extended version of this model, the so-called accumulation model (James & Gauthier, 2006), comes from an electrocorticography (ECoG) study in which RS effects for up to six stimulus presentations were investigated (Rangarajan et al., 2020). The authors found faster peak times for the second and subsequent presentations of a face image in face-selective units. However, this paradigm examined long-lagged RS effects. Intracranial recordings of macaques MT (Kar & Krekelberg, 2016) and IT neurons (De Baene & Vogels, 2010) found no support for the facilitation model for immediate repetitions.

The *synchronization model* (Gotts et al., 2012) states that neuronal processing achieves higher efficiency due to enhanced synchronization (Figure 10). Although firing rates generally decrease with repetition, their synchrony leads to sharper neuronal responses. This model provided a reasonable explanation for RS as it mediates between a neuronal response reduction and enhanced behavioral performance, hence repetition priming (Henson, 2003). However, findings of long-lagged repetition suppression (Rangarajan et al., 2020) are difficult to explain within this model.

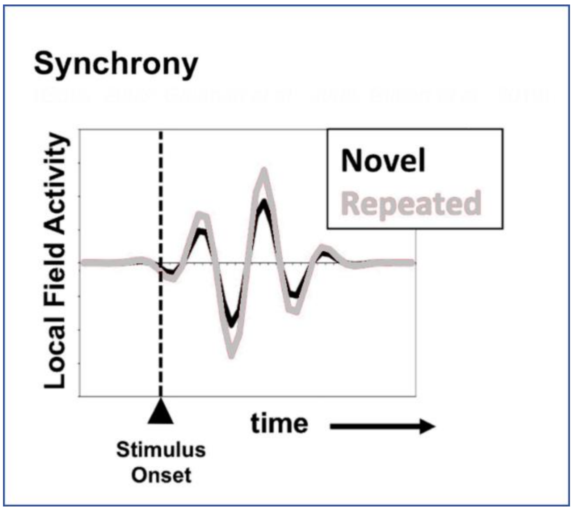


Figure 10. Local field activity to first and second stimulus presentation as predicted by synchronization models. Adapted from Gotts et al. (2012), with permission.

Recently, the importance of more refined models, including neuronal circuit processes, has been emphasized (Bastos et al., 2012; Whitmire & Stanley, 2016). Intracranial recordings from macaque monkeys MT during a repetition paradigm revealed that the underlying mechanisms are complex and cannot be accounted for by simple models such as fatigue or facilitation (Kar & Krekelberg, 2016). Therefore, adaptation processes on a single-neuron level seem unlikely, and neuronal circuit computations provide a better explanation. Models based on neuronal circuit computations are implemented in the PC framework (Bastos et al., 2012). However, measurements of RS in inferotemporal neurons (IT) suggest the differentiation of adaptation mechanisms from top-down influences suggested by predictive coding models (Vogels, 2016). Results from univariate and multivariate human fMRI analyses for RS to face stimuli also correlated best with a local neuronal scaling model of RS (Alink et al., 2018). However, none of these models for mechanisms behind response suppression to repeated stimuli can explain all observed RS effects. Therefore, it is more likely that a combination of mechanisms applies. As response suppression over repetitions was related to perceptual learning and inference processes, and therefore explained within the predictive coding framework (Auksztulewicz & Friston, 2016), the next chapter will further describe how RS can be explained within dynamic hierarchical systems.

In addition, different models make predictions about different RS characteristics. Most of them predict changes in the amplitude of the response to repetitions (e.g., sharpening and fatigue models). Still, some make predictions about the timing of responses to repetitions (facilitation model). Therefore, it is essential to think about what methods are useful to differentiate between these models. A reduced BOLD signal is proposed for all of them (Figure 9). Therefore, fMRI experiments can only distinguish between different models to a limited extent (Weiner et al., 2010). However, the influence of higher-order feedback information as proposed by PC explanations can be measured in the activity of neuronal populations. Some studies showing such effects will be introduced in the next chapter.

1.2.2.2 Repetition suppression explained as the result of top-down processes

As mentioned before the previously introduced models cannot explain all measures of RS under specific conditions. Thus, it has been suggested that RS could also be an expression of a reduced prediction error. The first and very influential study hinting towards a top-down component in RS effects as measured in fMRI experiments came from Summerfield and colleagues (2008). The authors of this study presented pairs of faces that could either be identical (repetition trials; Rep) or different (alternation trials; Alt) and measured the BOLD signal in the bilateral FFA while participants executed a behavioral task on target trials. The new and – as the next years were about to show – insightful trick in their design was to group trails in repetition and alternation blocks and vary the proportion of trial types, respectively. In repetition blocks, 75% of non-target trials were Rep and 25% Alt trials – and vice versa for Alt blocks (compare Figure 12A). This way, participant's expectations about encountering a specific trial type were manipulated. The results - significant RS for both block types but also a modulation of RS by repetition probability (P_{REP}) - were difficult to explain with most theories available at that timepoint. The P_{REP} effect was replicated in several following neuroimaging studies and further characterized, for example, as being dependent on experience with the stimulus material (Grotheer & Kovacs, 2014) and position invariant (Kovács et al., 2012). On the basis of evidence for this effect, the authors argued that RS might be the product of perceptual inference and top-down modulations by expectations (Summerfield et al., 2008). In other words, the higher expectation of encountering a stimulus repetition in Rep blocks leads to a bigger RS magnitude in those blocks in comparison to Alt blocks, in which repetitions are surprising. In an EEG study, those results were replicated for later event-related potentials, around 300 ms after stimulus onset, which showed an expectation modulation effect in central electrodes (Summerfield et al., 2011). In this study, also environmental volatility was investigated and found to modulate the expectation modulation of RS. Whereas in stable periods the P_{REP} effect was found reliably, in the volatile context, in which probabilities of repetition and alternation trials changed frequently, this effect was absent. This hints towards updating feedback mechanisms from higher-order cortical areas because the surprising events (e.g., repetitions trials in an alternation block) are weighted according to their informative value. A Bayesian model for response suppression phenomena is depicted in Figure 11.

Kaliukhovich and Vogels (2011) tried to replicate the P_{REP} effect in a single-cell study. They used objects and fractals in a repetition design where repetitions could be expected or not (similar to Summerfield et al., 2008) and recorded macaques IT neuron activity. Robust RS effects were found, but no modulatory effect of expectation was present for either spiking or local field potentials. Other fMRI studies failed to replicate the P_{REP} effect for faces even though a behavioral priming effect of repetition probability was present (Olkkonen et al., 2017), for everyday objects in LO (Kovacs et al., 2013), and for other non-face or unfamiliar stimuli (Grotheer & Kovacs, 2014). Furthermore, there is evidence that P_{REP} effects, but not RS, vanish if the participant's attention is diverted from the stimuli (Larsson & Smith, 2012). However, in an EEG study, RS was investigated in a priming paradigm and found to be modulated by the effect of expectation, even when participants were not consciously aware of the prime (Barbosa & Kouider, 2018).

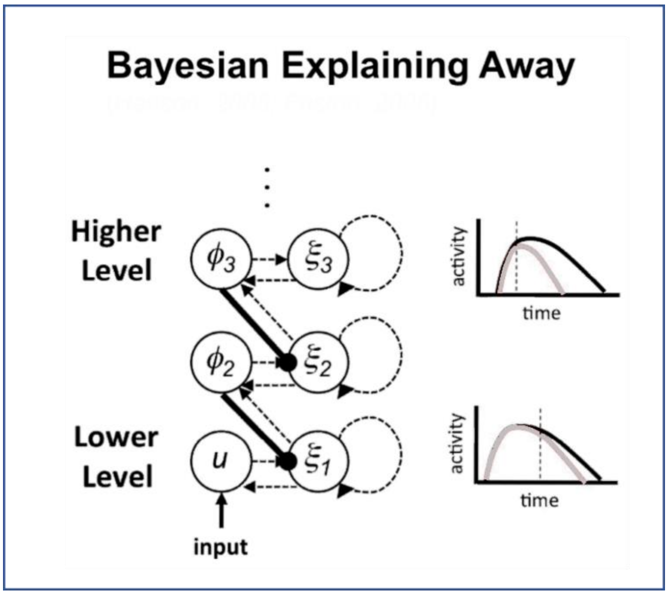


Figure 11. Bayesian Explaining Away. Hypothetical novel and repeated conditions are shown with black and grey curves, respectively. Suppressive feedback from higher levels to lower levels in the network structure is highlighted using thick black lines. The earlier separation of novel and repeated conditions in higher levels relative to lower levels is indicated with vertical dashed lines in the activity plots to the right. Adapted from Gotts et al. (2012), with permission.

The Bayesian explanation, hence predictive processing, for information processing in the cortical hierarchy has already been explained (see chapter 1.2.1.2). It is clear that higher-order expectations can modulate RS at least in some studies (e.g., Grotheer & Kovacs, 2014; Summerfield et al., 2011), but how can RS itself be explained by predictive theories? Grotheer and Kovács (2016) already asked this question and gave an overview on response suppression phenomena. They developed a two-stage model of response suppression to explain those effects including lower and higher-level expectations. However, the question remained unanswered, as mainly effects of expectation suppression, which will be explained in the next chapter, are accounted for by this model. RS itself fits within the framework of predictive coding when considered as an example of expectation-based response suppression (see chapter 1.2.3). Repetition of the same stimulus makes the prediction and the incoming sensory data more and more similar, and consequently improves predictions (Grill-Spector et al., 2006). Recently encountered stimuli serve as the basis for expectations, which in turn are weighted towards the already observed input. RS is therefore assumed to reflect the reduced prediction error when those expectations are fulfilled (Auksztulewicz & Friston, 2016). Within the theories of predictive processing, repetition is the default prior. This also refers to our outside world, which is stable on a short-term scale and highly interdependent in space and time (Dong & Atick, 1995). Measures of RS are accordingly prediction error signals shaped by expectations (Friston, 2005).

To sum it up, measures of RS are likely the outcome of several different mechanisms that are also likely to interact (Alink et al., 2018; Grotheer & Kovács, 2016). Forward models, such as facilitation or sharpening, can account for different measures of repetition effects. Critically, none of the proposed models can account for all measures of RS. However, their proposed neuronal mechanisms, such as sharpening of representations (Kok et al., 2012a), or synchronization of neuronal firing rates, can be integrated into the predictive coding framework (Auksztulewicz & Friston, 2016; Kok et al., 2012a). In light of an explanation for repetition effects within the PC framework, it is still crucial to consider inherited adaptation effects. Such effects from lower-level areas, for example, caused by synaptic depression, still need to be considered as contributing to measures of RS (Kohn, 2007; Vogels, 2016). However, they can be integrated into local circuit-based computations, in which RS can occur without modulatory influence from higher-level areas and therefore fit in with PC explanations of RS (Auksztulewicz & Friston, 2016; Bastos et al., 2012).

1.2.3 Expectation suppression, as a separate phenomenon

In their fMRI study, Summerfield and colleagues (2008) observed that the magnitude of RS changes by means of manipulating the probability of repetition events in different blocks. This modulation of RS due to high versus low expectation of encountering a repetition led to the idea that RS or its modulation could be an outcome of predictive processes. However, findings regarding the P_{REP} effect were inconsistent (e.g., Kovacs et al., 2013; Vinken et al., 2018). Shortly afterward, the term expectation suppression (ES) was coined. ES, related to RS, describes the suppressed response to a stimulus that is more likely to occur than another (neutral or surprising) stimulus. In other words, it is the difference between an expected event versus an unexpected one, regardless of whether the two events are identical or different. Predictive coding is a prominent theory to explain these kinds of effects (Grotheer & Kovács, 2016). Many studies investigating ES have shown that the degree to which a stimulus is expected is expressed in the magnitude of stimulus-evoked responses (Bendixen et al., 2009; Robinson et al., 2018; Wacongne et al., 2011). This inversely scaled relationship between expectations and response magnitude is assumed to constitute a key hypothesis of predictive processing (Walsh et al., 2020).

In an influential MEG study, Todorovic and de Lange (2012) investigated expectation suppression and repetition suppression and the relationship between these two phenomena. As both reveal similar results, that is to say, a reduced response, the authors tried to distinguish them in an auditory cue-target experiment. Each trial consisted of two tones, that could constitute either a repetition or a stimulus change. Additionally, the leading tone signaled whether the following tone will be more likely identical to or different from the first one. Their results showed that RS and ES could be differentiated on the basis of their temporal dynamics. RS effects were found in earlier time windows, whereas ES was present only later. As no interaction between RS and ES was found, Todorovic and de Lange (2012) suggested the independence of the two. Similar results were found in ERP data to repeated versus expected face stimuli (Feuerriegel et al., 2018a). These findings are further supported by an fMRI study in the visual domain. Grotheer and Kovács (2015) used pairs of face images, that similar to the previously described auditory experiment could be identical or different within one trial. The gender of the face served as a cue for the probability of encountering a repetition of the leading image or a transition to another face (comparable to high and low repetition probability in Figure 12B). Again, no interaction of RS and ES was found, and the two were therefore considered to be independent mechanisms. The driving mechanisms for this effect are still not fully explained (for a review, see Feuerriegel et al., 2021a). One possible explanation for expectation suppression is the suppressed response for correct predictions or fulfilled expectations. Another possible explanation is that violated expectations function as a surprising event in the system. Therefore, and to further extend the results of Grotheer and Kovács (2015), Amado et al. (2016) implemented a neutral condition in the previously described design, in which an infant's face signaled equal probability of repetition (comparable to middle repetition probability in Figure 12B). This way, they revealed that the surprise of unexpected events determines the neural response more than it can be suppressed by precise predictions. This issue is revisited in chapters 1.2.5 and 1.2.6.

Contradictory to RS, ES does need, as the name already suggests, prior expectations about upcoming events to generate predictions. How exactly expectations are computed and implemented on a neuronal level is still unclear (de Lange et al., 2018). In experiments, however, expectations are assumed to evolve based on a cue (cue-based; Figure12B) or overall probability (probability-based; Figure12A) that is assigned to certain events in a context (block). In each case, expectations need to be learned through extensive exposure to specific statistical relationships (Meyer & Olson, 2011; Schwiedrzik & Freiwald, 2017) or rely on explicit knowledge (Amado et al., 2016; Egner et al., 2010; Grotheer & Kovács, 2015). For example, Egner et al. (2010) informed participants in an fMRI experiment about the probabilities (high, middle, and low) with which a color frame is associated with either faces or houses. The more a face was expected, the lower the elicited BOLD response in the FFA was. Similar effects of cue-based predictions about stimulus content were found in an audio-visual learning paradigm (Den Ouden et al., 2010). The authors used auditory cues associated with either faces or houses and found expectation-related response modulation in the FFA for faces and the parahippocampal place area (PPA) for houses.

The cues used in the previously described studies signaled stimulus repetition or stimulus class (e.g., face or house). In a monkey study, Meyer and Olson (2011) implemented a different kind of cueing paradigm. The macaque monkeys were trained with image pairs, so specific images were associated. Thus, the monkeys could predict the second image of one trial based on the leading image. If these predictions were violated in a test run with simultaneous IT neurons recording, those neurons showed stronger responses compared to fulfilled image transition expectations. This kind of design is also known as statistical learning of transition probabilities and will be addressed again in chapter 1.2.5 and in Study II (chapter 3.2).

Attention is often a candidate for explaining contradictory results in PC studies, as the role it plays in predictive processes is not fully understood and might often be confounded with expectation-related effects (Feuerriegel et al., 2021a). For example, accurately predicting sensory input is assumed to lower the neuronal response to that stimulus because the prediction error is smaller (de Lange et al., 2018). But there is evidence for the opposite effect as well (Henson et al., 2000; Kok et al., 2012b). Lately, it has been suggested that attention may account for a considerable amount of results that were labelled as expectation suppression effects (Alink & Blank, 2021; Feuerriegel et al., 2021a)

1.2.3.1 Different manipulations of predictions about stimuli

Different paradigms to test for PC assumptions of response alterations are schematically illustrated in Figure 12. In the literature, probability-based expectations refer primarily to an experimental situation, in which probabilities of encountering a specific stimulus or trial-type in a given context can be implicitly learned. Such contexts can be repetition or alternation blocks similar to the design in Summerfield et al. (2008) and as depicted in Figure 12A. We did not implement such a design in the current studies.

Cue-based expectations, on the other hand, refer to stimulus combinations in which S1 signals the likelihood of a specific characteristic of the second stimulus. Studies investigating RS and ES with a cueing design often use a mixture of cue-based and probability-based manipulations, as depicted in Figure 12B. This kind of design was used in the study by Todorovic and de Lange (2012), who showed that RS and ES are separable in early and later time windows, respectively. Also, Amado and colleagues (2016) combined a cue (sex of presented face) with a probability manipulation regarding the likelihood of encountering a repetition of the stimulus. An example for a cue-based expectation manipulation without changing the probability of repetitions of stimuli is the study by Egner and colleagues (2010), in which a color frame cue signaled the stimulus category ‘face’ with low, middle, and high probability. The cueing of an exact image is more difficult because it needs extensive training (Meyer & Olson, 2011).

Figure 12C depicts two examples of an oddball design, which will be further explained in chapter 1.2.6. All designs in Figure 12 are examples of designs previously used to test for prediction-related response alterations. For a more detailed overview of designs used to investigate prediction-related effects, see Feuerriegel (2016)and Feuerriegel et al. (2021a). We used variations of a cue-based probability manipulation (Figure 12B) and an oddball design (Figure 12C) in Study II and III. The exact paradigms will be explained in the related manuscripts (chapter 3), and an overview is depicted in the discussion (chapter 4.1.4).

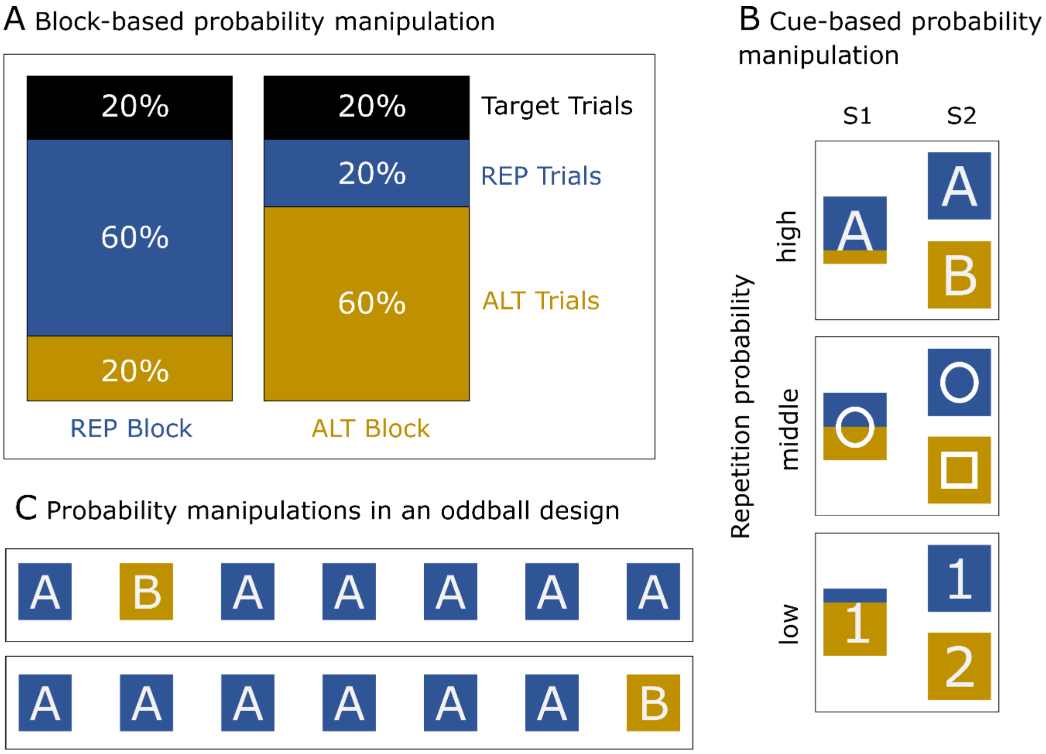


Figure 12. Examples for experimental designs investigating probability manipulations. Blue = repetitions, Gold = alternations. A Block-based probability manipulation: blocks of trials are characterized by the probability of stimulus repetition. B Cue-based probability manipulation: the first stimulus (S1) signals high (letter), middle (symbol), or low (number) probability of repetition. C Probability manipulations in an oddball design: sequences of stimuli are composed of standards (A) and deviants (B).

1.2.4 The role of precision in predictive coding

The weighting of prediction errors according to their precision is an important estimation of their reliability (Auksztulewicz & Friston, 2016). For every PE that is generated, the information on how much a prior should be updated according to this error is necessary and at least equally important. A mismatch between expectations and the sensory input could result from a wrong prediction but just as well from noisy sensory data. In other words, PEs can have different functional roles in decision-making, learning, perceptual inference, or cognition in general (Den Ouden et al., 2012). Therefore, an inferential system can achieve effectiveness only if this is taken into account. Logically, prior knowledge and new observations are combined in relation to their respective precision. One could also refer to this concept as having first-order predictions about incoming signals and second-order estimations about their precision. In short, sensory input is the sum of predictions and noise.

Predictions, based on prior information, as well as sensory input can be visualized as Gaussian probability distributions. Graphs charting probability distributions of estimations are termed probability density functions (PDFs). Figure 13 shows examples for PDFs representing prior beliefs and sensory inputs, where the curve demonstrates all possible states. Uncertainty can therefore be modeled as the variance of a distribution and is represented by the width in PDFs. The wider the distribution, the less precise the prior or the sensory signal is. For better understanding, I will give an example of the precision of priors and sensory data.

Imagine you are visiting New York and being surrounded by skyscrapers. You are trying to figure out the height of a specific building, but you have very little prior knowledge about them. This would refer to uncertainty or low precision in priors (Figure 13C). Hence, the posterior belief or prediction, is based more on the observation. Now imagine your friend grew up in New York and has more knowledge about the average height of buildings through long-term experience processes. In this case, the prior belief is more precise, although, of course, there is still some uncertainty in it. The resulting estimation about a specific building now depends on the sensory input, in other words, how the two sources of information are weighted. If our visual system were free of noise, the corresponding PDFs (sensory input in Figure 13) would be narrow and appear like vertical lines. However, in our daily life, we always have to deal with noise. The noise or uncertainty about the sensory input could come from the visual signal process itself or deception or obfuscation in the visual signal. Imagine your friend tries to estimate the building's height in front of you, but it is raining and foggy and, therefore, difficult to see its top and maybe compare it to surrounding buildings. This would refer to low precision in sensory input, and the prediction is rather based on previous experiences as they are more precise as depicted in Figure 13B. The posterior beliefs are calculated as the combination of previous knowledge and the observed input. Hence, Posterior beliefs are an estimation between the two PDFs. They are illustrated as dashed lines in Figure 13but can be depicted as a third PDF, whose precision is a sum of priors and sensory input precision. The exact location of these PDFs is also determined by the precision of both sources of information. Figure 13A illustrates the posterior belief right between the prior prediction and the sensory input, as their precision is equal. Depending on which source is more reliable, the posterior belief is shifted in that direction. Notably, the precision of prior beliefs plays a role in how much they themselves are updated. If a prior is imprecise, it will be changed to a greater extent than when it is already very precise itself (Auksztulewicz & Friston, 2016). The combination of priors and sensory input to calculate the posterior distribution will be updated by the following observation logically. Then, the previously posterior PDF will serve as the new prior. These processes are the basic principle for learning.

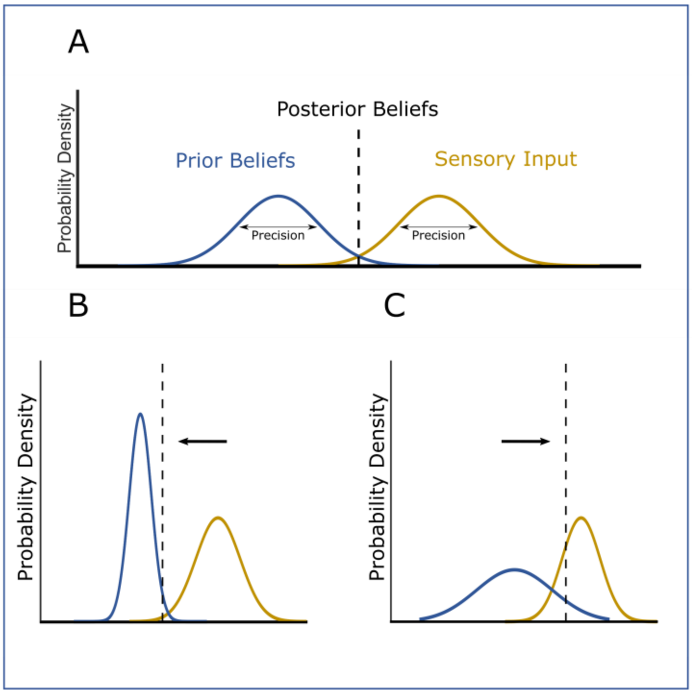


Figure 13. Schematic illustration of PDFs characterized by precision. Y-axis is probability density. Blue graphs represent prior beliefs, gold graphs represent sensory input, and dashed lines represent the posterior belief. Precision is defined as the width of the graphs. A. Precision in prior beliefs and sensory input is equal. B and C with equal precision in sensory input, posterior beliefs will be shifted towards the more precise prior belief in B and towards the sensory input as prior is less precise in C.

Precision in PC is strongly connected to attention, as attention can be seen as an outcome of precision weighting (Feldman & Friston, 2010). Imagine a simple visual attention task, like a Posner task (Posner et al., 1980) in which specific locations are cued. Targets are then expected to occur at the cued position, which can be interpreted as a prior belief about its spatial position. The general assumption is that attention will be allocated accordingly to reduce uncertainty in the system (Rao, 2005) which was supported by a neuroimaging study by Jiang et al. (2013).

The importance of well-functioning precision-weighting mechanisms becomes clear when we look at what happens when they are disrupted or out of balance. Adams et al. (2013; see also Adams et al., 2014) explained psychotic symptoms such as hallucinosis or delusions as expressions of false inference processes. According to their proposed explanation, the optimal integration of prior beliefs and sensory data is impaired because the relative precision of both sources of information is abnormally encoded. For example, overweighting sensory input due to depletion of precision of prior beliefs could lead to attention being drawn to irrelevant sensory events. The resulting feeling of something is odd can be compared to a delusionary state. The exact neurocomputational mechanisms underlying psychosis remain to be discovered. Still, they are now thought to be related to maladaptive inferences, and thus predictive coding provides insights into this multifaceted disorder (for a review, see Sterzer et al., 2018).

The previously described effect of P_{REP} can be explained by changes in precision as well (Summerfield et al., 2011). Effects of repetition can be modulated by the precision of sensory predictions (Auksztulewicz & Friston, 2016). Imagine a situation where recurrence or stability of sensory input is unusual. For example, going up in the lift in a high-rise building, one expects the floor display to change every few seconds. Sensory input that is at odds, in this case, would therefore be estimated as unreliably. In this manner, precision of prior beliefs can be modulated by perceptual expectations, such as developed in block designs (Figure 12A). Moreover, a volatile environment such as a trial sequence in which the probability of repetition changes every now and then would reduce the precision of prior beliefs (Auksztulewicz & Friston, 2016). This relationship allows further testing of the underlying mechanisms of RS as prediction errors are assumed be exogenously changed based on the reliability of sensory input. We address this hypothesis in Study I (chapter 3.1).

1.2.5 The role of stimulus novelty and surprise in predictive coding

Novel stimuli evoke greater neuronal responses compared to recently encountered or well-known stimuli (Li et al., 1993). The same also applies to surprising stimuli (Amado et al., 2016; Bunzeck & Thiel, 2016). Therefore, both must be considered as properties of stimuli in experimental designs investigating effects as measures of response amplitudes. In this chapter, I will briefly introduce pieces of evidence for the impact of stimulus predictability in the light of novelty and surprise[3].

Already Summerfield, in his influential study on the influence of P_{REP} on RS (Summerfield et al., 2008), stated that RS in Alt blocks (in which repetitions were unexpected) was not entirely abolished. Although a block with a high probability of alternations was presented, the authors suggested that encountering a repetition of the same face was still more expected than a specific alternated face. Imaging a stimulus pair that shows identical images being compared to a stimulus pair in which the second image is drawn from a large set of images, as in most previous studies (Amado et al., 2016; Grotheer & Kovács, 2015; Summerfield et al., 2008). Effects of novelty and familiarity can blend with stimulus repetition effects. Depending on the method with which data is acquired, the novelty of one stimulus could result in a larger RS effect and act as a confounding factor in repetition designs (for a review on repetition paradigms, see Feuerriegel, 2016). This aspect has not been addressed in many experimental designs aiming at investigating RS and ES. Whereas repetition always offers the prediction of the second image on the basis of the first in a detailed manner, alternating images are not predictable. It is important to differentiate studies in which the outcome of a trial (being a repetition or a change of a stimulus) can be predicted quite reliably (Amado et al. 2016) from studies in which the second image of a trial can be not only expected, but predicted with high precision as in an EEG experiment by Feuerriegel et al. (2018) or fMRI experiments, such as Pajani et al., (2017). This is the primary topic of Study II (3.2). The term predictability is therefore defined as the capacity to predict the image itself in an experimental design investigating RS or ES. In general, predictable sensory input results in less neuronal activation than unpredictable. This is in line with theories of predictive processing and was first shown for early visual areas (V1) (Alink et al., 2010). When onset or orientation of upcoming stimuli could be predicted from spatiotemporal context information, the stimulus-related BOLD response was smaller than for unpredictable stimuli. This result shows the importance of predictability at early stages of cortical processing. However, the prediction of a position or orientation of a stimulus is far away from predicting a complex stimulus such as a face. Additionally, in this study, the

---

[3] The term surprise in studies investigating phenomena such as RS and interpreting related results in the predictive coding framework is used to describe unpredictable sensory input. This can only be compared to a limited extent with the term as it is used in everyday life. Minimizing surprise in predictive processes to reduce free energy cannot be equated with everyday surprises. Avoiding surprises seems to be something unusual for a species that seeks novelty, likes to explore, and avoids boredom. However, according to PC theories, it is necessary to minimize the average of surprise on a long-term scale to achieve an overall model that allows explorative engagement with the world (Clark, 2018; Schwartenbeck et al., 2013).

predictability of the investigated stimulus features (orientation and position/timing) depended on the trajectory information within each trial.

To predict an image in an alternating trial within a repetition paradigm, only the first image of the trial can be informative. Naturally, training is needed to familiarize participants with specific stimulus combinations. Meyer and Olson (2011) showed an effect of implicitly learned image transition rules in IT neuron activity recorded in two monkeys as a lower response when associated images were shown as trailing image compared to other unpaired images. This unsupervised form of learning transition rules by mere exposure to statistical contingencies is mainly known from learning your first language as a child (Trainor, 2012). In an fMRI experiment on this issue, Pajani et al. (2017) introduced prediction blocks additionally to repetition and alternation blocks. Rep and Alt blocks were similar to those in previous designs (Grotheer & Kovacs, 2014; Summerfield et al., 2008). Prediction blocks were identical to Alt blocks in terms of probability manipulations. Still, they differed as participants could precisely predict alternating images. The authors trained participants on specific face stimulus combinations in an initial behavioral session. This resulted in a much smaller RS magnitude in prediction blocks due to a reduced signal to alternating stimuli compared to alternating stimuli in Rep and Alt blocks. The authors termed this effect exemplar-specific expectation suppression. A similar effect was found in an EEG study, in which RS and ES were investigated (Feuerriegel et al., 2018a). In Study II we investigated the effect of stimulus predictability on measures of RS with an adapted version of the design used by Feuerriegel et al. (2018a) to control for stimulus novelty effects.

1.2.6 Mismatch responses under the predictive coding framework

The mismatch negativity (MMN; Näätänen, 1992) was initially found in the auditory domain. The MMN describes a negative EEG component observed for a deviant stimulus occurring in a series of standards or, in other words, a mismatch to preceding stimuli. This response alteration has been explained as the outcome of a memory-based mechanism (Winkler & Czigler, 1998) or adaptation of neurons responding to standard stimuli (May & Tiitinen, 2010). The visual mismatch negativity (vMMN) is the same phenomenon observed in the visual domain accordingly (for reviews, see Czigler, 2007; Pazo-Alvarez et al., 2003). In the following, I will refer to this and similar effects as (visual) mismatch responses ((v)MMR), as the description of a negative ERP does not apply to responses measured with fMRI. Designs that are used to investigate this phenomenon are called oddball designs.

Typically, two stimuli are used that serve as standard and deviant within one experiment. In other words, stimulus **A** can be presented in one context with a high probability and, therefore, is the standard and a rarely occurring stimulus **B** is the deviation accordingly (Figure 12). Usually, these probability assignments are then reversed for another context or block. The difference in responses to the same stimulus (e.g., stimulus **A**) when presented as a standard versus deviant is called the oddball effect. The effect describes a context-dependent modulation since the visual features of the stimuli are the same. It can also be described as stimulus-specific adaptation effect and interpreted in the PC framework. Connected to hierarchical inference processes and learning, MMRs are assumed to reflect error detection to surprising or unexpected events (Baldeweg, 2007; Garrido et al., 2009a) after adaptation over multiple repetitions, which improves prediction (Grill-Spector et al., 2006). This idea is further supported by findings showing that the omission of an expected stimulus results in a mismatch response (Bendixen et al., 2009; Wacongne et al., 2011). For a comprehensive review of possible underlying mechanisms inclusive processes related to PC, see Garrido et al. (2009).

Effects of repetition can account for part of mismatch responses, as well enhancement effects due to unexpectedness of deviant stimuli. As mismatch responses were mainly investigated with electrophysiological techniques, their relationship to fMR-a was only studied recently. Amado and colleagues (2018) applied oddball sequences of characters in EEG and fMRI experiments and acquired data from the same participants. They found the observed vMMN and fMRI adaptation to correlate in cortical regions processing characters.

A related paradigm has also been used to investigate expectations about higher-level stimulus features. In an EEG experiment, instead of using the same image of a famous identity, different so-called ambient images were shown in a stream of images (rapid periodic stimulation paradigm) with a rarely occurring (deviant) different same-sex identity face (Johnston et al., 2016). The face-sensitive N170 amplitude was modulated by the probability of occurrence of one identity. The authors concluded that expectations about person identity led to changes in the N170 amplitude when violated, signaling prediction error responses. In other words, a response to a mismatch in identity is measurable in early event-related potentials.

Furthermore, similar designs were used to investigate prediction error responses by stimulus presentations in which a given trajectory was maintained or violated (Robinson et al., 2020, 2018). This allows for testing prediction-related responses without repeating a specific stimulus. However, the contribution of adaptation and surprise effects to MMR is still not fully clarified (Feuerriegel et al., 2018b). A neutral condition was added to an oddball design in a recent EEG study, allowing to better classify observed effects as expectation suppression to repeated stimuli or surprise response to deviant stimuli (Feuerriegel et al., 2021b). The authors state that mismatch responses to visual stimuli are best explained by a surprise response, which contrasts findings from single-unit recordings of macaque IT neurons (Kaliukhovich & Vogels, 2014) but stands in agreement with studies on P_{REP} (Amado et al., 2016).

In Study III, we applied a simple version of the design depicted in Figure 12C and investigated adaptation and expectation effects to ambient images of famous identities (chapter 3.3).

So, in summary, predictive coding models constitute a powerful framework that can explain phenomena in visual perception like bistable perception, perceptual illusions, neural response patterns to violated expectations, or surprising input as in an MMN design (for a review, see Walsh et al., 2020). Also, predictive processing as a strategy seems to fit perfectly to the way our cortex is structured. Still, the predictive coding framework is lacking direct empirical evidence (Walsh et al., 2020) because neurophysiological results are often contradictory or can be explained by other models of perception. Because the current studies were all conducted in the visual domain, most of the literature that has been summarized refers largely to studies that examined the visual system.

Different hypotheses about phenomena explained (among others) within the predictive processing framework are tested in the present work. As we used face stimulus sets in all current experiments, the next chapter will introduce some important aspects about the relationship between face perception and predictive processing.

1.3 Face perception and predictive processing

Theories of predictive coding have explained processes on low-level visual phenomena, even going back to cellular-level processing in the retina (Hosoya et al., 2005), until higher-level phenomena like binocular rivalry (Hohwy et al., 2008). Also, face-stimuli have been applied extensively to investigate neural phenomena, such as RS (for reviews on neuroimaging and EEG studies, see Henson, 2016; Schweinberger & Neumann, 2016) and ES, suspected to display processes of PC (Feuerriegel et al., 2018b; Grotheer & Kovács, 2015; Summerfield et al., 2008). RS effects for faces measured with fMRI highly correlate with face perception ability, implying a role of response reduction mechanisms in face perception processes (Hermann et al., 2017). Still, the connection between models of the face perception network and predictive coding theories has been enlightened only lately (Trapp et al., 2018). In the last chapters, I have only touched on why faces are very well suited for the investigation of prediction-related processes. Here I will explain some aspects in a detailed manner.

First of all, humans are face experts because processing facial information is essential from the very beginning of our lives. PC models suggest an internal model of the world that is continuously updated to predict sensory input (Rao & Ballard, 1999). Logically, the more experience we can draw on, the better our predictions should be. MEG recordings by Brodski and colleagues (2015) provide an insightful example of how lifelong experiences with stimuli, such as faces, can be used to study predictive mechanisms. They asked participants to perform a Mooney (Figure 4) face-detection task and manipulated orientation and illumination of the stimuli. The priors based on lifelong experiences in this study were the upright orientation for faces and a top-down direction for illumination, accordingly. Increased gamma-band activity is assumed to reflect increased prediction errors. Indeed, the authors found signals of violating priors of face orientation in occipital and fusiform regions, whereas violating illumination priors led to increased signals in the medial frontal cortex, superior frontal gyrus, and anterior cingulate gyrus.

Second, the visual system is hierarchically organized, and so is the face-processing network (see chapters 1.1.1 and 1.1.2). This characteristic is an essential condition for predictive processing. Besides, areas belonging to the face-processing network are highly interconnected (Nagy et al., 2012; Pyles et al., 2013). Therefore, top-down influences are assumed to play an important role in face recognition. For example, the OFA was considered to constitute one of the first areas dealing with facial information in a specific way but was then found to be related to identity learning in TMS experiments (e.g., Ambrus et al., 2017b;

Eick et al., 2020). Top-down influence from higher-level areas, such as the FFA, cannot be ruled out as partial explanations for these findings. The same applies to other areas processing faces. Although the face-processing network is widespread, hierarchical, and highly interconnected, its single pieces are functionally specialized to a certain extent (Duchaine & Yovel, 2015; Haxby et al., 2000). This fact allows hypothesizing about the internal representations in each area and the related predictions that can be made by it. In summary, these features make the face-processing network perfect for investigating processes and consequences of predictive coding. Electrophysiological recordings from macaques' face-processing areas provide evidence for the effects of higher-order perceptual expectations and related prediction errors measured in a lower-level face responsive area (Schwiedrzik & Freiwald, 2017). Furthermore, the authors found prediction errors to reflect viewpoint invariant identity-specific violation. Their results show clear evidence for predictive processing in face processing areas.

Third, face perception processes are strongly influenced by context and prior perceptual experiences. Kok and colleagues (2017) had participants rate attractiveness in familiar and unfamiliar faces and found an effect known as serial dependence. More precisely, they showed that attractiveness ratings were dependent on the previously seen face images. However, this effect was more pronounced for unfamiliar than familiar faces, suggesting an identity representation that is less susceptible to short-term context effects for familiar faces. This finding also fits in perfectly with the precision weighting mechanisms in PC as identity can be defined as a precise prior. Whereas visual appearance (e.g., hairstyle, make-up, outfit) varies continuously, identity is a stable concept. In this context, a single observation of visual appearance (an unexpected piece of clothing) that deviates from our internal model would elicit a prediction error, but the posterior belief would be mainly driven by the prior belief (Figure 13C). However, a change of a stable characteristic, such as sex, in theory would be highly informative and signal the urgent necessity to update prior beliefs.

More evidence for context-dependent modulation of face perception comes from studies investigating face aftereffects, for example, for identity (Carbon et al., 2007; Hole, 2011; Walther et al., 2013). In a study by Walther and colleagues (2013), participants had to match a face image resulting from morphing two celebrities (with equal contribution) to either of them. In other words, they had to decide which of the two famous identities was more represented in the image after seeing either a veridical or a morphed image. The authors found a strong contrastive bias when the S1 was veridical (original) face but no such effect when S1 was an ambiguous face image in behavioral data. Such identity-related aftereffects were also shown to be long-lasting (Carbon et al., 2007).

Last, familiar faces are familiar because we already hold a neural representation for them that can be activated as soon as we see them. A stable face representation facilitates face perception processes (Jenkins & Burton, 2011). However, this also means that everything we experience in face-to-face interaction is interpreted in the context of the relationship with that person, or, in the case of famous identities, in the context of what we know about them. Face identification processes always involve bottom-up and top-down mechanisms, even in the case of unfamiliar faces, where identification occurs with little information about visual appearance, and identification may involve only deciding whether or not this person is familiar (Trapp et al., 2018). How exactly familiarity with individual faces develops is not fully clarified. That predictive coding models provide the basis for such a process seems logical since many other forms of learning can be explained similarly (Köster et al., 2020). Indeed, Apps and Tsakiris (2013) showed that face learning is accompanied by activity changes in the FFA that are best modeled by predictive coding. Naturally, for familiar faces, the process of recognizing and identifying a person is always connected to contextual, semantic information and requires retrieval of memories. The question arises of how much top-down information can be available when processing unknown faces. Brodski-Guerniero and colleagues (2017) provided evidence for a pre-activation of face knowledge measured with MEG. This activation of prior knowledge was manifest in measures of active information storage and led to increased predictable information in face processing regions, such as FFA and OFA, when participants performed a face but not a house detection task on Mooney stimuli (compare to Figure 4). Additionally, the authors proved increased information transfer to the FFA from higher-level visual areas, such as the anterior inferior temporal cortex. Their results suggest top-down influences for face detection. Although this might be essential due to the two-tone nature of the stimulus material, it seems logical that inferential processes play a role in unfamiliar face perception as well. Contextual information facilitates the recognition and perception of visual objects in general (Bar, 2004). Therefore, the processing of unfamiliar faces and facial information cannot happen independently from context information in real life. Just imagine meeting somebody that reminds you of a good friend. Although this person is unknown to you, an internal representation is activated immediately. Or, to give another example, you meet a friend on the street, and he or she is with another person you do not know. This person could be his or her partner, a sibling, or a colleague, to name a few examples. Depending on your knowledge and previous experience, you will automatically favor one of the possibilities - depending on what is most likely.

To sum it up, the identification process engages a distributed and recurrent network of brain areas which involves interactions with more regions the more familiar we are with a person (Kovács, 2020). This architecture is crucial to predictive processes and mechanisms that engage different functionally specialized areas that can contribute to the perceptual process as experiences are added, and representations evolve.

Additionally, the processing of faces has been extensively studied. A relatively large amount of knowledge is now available about the areas involved (Kovács, 2020), as well as on face-specific components responses of electrophysiological measures (Schweinberger & Burton, 2003). The large number of studies that already exist provides a valuable basis for generating clear hypotheses for neuronal behavior in different cortical areas.

1.4 Aims and questions of the thesis.

The studies of the current thesis focus on several possible consequences of predictive processing, investigated with experimental designs involving faces and measuring neural activity in face-sensitive regions with fMRI. From the large body of research in which faces were used as stimuli, it is evident that previously described phenomena as RS, the modulation of RS by repetition probability, and ES can be observed with facial stimulus material and recording of functional MRI in face responsive regions (Grotheer & Kovács, 2015; Henson, 2016; Summerfield & Koechlin, 2008). Therefore, they are ideal for studying prediction-related response alterations in further detail.

In these experiments, we investigate different neuronal phenomena, tested using faces in various stages of familiarization processes. Faces lend themselves as stimulus material because they are very complex on the one hand, but their processing is relatively effortless for humans. All three experimental studies focus on the fusiform face area, a key area of face processing but certainly not the only one in the ventral visual stream (see chapter 1.1.2). Still, its manifold connections to lower and higher-level areas processing visual input and its crucial role in processing faces make it perfect for investigating phenomena explained under the PC framework, such as response suppression to repetition or expectation.

The first empirical contribution (Study I, chapter 3.1) addresses RS and the question under which circumstances it can be modulated. RS has been explained within the predictive coding network in many studies because its magnitude can be modulated by changing aspects of the temporal context. While RS itself is a robust effect, its modulation by expectation was not found when using objects (Kaliukhovich & Vogels, 2011; Kovacs et al., 2013) or unfamiliar stimuli (Grotheer & Kovacs, 2014). However, in some studies, the so-called P_{REP} effect was not replicable when face stimuli were used (human BOLD responses: Olkkonen et al., 2017; monkey spiking activity and local field potentials: Vinken et al., 2018). Furthermore, many other theories provide adequate answers for the underlying mechanisms of RS (see chapter 1.2.2.1), and this effect was further shown to be independent of ES (Feuerriegel et al., 2018a; Grotheer & Kovács, 2015; Todorovic & de Lange, 2012). In order to prove whether RS itself relies on top-down prediction processes, other key features of PC models need to be tested in simple repetition designs without adding expectation-related manipulations. Therefore, the first study in this thesis investigates repetition suppression to unfamiliar faces under two conditions with varying sensory precision. Precision weighting is an essential component of predictive processing (see chapter 1.2.4), and we modeled it by manipulating the visibility of sensory data. According to predictive theories, RS is the manifestation of a reduced prediction error. Therefore we hypothesize that the manipulation of prediction errors by varying sensory precision should be reflected in measures of RS (Auksztulewicz & Friston, 2016). This could provide new evidence for underlying mechanisms of RS.

Study II (chapter 3.2) focuses on a specific feature of stimuli that is often present in RS studies and related to predictive precision: image predictability. Stimulus predictability refers to the ability to predict the incoming sensory data in a detailed and accurate manner. As described in chapter 1.2.5, it is related to the novelty or surprise of alternating stimuli within experimental designs. Stimulus novelty is often unequal between conditions depending on the available stimulus data set and the critical manipulations of studies investigating prediction-related phenomena. Therefore, a stronger surprise component in one condition compared to others can confound measures of RS. In most experimental settings, a repetition trial is defined as the presentation of the same image. Therefore, the second stimulus is always predictable in a detailed manner in repetition trials compared to alternation trials. Our empirical Study II (chapter 3.2) investigates this effect in a repetition paradigm. We implemented a statistical learning paradigm to train participants on specific stimulus associations. By employing rules of transitional statistical learning, alternating stimuli are equally predictable as repetitions in one context and can be compared to more standard measures of RS (i.e., non- or less predictable alternations). This way, we show whether and how much differences in stimulus predictability between repetition and alternation trials contribute to measures of RS.

Whereas the first two studies use unfamiliar faces as stimulus material in repetition designs, the third study investigates adaptation and expectation-related phenomena on higher-order face processing levels such as personal identity. We use a design comparable to procedures testing MMRs and ambient images of famous identities. As explained in chapter 1.2.6, MMR designs rely on multiple repetitions of a specific stimulus, making this stimulus more and more expected and turns a deviation from this stimulus into an unexpected event that elicits a mismatch response. Contrary to the previous studies, we did not repeat a specific face stimulus but a given famous identity. By this means, adaptation effects and expectation (violation) effects are investigated for higher than image-based face processing. We present participants with short blocks of different face images of one ID or different IDs to investigate image-independent adaptation effects in occipito-temporal regions. Furthermore, prediction error signals to unexpected changes in identity rather than specific images have not yet been investigated using fMRI. Therefore, we included a condition in which a trial composed of ambient images of identity (A) was ended by an image of identity(B). We hypothesize adaptation effects to trials depicting one identity compared to trials with alternating IDs and identity-specific expectation violations when identity changed unexpectedly.

So briefly, the central aims of the current thesis are the following.

1. *Does sensory input precision affect repetition suppression magnitude?*
2. *How can stimulus predictability confound repetition suppression magnitude when not equalized between repetition and alternation trials?*
3. *Can image-independent identity changes alter responses in ventral occipito-temporal regions?*

2 GENERAL METHODS

Since all the experiments presented use the same methods, I will briefly introduce the fMRI technique in this chapter. More importantly, this chapter includes the general analyses we used to process the neuroimaging data. While the individual designs are explained in detail in the corresponding manuscripts, procedures and analyses performed on the acquired fMRI data are often presented in an abbreviated form. Therefore, here I will detail not the method itself but the way we processed acquired data of all three studies.

2.1 The technique of functional magnetic resonance imaging

Functional magnetic resonance imaging exploits the relationship between neuronal activity and cerebral blood flow (CBF). This relationship is called neurovascular or neurometabolic coupling (Buchbinder, 2016). It is assumed that if any cerebral area is active, it will consume more energy in terms of oxygenated blood. As blood flow is relatively slow in comparison with neuronal activity changes, an area that is actively involved in processing a visual scene, for example, will first show a decrease in oxygen level before it will receive more oxygenated blood. The level of oxygen can be measured with functional MRI. The resulting signal used to inform about the activity level of specific regions is the blood oxygen level dependent (BOLD) signal. In other words, the hemodynamic response function (HRF) is imaged. As only the activity-related blood flow is measured with this technique, fMRI does not provide a direct measurement of neuronal activity.

Also, the BOLD signal is a measure of neuron population responses. In contrast to single-unit recordings, the BOLD signal will always reflect the neuronal activity of a large neuronal population. How many neurons we are talking about here depends on the voxel size, hence the spatial resolution. A detailed description of different preprocessing steps and the features of the resulting images are the content of chapter 2.2.

2.2 Functional preprocessing

Many data preparation steps are needed before meaningful statistical analyses to explore differences between conditions or groups can be conducted. We performed all preprocessing steps in SPM12 (Wellcome Department of Imaging Neuroscience, London, United Kingdom; Ashburner et al., n.d.), a software for **S**tatistical **P**arametric **M**apping designed to work with MATLAB (Mathworks). Descriptions of how all steps are performed in SPM are described in (Ashburner et al., n.d.). All measured data are saved in Digital Imaging and Communications in Medicine format (DICOM) and imported to SPM. The resulting 3D volumes of our functional data then need to be corrected for differences in acquisition time. In the sequences used to obtain functional data, a 3D volume is not acquired all at once but in a sequence of 2D slices acquired within the so-called repetition time (TR). The TR in all fMRI experiments reported here is two seconds. This results in a delay of approximately two seconds for the last slice being acquired in comparison to the first. To correct for these delays in slice acquisition, all slices of one volume can be interpolated to one time point for which a reference slice is selected. Therefore, it is crucial to know the exact order of slices being acquired. This first step is called slice-time correction and is the only preprocessing step on temporal data characteristics.

All following steps concern spatial corrections, the first being correction for head movements during the measurement. When participants move within or between runs during an experiment, naturally, voxels will be acquired in a different location. To account for head motion results, volumes need to be realigned to one reference volume. Successively, volumes are co-registered to a reference volume using a rigid-body transformation. To achieve the best match between two volumes, six parameters are computed that describe the differences in orientation between those two volumes. Three parameters describe translational moving (each along one axis in space), and three represent the rotational moving of the image (rotation around each of the three axes in space). These parameters are saved and used additional regressors when the general linear model (GLM) is specified. It is also possible to realign the data right away, meaning the realignment parameters will not only be estimated, but images will be resliced (Churchill et al., 2012). In the current studies, motion-related components are included in the GLM as six nuisance regressors. The subsequent two steps are types of coregistration as well and are necessary to match activity differences to anatomical locations and compare these results within a group of subjects. Although more than one pre-processing step relies on the concept of coregistration, only the structural-functional coregistration is termed ‘coregistration’ in SPM. Here the preprocessed anatomical image is aligned to the mean functional volume. Conceptually this is similar to the realignment step but differs, as anatomical and functional images are very distinct, especially in terms of spatial resolution and liability to noise in the form of distortions. Normalized mutual information is used as a function to align the two types of images that differ mostly in contrast. As another spatial transformation step needs to be executed, the transformation parameters are saved and applied to the images later. The next transformation step is important to ensure not only of activity but any measured functional or anatomical images are comparable between subjects or even with results from other studies. Therefore, a reference brain or template offering a common coordinate space for all the individual brain to be spatially transformed to is necessary. Classically, one of two templates is used, Talairach or MNI. Talairach coordinates refer to one reference brain in which anatomical landmarks were defined and used to construct a three-dimensional space. Templates from the Montreal Neurological Institute (MNI) – the so-called MNI templates – offer a better representation of the population. In the current studies, all images were normalized to the MNI template ICBM152, which is the result of averaging 152 normal MRI scans. In the presented data sets, we first normalized the anatomical images to the MNI template and used the resulting so-called deformation field to normalize all the functional volumes of the same participant. In other words, the deformation field stores the information about how the anatomical volume was warped into the standard space to apply them to the functional data as well. As the transformation parameters of the last step (structural-functional-coregistration) are stored, this will write out all functional images co-registered and normalized. The final step, before statistical models are defined and estimated, is called spatial smoothing. Although a lot of previous preprocessing steps try to compensate for differences, especially between brains (normalization), there is always some variability left in the data. To cope with this and to increase the signal-to-noise ratio, spatial smoothing is applied. In the presented studies, we used a Gaussian filter with a kernel width of 8mm full-width at half-max (FWHM). Spatial smoothing, of course, results in a loss of spatial resolution but increases statistical power.

At this point, all the preprocessing of the obtained images is completed. Next, the data needs to be mapped to the experimental conditions presented when the data was obtained. Therefore, the exact timing and duration of conditions are entered together with the preprocessed data and the stored movement parameters as additional regressors. We specified general linear models (GLM) based on a chosen function which is the canonical hemodynamic response function for all experiments presented in this thesis.

The estimated models of the condition-related BOLD signals serve as the basis for subsequent analyses, such as region-of-interest (ROI) based or whole-brain analyses. In all presented studies, the stimulus material depicted faces, which is why ROI-based analyses are always the main analyses. How the respective ROIs are defined is described in chapter 2.3. A whole-brain analysis was added to not overlook effects somewhere else in the brain, in case literature suggested the involvement of other areas in the investigated effect.

2.3 Functional localization of face areas within the occipito-temporal cortex

In each of the presented studies, we added an fMRI sequence for localizing the ROIs to the experimental procedures. Those sequences only differ slightly in terms of stimulus material and paradigm from each other. All localizing paradigms utilized images of faces, objects, and Fourier transformed noise patterns. Whereas 40 greyscale images with a presentation time of 300ms each were used in the first two studies, the third study used colorful images presented with a frequency of 4Hz to reduce measuring time. As the manuscripts themselves only offer a short description of the stimulus material and design of the localizing sequences, an example is depicted in Figure 14.

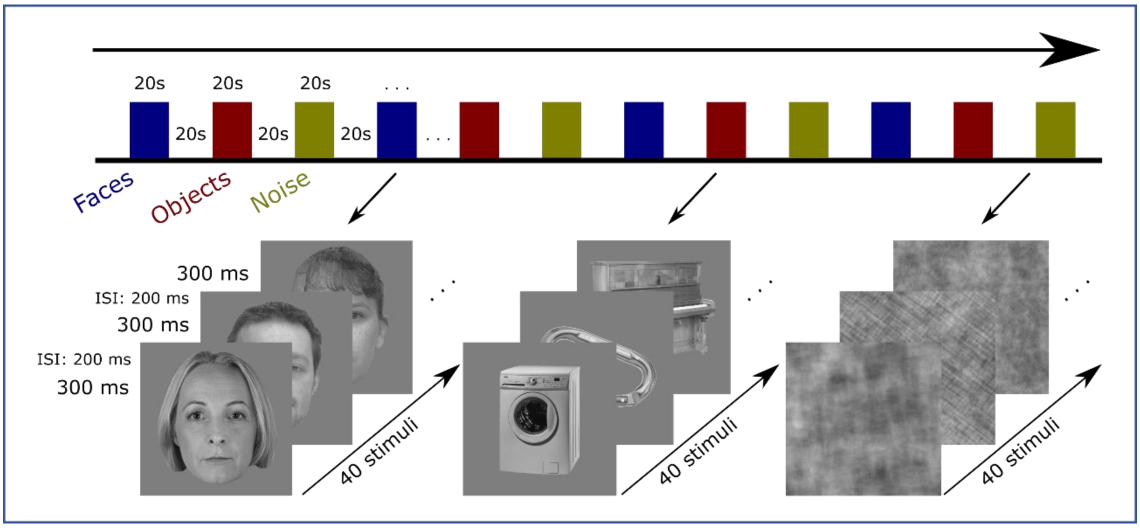


Figure 14. Example for a paradigm to functionally localize face-and object-processing regions. This type of paradigm was used in Studies I and II. Blocks of faces, objects and Fourier-randomized noise were repeated at least four time, interleaved by breaks of 20 second.

Figure 15 shows the results of the whole-brain analyses of the localizer data of Study III. Here, faces were contrasted with objects and Fourier randomized noise images and displayed with a threshold of $p<0.0001_{\text{uncorrected}}$ and a minimum cluster size of 50 voxels. Data from Study III were chosen for this illustration because we collected the largest sample in this study. Thirty participants took part in the study and for this analysis no one was excluded. Although the implemented localizer paradigm differed from the one depicted in Figure 14, Table 1 shows that the mean coordinates defined in all studies correspond to each other. Our localizer revealed another significant cluster for the contrast faces versus objects and noise in the frontal lobe (MNI[x,y,z]: 40, 8, 32), near the inferior frontal gyrus, which also assumed to be involved in processing faces (for comparison of coordinates, see Axelrod & Yovel, 2013; Ishai et al., 2005).

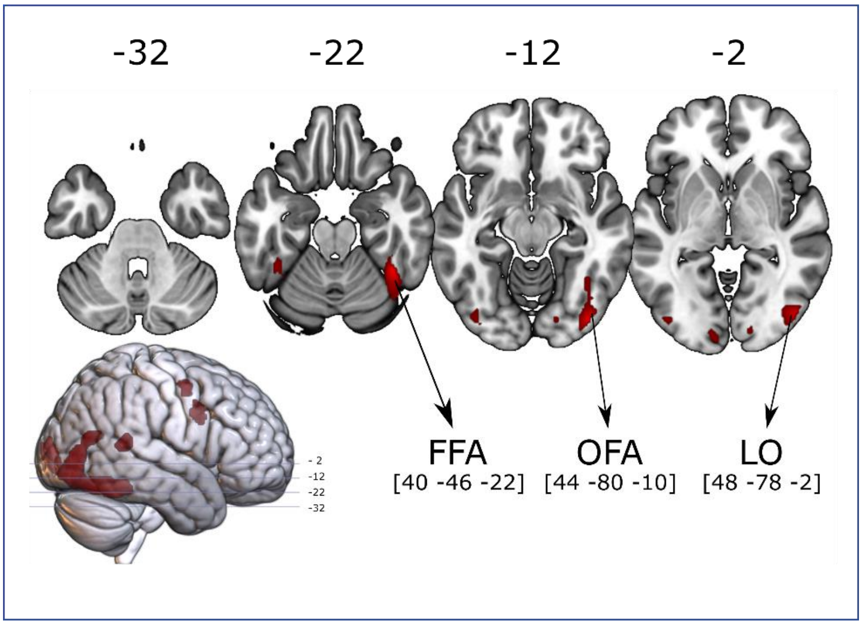


Figure 15. Results of the functional localizer on a group level. Faces were contrasted with objects and Fourier randomized noise images and displayed with a threshold of $p<0.0001_{unc.}$ *and a minimum cluster size of 50 voxels on a template brain using MRIcroGL (https://www.nitrc.org/projects/mricrogl/). Depicted coordinates refer to the MNI template.*

Table 1. Mean (±SE) MNI coordinates [x, y, z] for bilateral FFA and OFA for each study, defined on the basis of the respective functional localizer.

	Left FFA			Right FFA			Left OFA			Right OFA		
	x	y	z	x	y	z	x	y	z	x	y	z
Study I	-39 (1)	-54 (2)	-18 (1)	41 (1)	-53 (1)	-19 (1)	-40 (1)	-74 (2)	-13 (1)	41 (1)	-76 (2)	-12 (1)
Study II	-40 (1)	-51 (2)	-21 (1)	41 (1)	-48 (1)	-21 (1)	-40 (1)	-75 (2)	-13 (1)	42 (1)	-74 (2)	-13 (1)
Study III	-40 (1)	-52 (2)	-20 (1)	41 (0)	-53 (2)	-19 (1)	-40 (1)	-77 (2)	-12 (1)	42 (1)	-75 (2)	-12 (1)

3 EMPIRICAL CONTRIBUTION

The following sections present the current studies.

Study I, page 53 ff.

Repetition suppression for noisy and intact faces in the occipito-temporal cortex.

Rostalski, S.-M. M., Amado, C., & Kovács, G. (2019). Repetition suppression for noisy and intact faces in the Occipito-temporal cortex. Frontiers in Psychology, 10(JUN), 1348. https://doi.org/10.3389/fpsyg.2019.01348

Study II, page 63 ff.

Measures of repetition suppression in the fusiform face area are inflated by co-occurring effects of statistically learned visual associations.

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, 131, 123–136. https://doi.org/10.1016/j.cortex.2020.07.010

Study III, page 85 ff.

Person identity-specific adaptation effects in the ventral occipito-temporal cortex

Rostalski, S.-M., Robinson, J., Ambrus, G. G., Johnston, P. & Kovács, G. (2021). Person identity-specific adaptation effects in the ventral occipito-temporal cortex. European Journal of Neuroscience [under review]

3.1 Study I. Repetition suppression for noisy and intact faces in the occipito-temporal cortex

Sophie-Marie Rostalski[1], Catarina Amado[2] and Gyula Kovács[1]

[1]Department of Biological Psychology and Cognitive Neurosciences, Institute of Psychology,
Friedrich Schiller University Jena, Jena, Germany,

[2]Department of Computer Science, Experimental Cognitive Science Research Group, Eberhard Karls
Universität Tübingen, Tübingen, Germany

Main research question:

Does sensory input precision affect repetition suppression magnitude?

ORIGINAL RESEARCH
published: 14 June 2019
doi: 10.3389/fpsyg.2019.01348

Repetition Suppression for Noisy and Intact Faces in the Occipito-Temporal Cortex

*Sophie-Marie Rostalski[1], Catarina Amado[2] and Gyula Kovács[1]**

[1]Department of Biological Psychology and Cognitive Neurosciences, Institute of Psychology, Friedrich Schiller University Jena, Jena, Germany, [2]Department of Computer Science, Experimental Cognitive Science Research Group, Eberhard Karls Universität Tübingen, Tübingen, Germany

Repetition suppression (RS), the relative lower neural response magnitude to repeated as compared to non-repeated stimuli, is often explained within the predictive coding framework. According to this theory, precise predictions (priors) together with less precise sensory evidences lead to decisions that are determined largely by the predictions and the other way around. In other words, the prediction error, namely the magnitude of RS, should depend on the precision of predictions and sensory inputs. In the current study, we aimed at testing this idea by manipulating the clarity and thereby the precision of the incoming sensory data by adding noise to the images. This resulted in an fMRI adaptation design with repeated or alternating trials showing clear or noisy face stimuli. Our results show a noise effect on the activity in the fusiform face area (FFA), namely less activation for noisy than for clear trials, which supports previous findings. No such effects could be found in OFA or LO. Data also showed reliable RS in the FFA (bilateral) and unilaterally in OFA (right) and LO (left). Interestingly, the noise added to the stimuli did not affect the magnitude of RS in any of the tested cortical areas. This suggests that the clarity of the sensory input is not crucial in determining the magnitude of RS.

Keywords: repetition suppression, predictive coding, precision, noise, fusiform face area

Edited by:
Maria Olkkonen,
Durham University, United Kingdom

Reviewed by:
Stephen J. Gotts,
National Institute of Mental Health (NIMH), United States
Kasper Vinken,
KU Leuven, Belgium

Correspondence:
Gyula Kovács
gyula.kovacs@uni-jena.de

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Psychology

Received: *03 December 2018*
Accepted: *23 May 2019*
Published: *14 June 2019*

Citation:
Rostalski S-M, Amado C and Kovács G (2019) Repetition Suppression for Noisy and Intact Faces in the Occipito-Temporal Cortex. Front. Psychol. 10:1348. doi: 10.3389/fpsyg.2019.01348

INTRODUCTION

Repetition suppression (RS), the relative lower neural response magnitude to repeated as compared to non-repeated stimuli, is one of the most studied phenomena of cognitive neurosciences. Over the last years, not only RS but other stimulus repetition-related phenomena, such as expectation suppression or surprise-related response elevation, were explained under the framework of predictive coding (Summerfield et al., 2008; Todorovic and de Lange, 2012; Grotheer and Kovacs, 2014; Mayrhauser et al., 2014; Grotheer and Kovács, 2015). This theory states that perception is not determined solely by the incoming stimuli themselves, but it is also modulated by inferential processes (Rao and Ballard, 1999). In other words, the sensory inputs together with our prior experiences are used to form predictions of upcoming events to ensure efficient processing (Friston, 2005). For a better understanding of these processes, several prior studies manipulated the temporal context of a stimulus to alter predictions (Auksztulewicz and Friston, 2016; Grotheer and Kovács, 2016). These studies suggested that stimulus repetitions lead to lower prediction errors and this is manifest in RS while rarely presented, thereby surprising, stimuli lead to higher prediction errors and enhanced neural responses (for a review, see Grotheer and Kovács, 2016). Although recently, numerous studies explained repetition and expectation-related phenomena under the framework of predictive coding, this explanation is not unchallenged in the literature. While expectations seem capable of modulating RS in many cases, RS and expectation suppression (ES) were dissociated from each other in several studies (Todorovic and de Lange, 2012; Grotheer and Kovács, 2015; Feuerriegel et al., 2018) and therefore seem to reflect different neuronal mechanisms. Further, human fMRI studies with objects (Kovacs et al., 2013; Grotheer and Kovacs, 2014) and nonhuman primate single-cell studies with objects as well as recent single-cell (Vinken et al., 2018) or fMRI (Olkkonen et al., 2017) studies with faces failed to find any trace of modulatory effects of expectation on RS (see, however, Mayrhauser et al., 2014 and Kronbichler et al., 2018 for a different conclusion). Therefore, the role of top-down modulatory effects, such as predictions and expectation, in determining the magnitude of RS is under heavy discussion as of today.

Although RS seems to be a robust phenomenon, that has been investigated in several paradigms (for reviews, see Grill-Spector et al., 2006; Krekelberg et al., 2006), there are evidences for repetition enhancement (i.e., an enhanced neural response for repeated stimuli) as well (Henson, 2003; Turk-Browne et al., 2007; De Gardelle et al., 2013; Segaert et al., 2013; Recasens et al., 2015). For example, Turk-Browne et al. (2007) could show that the attenuating effect on the BOLD responses of showing two identical scenes compared to two different scenes in one trial could be reversed by reducing the contrast of the stimuli. This modulatory effect on neural responses to repetitions is introducing an important factor into the RS research field, namely precision or clarity of visual input.

Prior studies (Auksztulewicz and Friston, 2016) conceptualized prediction error as the magnitude of neural responses in certain "error units." The repeated presentations of a given stimulus would, in turn, reduce the activity of these neurons, leading to RS. According to theories of predictive coding, precise predictions (priors) together with less precise sensory evidences lead to decisions that are determined largely by the predictions; in other words, the prediction error is increased if predictions fail to come true. This would in turn reduce the magnitude of RS for noisy when compared to clear sensory inputs. However, if the predicted priors are less precise (for example due to the frequent occurrence of unexpected events) but the incoming sensory stimuli are clear

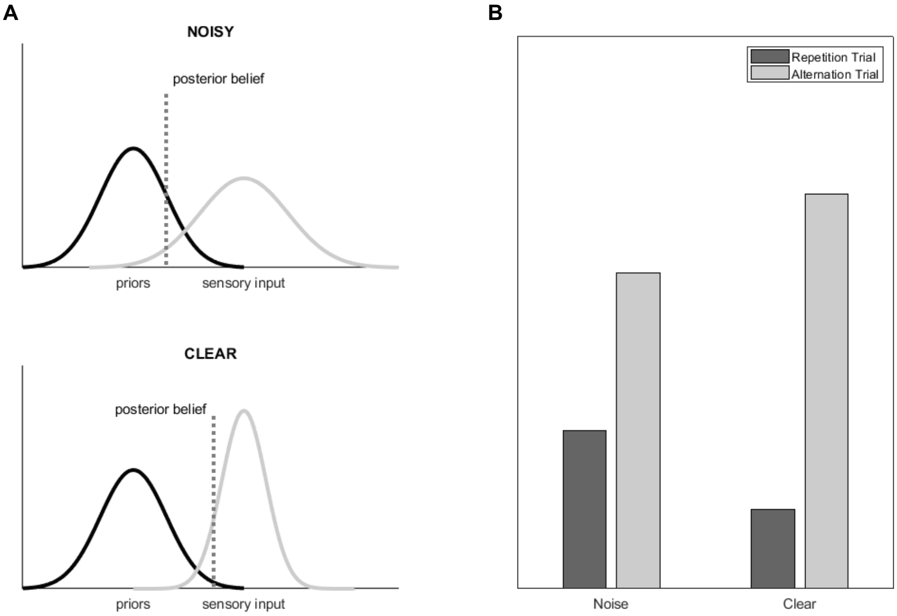


FIGURE 1 | (A) The graphs show the Gaussian probability distributions that represent the distribution of the priors (i.e., the *a priori* beliefs, black) and of the sensory evidences (gray) as well as the resulting posterior beliefs (dotted line) for a situation where the statistics-based priors are kept constant and the precision of the sensory stimuli is modulated (for example by adding noise to the images). Precision can be understood as the inverse of the distribution width. Reducing the relative precision of the sensory input biases the posterior beliefs toward the priors and thereby reduces predictive error (figure adapted from Adams et al., 2013).
(B) Theoretical BOLD signal magnitudes for alternating and repeating stimulus pairs and for stimuli with and without noise, separately. Note that these results assume that predictive error is reflected in the magnitude of RS (Grotheer and Kovács, 2016).

and precise, then the *a posteriori* decisions are rather determined by the sensory stimulation and the prediction errors are lower (O'Reilly et al., 2012; Adams et al., 2013; Auksztulewicz and Friston, 2016; Sterzer et al., 2018).

In the current study, we aimed at testing this idea by manipulating sthe precision of the incoming sensory data by adding noise to the images. We reasoned that the modulatory effect of stimulus precision on prediction errors might be reflected in the magnitude of RS. Because sensory uncertainty is assumed to reduce the difference of priors and posterior beliefs (**Figure 1A**; Adams et al., 2013; Sterzer et al., 2018), a smaller prediction error (RS magnitude) is expected for noisy, as compared to clear visual stimuli. In other words, alternations lead to much higher prediction errors and repetitions to lower prediction errors when the sensory input is clear as compared to noisy, which results in higher RS magnitude (**Figure 1B**) for clear visual inputs.

Indeed, previous studies suggest the differential processing of noisy stimuli (Wild and Busey, 2004; Banko et al., 2011). For example, Banko et al. (2011) manipulated task difficulty by decreasing the phase coherence of face stimuli and found that this affects early electrophysiological responses. The visually evoked P1 showed a higher amplitude to noisy stimuli, whereas the face-sensitive N170 showed a lower amplitude. In line with the P1 modulation, fMRI data showed increased activation in the lateral occipital cortex (LO) due to noise (Banko et al., 2011). Also, Hermann et al. (2015) found that noisy stimuli with lower phase coherence lead to increased activity in the LO. This suggests increased processing demands in the visual cortex due to added visual noise. However, authors also found reduced activity in the face-selective fusiform face area (FFA) when noise was added (Hermann et al., 2015). In addition, a linear increase in the amplitude of a face-sensitive ERP component (N170) (Jemel et al., 2003) could be observed by decreasing the level of a Gaussian distributed noise, added to face stimuli gradually. This result could later be confirmed with fMRI data by Horner and Andrews (2009) who manipulated phase coherence and found evidence for the principle of scaling for preferred stimuli in the FFA. This suggests the linearity of the BOLD response and the noise level in face stimuli. Altogether, these studies show that visual noise indeed affects neural processing, but it is not clear what impact that effect has on inferential processes and subsequent predictions.

To the best of our knowledge so far, no study compared the effect of stimulus repetitions for noisy and clear stimuli in the ventral temporal cortex. Therefore, in the present study, noise was added to face stimuli to manipulate the precision of sensory stimulation in a design containing repeated and alternating trials. Trials could therefore either consist of pairs of clear or noisy faces, which could either be the same or different. Activity in face-specific areas (FFA and OFA, occipital face area) as well as in LO was acquired using fMRI. Based on prior evidences (Horner and Andrews, 2009; Hermann et al., 2015), noisy stimuli were expected to elicit lower BOLD responses than clear ones in the regions of fusiform gyrus, but an enhanced response was expected in the lateral occipital regions (Banko et al., 2011; Hermann et al., 2015). Also, in line with the predictive coding theory, repeated trials were hypothesized to show a smaller neuronal response than alternating trials. We reasoned that this RS effect should be modulated by the clarity of the stimuli if predictions are indeed less precise for noisy as compared to clear stimuli (Auksztulewicz and Friston, 2016).

MATERIALS AND METHODS

Participants

Twenty-three subjects participated in this study. One subject was excluded from the analysis due to excessive movements during image acquisition. The remaining 22 participants (11 females, one left-handed and one both left- and right-handed) were between 18 and 31 years of age ($M = 22$; SD = 3.81) and all had normal or corrected to normal vision. Previous fMRI studies, using stimulus pairs and reporting significant RS, as well as significant predictive modulations of RS (Summerfield et al., 2008; Kovacs et al., 2013; Grotheer et al., 2014; Grotheer and Kovacs, 2014) were typically able to find modulatory effects of RS by other factors, such as probabilistic predictions with sample sizes between 11 and 26. Therefore, here, we reasoned, that with the tested number of participants, we could reliably detect any interaction of noise and RS and this is supported by the results of the Bayes factor analyses.

Participants were fully informed about the study and gave written consents to participate. They received course credits for participation. The experiment was conducted in accordance with the guidelines of the Declaration of Helsinki, and with the approval of the ethics committee of the University of Jena.

Stimuli and Procedure

A total set of 490 unfamiliar faces (246 clear (127 female) and 244 noisy (127 female)) were used as stimuli. Those were shown in the center of the screen with a superimposed grey scale mask, which additionally covered the hair and the shape of the face resulting in round-shaped faces including eyes, nose, and mouth. Noisy stimuli were generated by superimposing Fourier-transformed versions of the original images on the faces where phase coherence was reduced (45%) by the weighted mean phase technique (Dakin et al., 2002).

Participants completed three experimental runs, each including 120 trials of the four different trial types (**Figure 2**) in a randomized fashion. One trial included two stimulus presentations, which could either be the same face (repeated trial) or two different faces (alternating trial). Participants' task was to detect target trials, in which the second face stimulus was tilted by 10° either clockwise or counterclockwise and to indicate this direction by pressing a button (**Figure 3**). Such target trials were equally distributed across the four conditions and represented 20% of the overall trial amount and were excluded from any further analysis.

For defining regions of interests (ROI), a localizer sequence was performed. Grayscale images of faces, objects, and

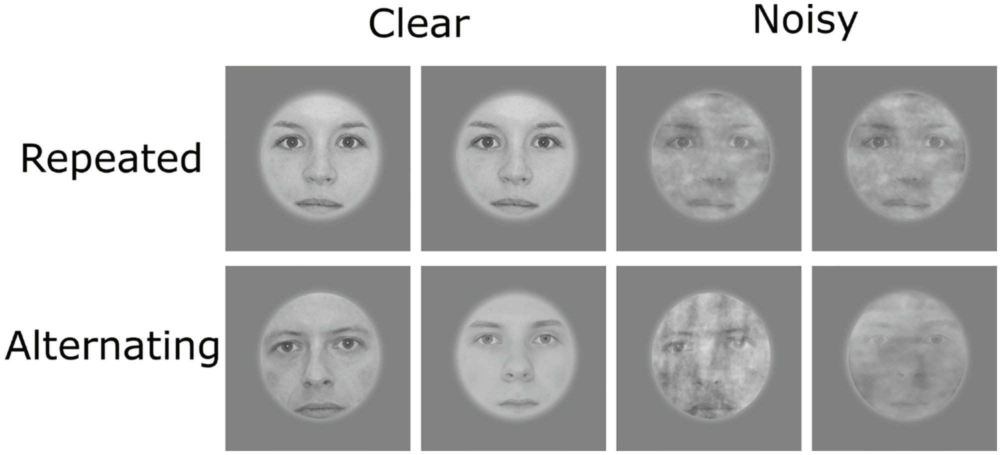


FIGURE 2 | Examples for the four possible trial types (excluding target trials). Written informed consent for publishing these images was given by the respective persons.

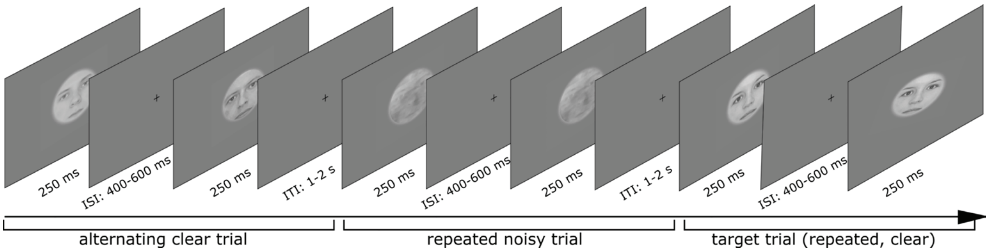


FIGURE 3 | Sequence of three trials including clear alternating, noisy repeated, and one target trial (clear repeated). Written informed consent for publishing these images was given by the respective persons.

Fourier-transformed noise patterns were presented (exposition time: 300 ms, interstimulus interval: 200 ms) in blocks of 20 s intermitted by a break of 20 s and were repeated four times.

Imaging Parameters

Neuroimaging was performed using a Prisma fit 3 T MRI Scanner from Siemens. During the functional runs, T2*-weighted images (35 slices, TR = 2,000 ms, TE = 30 ms, isotropic voxel size of 3 mm) were acquired continuously. High-resolution T1*-weighted simages (TR = 2,300 ms, TE = 3.03 ms, isotropic voxel size of 1 mm) were acquired to obtain a 3D structural scan. Data were preprocessed using SPM12 (Wellcome Trust Centre for Neuroimaging, University College London, UK). The functional images were slice-timed, realigned, co-registered to the structural scan, and afterward normalized to the MNI space and smoothed using an 8-mm Gaussian kernel.

ROIs were defined using the data from the localizer sequence and canonical hemodynamic response functions (HRFs) were extracted using MarsBaR (Brett et al., 2002). HRFs were estimated for all subjects and ROIs. Then, peak values were submitted to repeated measurement ANOVAs with the factors noise level (clear vs. noisy) and repetition (repeated vs. alternating).

RESULTS

Behavioral Results

A repeated measurements ANOVA with the factors noise (*clear* vs. *noisy*) and trial type (*repeated* vs. *alternating*) was conducted for the reaction times and accuracy. Regarding the reaction times, no significant effect was revealed from the analysis. However, a significant main effect for noise level was found for the accuracy rates, $F(1,21) = 12.00$, $p < 0.01$, $\eta = 0.36$, which shows better performance for clear ($M = 94.4\%$, SD = 10.3%) as compared to noisy trials ($M = 88.9\%$, s = 12.3%).

Neuroimaging Results

Neuroimaging results are depicted in **Figure 4**. We found a similar pattern in the FFA of the two hemispheres. A significant main effect of noise level was found, $F(1,19) = 14.18$, $p < 0.01$, $\eta = 0.43$ for right hemisphere and $F(1,20) = 19.45$, $p < 0.001$, $\eta = 0.49$ for left hemisphere, with clear trials eliciting larger BOLD signal than the noisy ones. Additionally, a significant main effect of trial type was observed in both hemispheres: $F(1,19) = 16.22$, $p < 0.001$, $\eta = 0.46$ for right hemisphere and $F(1,20) = 15.99$, $p < 0.001$, $\eta = 0.44$ for left hemisphere. This effect suggests a generally higher signal for alternating as

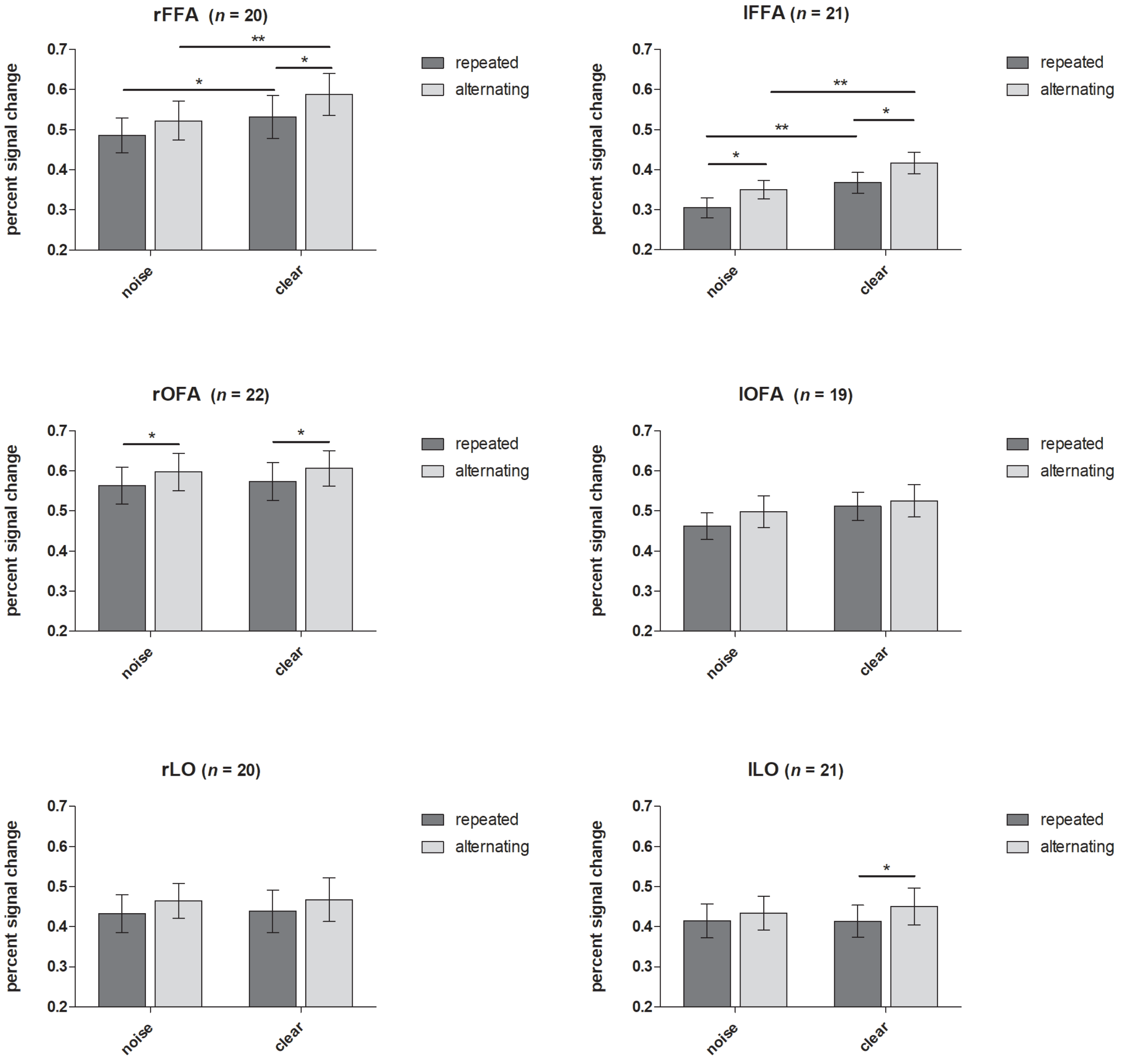


FIGURE 4 | Effects of noise and repetitions. Percent signal changes of FFA, OFA, and LO (left and right hemispheres for each) are presented for condition and trial type. Error bars indicate standard errors. Displayed significant differences refer to Fisher's LSD *post hoc* test. *$p < 0.05$, **$p < 0.01$.

compared to repeated trials. However, no interaction between noise level and trial type was found, neither for the right, $F(1,19) < 1$, $p = 0.44$, $\eta = 0.03$, nor for the left hemisphere, $F(1,20) < 1$, $p = 0.87$, $\eta = 0.00$, suggesting that the observed RS is similar for noisy and clear stimuli in the FFA.

The same analysis performed on the right OFA revealed a significant main effect of trial type for the right hemisphere, $F(1,21) = 8.14$, $p < 0.01$, $\eta = 0.28$, showing that alternating trials elicit greater signal changes than repeated ones. However, the main effect of noise remained nonsignificant, $F(1,21) < 1$, $p = 0.54$, $\eta = 0.02$, as well as the interaction effect did, $F(1,21) < 1$, $p = 0.94$, $\eta = 0.00$. The same analysis for the left OFA revealed no significant main effect or interaction.

In the LO, similar to the OFA, the main effect of trial type was found to be significant in one hemisphere, the left one only, $F(1,20) = 4.60$, $p < 0.05$, $\eta = 0.19$. Again, repetitions led to lower signal than alternations in the LO as well whereas the right hemisphere showed a strong tendency in this direction only, $F(1,20) = 3.59$, $p = 0.07$, $\eta = 0.16$. No other main effect or interaction reached significance.

We additionally evaluated the likelihood that there is no interaction between the two factors using a Bayesian repeated measures ANOVA to substantiate our conclusion. This analysis, performed in JASP (JASP Team, 2018), provides the Bayes factor, reflecting how much more likely a dataset reflects the null hypothesis compared to the alternative hypotheses. To get the Bayes factor for the interaction, we performed the division of BF01 of the model containing the two main effects and the interaction between these by the model containing the two main effects only. Values reported here show the Bayes factor for the null hypothesis against the hypothesis of an interaction.

The estimated Bayes factor (null/alternative) for an interaction of condition and trial type in the right FFA was 2.6, suggesting that the null hypothesis of no interaction is 2.6 times more likely than the alternative hypothesis. Bayes factor (null/alternative) for an interaction effect in the left FFA was 3.1, providing substantial evidence for the null hypothesis.

Regarding the right OFA, a Bayes factor of 3.3, meaning that data are 3.3 times more likely to occur under the null hypothesis, provides substantial evidence against the presence of an interaction effect between the two factors. In the left OFA, the calculated Bayes factor for the interaction between condition and trial type was 2.9, implying that the null hypothesis of no interaction is 2.9 times more likely than the alternative hypothesis.

The Bayesian repeated measures ANOVA in the right LO revealed a Bayes factor of 3.1, providing substantial evidence for the null hypothesis, that there is no interaction between the two factors noise level and trial type. For the left LO, the Bayes factor for the interaction effect model was 2.6, suggesting that the observed data are 2.6 more likely to occur under the null hypothesis.

Altogether, the Bayes factor analyses supported the conclusion that clarity does not affect the magnitude of RS in any of the tested ROIs.

DISCUSSION

The present study aimed at investigating the impact of added sensory noise on repetition suppression. First, the results show an effect of noise level on the activity in the FFA but not in OFA or LO. This is in line with other results showing lower FFA activity (Horner and Andrews, 2009) or a weaker electrophysiological signal in electrodes over the temporal cortex (Banko et al., 2011) in response to noisy faces. Regarding the lateral occipital regions, the same studies showed even an enhanced processing in these, earlier stages of visual processing (Banko et al., 2011; Hermann et al., 2015) when exposed to noisy stimuli. However, there is also evidence for the opposite result, namely a reduced activation with increasing noise level (Malach et al., 1995) or no noise effects at all (Jemel et al., 2003; Wild and Busey, 2004). In the current study, noise had no effect on the LO, which is at odds with prior studies (Malach et al., 1995; Banko et al., 2011; Hermann et al., 2015). The chosen noise level could be one factor leading to this result. We applied 45%, whereas prior studies applied slightly higher (55%) noise levels (Hermann et al., 2015). It is possible that more noise is necessary to affect the activity of the lower level visual areas, whereas higher level visual areas are more sensitive to added visual noise. Also, in studies finding an elevated BOLD signal for noise in the LO, participants had to perform more demanding tasks like gender categorization (Wild and Busey, 2004; Banko et al., 2011), and the higher task-difficulty might led to enhanced neural activity. In contrast, the target detection task in the current study was comparatively easy and this could lead to the similar activity for noisy and clear stimuli. The fact, however, that the behavioral results show an effect of noise on participants' performance argues against this interpretation.

Repetition suppression was found in all the tested regions, even if not always in both hemispheres (OFA and LO), regardless of the noise level. Therefore, this study joins the large body of evidence for this robust effect (for a review, see Grill-Spector et al., 2006). Interestingly, the noise added to the stimuli did not affect the magnitude of RS in any of the tested cortical areas. This suggests that the neural mechanisms driving RS are similar for clear and noisy stimuli. The exact nature of these mechanisms is highly debated currently in the literature. Specifically, while electrophysiological single-cell recording studies suggest that RS is explained by local or bottom-up mechanisms, such as fatigue (Carandini and Ferster, 1997), several current neuroimaging studies support the role of top-down mechanisms, such as predictions, in explaining RS (for a review, see Grotheer and Kovács, 2016).

Theories of predictive coding (Rao and Ballard, 1999; Friston, 2005) assume that the human central nervous system continuously makes inferences or predictions about the surrounding sensory environment and estimates the difference of the actual incoming and predicted inputs (prediction error). Many studies have suggested so far that this prediction error is reflected in the repetition-related response reduction of neurons (RS; Summerfield et al., 2008, 2011; Todorovic and de Lange, 2012; Grotheer and Kovacs, 2014; Mayrhauser et al., 2014; Grotheer and Kovács, 2015). In addition, recent theoretical (O'Reilly et al., 2012) and clinically motivated studies of predictive coding (Adams et al., 2013; Sterzer et al., 2018) suggested that the magnitude of the prediction error, therefore of RS, should depend on the precision of the predictions, as well as of the incoming sensory data. It has been suggested that precise predictions together with more noisy sensory evidence lead to enhanced prediction errors while less precise priors with more precise incoming sensory stimuli lead to reduced prediction errors (O'Reilly et al., 2012; Adams et al., 2013; Auksztulewicz and Friston, 2016; Sterzer et al., 2018). Here we modeled the precision of the sensory data by adding noise to our stimuli and we kept the precision of the predictions (i.e., the volatility of the system, Summerfield et al., 2011) constant. We asked if the modulatory effect of stimulus precision on prediction errors is reflected in the magnitude of RS. To our surprise, the results suggest that RS is insensitive to the manipulations of the precision of incoming sensory inputs, at least if the precision is modulated by adding noise to the stimuli.

The reason for the lack of modulatory effect of noise might be due to the fact that RS is the result of the interaction of multiple neural processes. While many prior human electrophysiological and neuroimaging studies explained RS in the framework of predictive coding (for a review, see Grotheer and Kovács, 2016), other studies explained RS by simpler, bottom-up, or local mechanisms (Kaliukhovich and Vogels, 2014; Vogels, 2016; Olkkonen et al., 2017; Vinken et al., 2018). Indeed, the separation of RS from its modulation by stimulus probabilities and thereby by expectation was confirmed by many studies (Larsson and Smith, 2012; Todorovic and de Lange, 2012; Grotheer and Kovács, 2015; Feuerriegel et al., 2018). We presented our Rep and Alt trials with equal probabilities; therefore, we did not modulate probabilistic expectations. Thus, it is possible that the manipulation of sensory precision affects only the modulation of RS by top-down factors, such as probabilistic expectations, but not the magnitude of RS *per se*. This would explain why we observed similar RS for noisy and clear stimuli and at the same time requires further specifically targeted studies to test. This fact, together with the currently heavily debated neural mechanisms of RS (Vinken et al., 2018), does not allow us to conclude that the precision of incoming sensory stimulation has no effect on predictive processes at all. Nonetheless, our results clearly show that the precision of the sensory input is not crucial in determining the RS magnitude *per se*.

Also, we did not observe repetition enhancement effects for the less visible, noisy stimuli which have been reported by Turk-Browne and colleagues in their study (Turk-Browne et al., 2007). However, there are several conceptual differences between their experiment and the current one. First, they used a different stimulus set (indoor and outdoor scenes) and therefore measured the BOLD response in different areas (parahippocampal place area) as we did. Second, the task was an orthogonal orientation discrimination task in our case, while an indoor-outdoor scene discrimination in the Turk-Brown study, meaning that it directed attention to the stimulus content. Third, and above all, while we used short-lagged stimulus pairs (with 500 ms average ISI) the Turk-Brown study used much longer, 3-s-long ISI with masked presentations, and it is not clear so far if these two types of presentations provoke the same neuronal mechanism or not. Altogether, these differences make the comparisons of the two studies difficult.

Manipulating the precision of sensory data was not sufficient to affect RS magnitudes at all in our study. Provided prediction errors are reflected in RS at all, posterior beliefs may be more determined by the precision of the predicted priors than by the precision of the sensory inputs. The precision of the priors can be modeled by applying stable, highly predictable or more variable, volatile stimulus sequences. Indeed, Summerfield et al. (2011) found that the repetition probability-induced modulation of RS (measured on visual evoked potentials) was present during stable stimulation segments but disappeared almost entirely when the stimulation became volatile. The aim of the current study was to test the precision of the sensory stimuli only; therefore, we did not make an effort to modulate stability/volatility here. Also, we assumed the priors to be the same for both noisy and clear conditions and for alternating and repeated stimulus pairs. In other words, we kept the probabilities of the four trial types equal and constant across the experiment. Still, we cannot exclude entirely the possibility, that the *a priori* hypotheses of the "noisy world" are different from those of a "clear world." In other words, introducing noise to the sensory input might have had an effect on the predictive priors as well. Therefore, the lack of a modulatory effect of sensory data precision on RS suggests that future studies should manipulate sensory data precision together with the precision of prior predictions. Including precision manipulations into probabilistic prediction paradigms (e.g., as in Summerfield et al., 2008) will provide more insight into predictive processes.

CONCLUSION

The findings of this study are in agreement with previous studies showing a reducing effect of noise in the region of the fusiform gyrus (Horner and Andrews, 2009; Banko et al., 2011). The enhanced activation in more lateral occipital regions found in earlier investigations (Banko et al., 2011; Hermann et al., 2015) could not be confirmed. This suggests a different sensitivity to noise of the different regions. Significant RS was present in the FFA (bilateral), right OFA, and left LO. Evidence for a modulatory effect of precision on RS could not be proved. Therefore, future studies should focus on independently manipulating the precision of prior beliefs and sensory inputs for a better understanding of its impact on predictive processes.

ETHICS STATEMENT

Participants were fully informed about the study and gave written consents to participate. They received course credits for participation. The experiment was conducted in accordance with the guidelines of the Declaration of Helsinki and with the approval of the ethics committee of the Faculty of Social and Behavioural Sciences of the University of Jena.

AUTHOR CONTRIBUTIONS

S-MR, CA, and GK designed the concept of the article. S-MR and CA ran the experiments and analyzed the data. S-MR and GK wrote the article.

FUNDING

This work was supported by grants from the Deutsche Forschungsgemeinschaft (KO 3918/5-1).

ACKNOWLEDGMENTS

We thank Sandrine Hinrichs for helping with data collection.

REFERENCES

Adams, R. A., Stephan, K. E., Brown, H. R., Frith, C. D., and Friston, K. J. (2013). The computational anatomy of psychosis. *Front. Psych.* 4, 1–26. doi: 10.3389/fpsyt.2013.00047

Auksztulewicz, R., and Friston, K. (2016). Repetition suppression and its contextual determinants in predictive coding. *Cortex* 80, 125–140. doi: 10.1016/j.cortex.2015.11.024

Banko, E. M., Gal, V., Kortvelyes, J., Kovacs, G., and Vidnyanszky, Z. (2011). Dissociating the effect of noise on sensory processing and overall decision difficulty. *J. Neurosci.* 31, 2663–2674. doi: 10.1523/JNEUROSCI.2725-10.2011

Brett, M., Anton, J.-L., Valabregue, R., and Poline, J.-B. (2002). Region of interest analysis using an SPM toolbox (abstract). *Presented at the 8th International Conference on Functional Mapping of the Human Brain, June 2-6, 2002, Sendai, Japan.* Available on CD-ROM in Neuroimage, Vol. 16, No 2.

Carandini, M., and Ferster, D. (1997). A tonic hyperpolarization underlying contrast adaptation in cat visual cortex. *Science* 276, 949–952. doi: 10.1126/science.276.5314.949

Dakin, S. C., Hess, R. F., Ledgeway, T., and Achtman, R. L. (2002). What causes non-monotonic tuning of fMRI response to noisy images? *Curr. Biol.* 12, R476–R477. doi: 10.1016/S0960-9822(02)00960-0

De Gardelle, V., Waszczuk, M., Egner, T., and Summerfield, C. (2013). Concurrent repetition enhancement and suppression responses in extrastriate visual cortex. *Cereb. Cortex* 23, 2235–2244. doi: 10.1093/cercor/bhs211

Feuerriegel, D., Churches, O., Coussens, S., and Keage, H. A. D. (2018). Evidence for spatiotemporally distinct effects of image repetition and perceptual expectations as measured by event-related potentials. *NeuroImage* 169, 94–105. doi: 10.1016/j.neuroimage.2017.12.029

Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 815–836. doi: 10.1098/rstb.2005.1622

Grill-Spector, K., Henson, R., and Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci.* 10, 14–23. doi: 10.1016/j.tics.2005.11.006

Grotheer, M., Hermann, P., Vidnyánszky, Z., and Kovács, G. (2014). NeuroImage repetition probability effects for inverted faces. *NeuroImage* 102, 416–423. doi: 10.1016/j.neuroimage.2014.08.006

Grotheer, M., and Kovacs, G. (2014). Repetition probability effects depend on prior experiences. *J. Neurosci.* 34, 6640–6646. doi: 10.1523/JNEUROSCI.5326-13.2014

Grotheer, M., and Kovács, G. (2015). Neuropsychologia the relationship between stimulus repetitions and ful filled expectations. *Neuropsychologia* 67, 175–182. doi: 10.1016/j.neuropsychologia.2014.12.017

Grotheer, M., and Kovács, G. (2016). Can predictive coding explain repetition suppression? *Cortex* 80, 113–124. doi: 10.1016/j.cortex.2015.11.027

Henson, R. N. A. (2003). Neuroimaging studies of priming. *Prog. Neurobiol.* 70, 53–81. doi: 10.1016/S0301-0082(03)00086-8

Hermann, P., Banko, E. M., Gal, V., and Vidnyanszky, Z. (2015). Neural basis of identity information extraction from noisy face images. *J. Neurosci.* 35, 7165–7173. doi: 10.1523/JNEUROSCI.3572-14.2015

Horner, A. J., and Andrews, T. J. (2009). Linearity of the fMRI response in category-selective regions of human visual cortex. *Hum. Brain Mapp.* 30, 2628–2640. doi: 10.1002/hbm.20694

JASP Team (2018). JASP (Version 0.9)[Computer software].

Jemel, B., Schuller, A. M., Cheref-Khan, Y., Goffaux, V., Crommelinck, M., and Bruyer, R. (2003). Stepwise emergence of the face-sensitive N170 event-related potential component. *Neuroreport* 14, 2035–2039. doi: 10.1097/00001756-200311140-00006

Kaliukhovich, D. A., and Vogels, R. (2014). Neurons in macaque inferior temporal cortex show no surprise response to deviants in visual oddball sequences. *J. Neurosci.* 34, 12801–12815. doi: 10.1523/JNEUROSCI.2154-14.2014

Kovacs, G., Kaiser, D., Kaliukhovich, D. A., Vidnyanszky, Z., and Vogels, R. (2013). Repetition probability does not affect fMRI repetition suppression for objects. *J. Neurosci.* 33, 9805–9812. doi: 10.1523/JNEUROSCI.3423-12.2013

Krekelberg, B., Boynton, G. M., and van Wezel, R. J. A. (2006). Adaptation: from single cells to BOLD signals. *Trends Neurosci.* 29, 250–256. doi: 10.1016/j.tins.2006.02.008

Kronbichler, L., Said-yürekli, S., and Kronbichler, M. (2018). Perceptual expectations of object stimuli modulate repetition suppression in a delayed repetition design. *Scientific Reports* 8:12526. doi: 10.1038/s41598-018-31091-4

Larsson, J., and Smith, A. T. (2012). fMRI Repetition Suppression: Neuronal Adaptation or Stimulus Expectation? *Cerebral Cortex* 22, 567–576.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc. Natl. Acad. Sci. USA* 92, 8135–8139. doi: 10.1073/pnas.92.18.8135

Mayrhauser, L., Bergmann, J., Crone, J., and Kronbichler, M. (2014). Neural repetition suppression: evidence for perceptual expectation in object-selective regions. *Front. Hum. Neurosci.* 8, 1–8. doi: 10.3389/fnhum.2014.00225

O'Reilly, J. X., Jbabdi, S., and Behrens, T. E. J. (2012). How can a Bayesian approach inform neuroscience? *Eur. J. Neurosci.* 35, 1169–1179. doi: 10.1111/j.1460-9568.2012.08010.x

Olkkonen, M., Aguirre, G. K., and Epstein, R. A. (2017). Expectation modulates repetition priming under high stimulus variability. *J. Vision*. 17, 1–16. doi: 10.1167/17.6.10

Rao, R. P. N., and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87. doi: 10.1038/4580

Recasens, M., Leung, S., Grimm, S., Nowak, R., and Escera, C. (2015). Repetition suppression and repetition enhancement underlie auditory memory-trace formation in the human brain: an MEG study. *NeuroImage* 108, 75–86. doi: 10.1016/j.neuroimage.2014.12.031

Segaert, K., Weber, K., de Lange, F. P., Petersson, K. M., and Hagoort, P. (2013). The suppression of repetition enhancement: a review of fMRI studies. *Neuropsychologia* 51, 59–66. doi: 10.1016/j.neuropsychologia.2012.11.006

Sterzer, P., Voss, M., Schlagenhauf, F., and Heinz, A. (2018). Decision-making in schizophrenia: a predictive-coding perspective. *NeuroImage* 190, 133–143. doi: 10.1016/j.neuroimage.2018.05.074

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M.-M., and Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nat. Neurosci.* 11, 1004–1006. doi: 10.1038/nn.2163

Summerfield, C., Wyart, V., Johnen, V. M., and de Gardelle, V. (2011). Human scalp electroencephalography reveals that repetition suppression varies with expectation. *Front. Hum. Neurosci.* 5, 1–13. doi: 10.3389/fnhum.2011.00067

Todorovic, A., and de Lange, F. P. (2012). Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *J. Neurosci.* 32, 13389–13395. doi: 10.1523/JNEUROSCI.2227-12.2012

Turk-Browne, N. B., Yi, D. J., Leber, A. B., and Chun, M. M. (2007). Visual quality determines the direction of neural repetition effects. *Cereb. Cortex* 17, 425–433. doi: 10.1093/cercor/bhj159

Vinken, K., Op de Beeck, H., and Vogels, R. (2018). Face repetition probability does not affect repetition suppression in macaque inferotemporal cortex. *J. Neurosci.* 38, 7492–7504. doi: 10.1523/JNEUROSCI.0462-18.2018

Vogels, R. (2016). Sources of adaptation of inferior temporal cortical responses. *Cortex* 80, 185–195. doi: 10.1016/j.cortex.2015.08.024

Wild, H. A., and Busey, T. A. (2004). Seeing faces in the noise: stochastic activity in perceptual regions of the brain may influence the perception of ambiguous stimuli. *Psychon. Bull. Rev.* 11, 475–481.

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

3.2 Study II. Measures of repetition suppression in the fusiform face area are inflated by co-occurring effects of statistically learned visual associations

Sophie-Marie Rostalski [a], Catarina Amado [b], Gyula Kovács [a,+++], Daniel Feuerriegel [c,+++]

[a] Institute of Psychology, Friedrich Schiller University Jena, Jena, Germany

[b] Experimental Cognitive Science, Eberhard Karls Universität Tübingen, Tübingen, Germany

[c] Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, Australia

[+++] These two authors contributed to the publication equally as senior authors.

Main research question:

How can stimulus predictability confound measures of repetition suppression when not equalized between repetition and alternation trials?

CORTEX 131 (2020) 123–136

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research Report

Measures of repetition suppression in the fusiform face area are inflated by co-occurring effects of statistically learned visual associations

Sophie-Marie Rostalski [a,*], *Catarina Amado* [b], *Gyula Kovács* [a,1] *and Daniel Feuerriegel* [c,1]

[a] *Institute of Psychology, Friedrich Schiller University Jena, Jena, Germany*
[b] *Experimental Cognitive Science, Eberhard Karls University Tübingen, Tübingen, Germany*
[c] *Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, Australia*

ARTICLE INFO

Article history:
Received 23 October 2019
Reviewed 20 January 2020
Revised 20 March 2020
Accepted 3 July 2020
Action editor Pia Rotshtein
Published online 29 July 2020

Keywords:
Repetition suppression
Expectation
Prediction
fMRI

ABSTRACT

Repeated presentation of a stimulus leads to reductions in measures of neural responses. This phenomenon, termed repetition suppression (RS), has recently been conceptualized using models based on predictive coding, which describe RS as due to expectations that are weighted toward recently-seen stimuli. To evaluate these models, researchers have manipulated the likelihood of stimulus repetition within experiments. They have reported findings that are inconsistent across hemodynamic and electrophysiological measures, and difficult to interpret as clear support or refutation of predictive coding models. We instead investigated a different type of expectation effect that is apparent in stimulus repetition experiments: the difference in one's ability to predict the identity of repeated, compared to unrepeated, stimuli. In previous experiments that presented pairs of repeated or alternating images, once participants had seen the first stimulus image in a pair, they could form specific expectations about the repeated stimulus image but not the alternating image, which was often randomly chosen from a large stimulus set. To assess the contribution of stimulus predictability effects to previously observed RS, we measured BOLD signals while presenting pairs of repeated and alternating faces. This was done in contexts whereby stimuli in alternating trials were either i.) predictable through statistically learned associations between pairs of stimuli or ii.) chosen randomly and therefore unpredictable. We found that RS in the right fusiform face area (FFA) was much larger in trials with unpredictable compared to predictable alternating faces. This was primarily due to unpredictable alternating stimuli evoking larger BOLD signals than predictable alternating stimuli. We show that imbalances in stimulus predictability across repeated and alternating trials can greatly inflate measures of RS, and potentially even mimic RS effects. Our findings indicate that stimulus-specific expectations as described by predictive coding models may account for a sizeable portion of observed RS effects.

---

[*] *Corresponding author.*
E-mail address: sophie-marie.rostalski@uni-jena.de (S.-M. Rostalski).

[1] These two authors contributed to the publication equally as senior authors.

https://doi.org/10.1016/j.cortex.2020.07.010

1. Introduction

Repeated presentation of a stimulus leads to reduced measures of neural responses, as observed using a variety of electrophysiological and neuroimaging techniques (for a review see Grill-Spector et al., 2006). Such effects are commonly known as repetition suppression (RS) or adaptation. Similarly, the correct and fulfilled expectation of a forthcoming stimulus also leads to reduced responses when compared to unexpected or surprising stimuli, for several stimulus categories and measures (known as expectation suppression, ES; for a review see Summerfield & Egner, 2009).

Explanations of repetition-as well as expectation-related phenomena under the framework of predictive coding (Rao & Ballard, 1999) have gained traction in recent years. This is because, in contrast to several other neurobiologically-plausible models of RS (for review see Grill-Spector et al., 2006), predictive coding models describe mechanisms that can potentially account for observed RS effects, and also how RS might be modulated by processes related to perceptual expectations and attention (e.g., Eger, 2004; Murray & Wojciulik, 2004). Predictive coding models conceptualize RS as a reduction of prediction error signals, due to perceptual expectations that are weighted toward recently-encountered stimuli (e.g., Auksztulewicz & Friston, 2016; Friston, 2005; Grotheer & Kovács, 2015). Factors such as attention are hypothesized to modulate the precision of sensory predictions (Feldman & Friston, 2010), which in turn influence the extent of observed RS. Predictive coding models describe different mechanisms than those in recently-formulated local circuit models of RS (Dhruv et al., 2011; Kaliukhovich & Vogels, 2016; Solomon & Kohn, 2014). However, the notion of precision in predictive coding models allows us to test hypotheses about how attention and perceptual expectations affect RS, while specific hypotheses have not yet been derived for the above-mentioned local circuit models.

Summerfield et al. (2008) was the first to provide empirical support for the predictive coding model by showing that neuroimaging measures of RS can be modulated by contextual factors, such as the probability of stimulus repetition. They presented pairs of faces in each trial and reported that BOLD signal differences between repeated and unrepeated stimuli (i.e., repetition effects) were larger in blocks with high (75%), compared to blocks with low (25%) repetition probability. This interaction involving repetition probability was replicated several times using faces (for a review see Grotheer et al., 2014), and also for other stimulus categories such as letters (Grotheer & Kovács, 2014) and other non-face objects (Kronbichler et al., 2018; Mayrhauser et al., 2014). Notably, this interaction has mostly been reported in studies using functional magnetic resonance imaging (fMRI); when using electrophysiological measures researchers have found separable, non-interacting repetition and expectation effects (Feuerriegel, Churches, et al., 2018; Kaliukhovich & Vogels, 2014; Todorovic & de Lange, 2012; Vinken et al., 2018, with the exception of Summerfield et al., 2011; but see Feuerriegel, Churches, et al., 2018 for an alternative explanation of this result).

When interpreting these findings, it is important to differentiate the neural mechanisms of RS from how RS is measured within an experiment (typically as a difference between a comparable repeated and unrepeated stimulus condition). In such experiments, any effect that will influence repeated and unrepeated stimulus-evoked responses in different ways will also contribute to the measured magnitude of RS, even if that effect is unrelated to the underlying processes responsible for RS (reviewed in Feuerriegel, 2016). In Summerfield et al. (2008) and similar experiments, participants could learn to expect stimulus repetitions in the 75% repetition blocks, whereby in the same block unrepeated stimulus trials were relatively rare and surprising. Conversely, in the 25% repetition blocks the unrepeated stimuli were instead expected, and the repeated stimuli relatively surprising. Accordingly, the observed RS by expectation interaction could actually be produced by additive effects of genuine RS and another, expectation related suppression effect (ES; Kaliukhovich & Vogels, 2011; Larsson & Smith, 2012), with expectations suppressing responses to either repeated or unrepeated stimuli in different block types.

More recent studies have used “cue” stimuli, whereby the first stimulus in each trial signals the probability of stimulus repetition, in order to distinguish between additive and interactive effects of ES and RS. Todorovic and de Lange (2012) presented pairs of auditory tones, which could either repeat or change within a trial. The pitch of the first tone predicted stimulus alternation or repetition with 75% probability. They reported that RS and ES, as indexed by magnetoencephalography (MEG), were separable and occurred at distinct time windows. In a similar design using face stimuli Grotheer and Kovács (2015) reported that effects of RS and ES on BOLD signals did not interact and were partly dissociable in the time course of their effects on the hemodynamic response. In a follow-up study Amado et al. (2016) added a ‘neutral’ condition, in which expectations were not weighted toward either repeated or alternating stimuli, to separately quantify effects of fulfilled expectations and surprise. They found that surprise had a much larger effect on BOLD signals than fulfilled expectations, and that this effect of surprise was apparent for alternating (but not repeated) stimulus conditions (see also e.g., Figure 2 in De Gardelle et al., 2013; Figure 2 in Larsson & Smith, 2012). This suggests that, instead of ES modulating repetition effects, RS might in fact inhibit surprise-related response enhancement, as found in a recent electroencephalography (EEG) study (Feuerriegel, Keage, et al., 2018). These results, combined with the inconsistency of findings across fMRI and electrophysiological recording methods, does not provide clear evidence that expectations modulate RS in the way previously specified by predictive coding models.

Besides expectations relating to stimulus repetition probability, there is another type of expectation that is prevalent in studies of RS, and relevant for evaluating predictive coding models of repetition effects. There is evidence from single-cell recordings of non-human primates (Meyer & Olson, 2011) as well as human electrophysiological and neuroimaging experiments (Feuerriegel, Churches, et al., 2018; Hall et al., 2018; Pajani et al., 2017; Turk-Browne et al., 2009) indicating that associations are formed between images that are shown temporally close together, and this association modulates neural responses. The proposed underlying mechanism is that the observers learn about the transitional statistics or rules of the stimulation sequences, as humans do from early childhood on to learn about their environment (Fiser & Aslin, 2002; Romberg & Saffran, 2011). As a seminal example, Meyer and Olson (2011) trained macaques to associate originally unrelated images by presenting the same stimulus pairs over a prolonged time period. The animals learned that one leading image was always followed by a specific trailing image. In a subsequent session, single-neuron activity was recorded from inferotemporal cortex (IT) while the animals viewed stimulus pairs which were either previously associated or randomly paired. IT neurons exhibited higher firing rates following stimuli which violated previously learned transitional rules, compared to those that were associated with the previous image.

This type of statistically learned expectation is relevant to a large number of stimulus repetition designs that have been used in the past. In these designs, participants are presented with two stimuli in each trial, which may be of the same or different identities. In repetition trials the identity of the second stimulus can be predicted after seeing the first stimulus in the trial, however the alternating (unrepeated) stimulus is often randomly-chosen from a set of multiple stimuli, and is very difficult to predict with any certainty (Feuerriegel, 2016). In these cases, there is a difference in the predictability of repeated and alternating stimulus images. Importantly, this difference in predictability is distinct from the effects of fulfilled expectations and surprise mentioned above. Predictability is defined as whether a specific stimulus image can be anticipated prior to its appearance, whereas expectation and surprise effects relate to whether an observer's expectation is fulfilled or violated upon presentation of a stimulus.

This imbalance in predictability across repetition and alternation trials could theoretically inflate the magnitude of, or even produce, many previously observed RS effects. Pajani et al. (2017) investigated this using a design that manipulated the predictability of the alternating stimuli. They presented stimuli in repetition blocks, composed of 75% repetition and 25% alternation trials, and alternation blocks, with only a 25% portion of repetition trials. Crucially, in a third block type 25% of trials were repetitions and 75% were predictable alternations, where the second stimulus in the trial was repeatedly paired with the first stimulus during a prior training session. They observed large differences in the magnitude of repetition effects that were apparently due to reductions in BOLD signals for predictable compared to unpredictable alternating faces. Further evidence for predictability effects came from a recent EEG study (Feuerriegel, Churches, et al., 2018), who used a similar blocked design with predictable and unpredictable alternating faces. In the so-called “AB” blocks in that experiment the second stimulus in each trial could either be the same image as the first (repetition trials), or a specific same-sex face (predictable alternation trials). In the “AX” blocks, however, the second stimulus could either be a repetition of the first one, or a same- sex face, selected randomly from a set of 23 stimuli (unpredictable alternation trials). Differences in event-related potential (ERP) repetition effect magnitudes across AB and AX blocks were found during multiple time windows post stimulus onset. Importantly, these differences in observed repetition effects were due to differences in ERP responses to alternating stimuli across block types, while no differences across AB and AX blocks were found for repeating stimuli.

Critically, this study did not equate the relative novelty of AB and AX alternating stimuli, as each individual face identity was presented many more times in the AB compared to AX conditions. Similarly, in Pajani et al. (2017) the predictable alternating stimuli were presented many more times during the experiment than the unpredictable alternating stimuli, which were trial-unique. Because of this, it is unclear whether the observed effects were primarily due to effects of stimulus predictability or stimulus novelty, both of would have similar hypothesised effects on neural responses (e.g., Feuerriegel, 2016; Mur et al., 2010; Xiang & Brown, 1998).

We used a similar design to investigate the interplay of stimulus repetition and predictability effects using fMRI, while controlling for the relative novelty of predictable and unpredictable alternating stimuli. The previously introduced conditions in Feuerriegel, Churches, et al. (2018) were adopted, including predictable (AB) and unpredictable (AX) alternating trials. RS was measured by comparing BOLD signals in trials with repeated and alternating stimulus pairs. Importantly, prior to the fMRI scanning session participants underwent 4 training sessions on consecutive days, during which they were presented with 6 predictable alternating face pairs (i.e., the first face of a pair was always followed by a given same-sex face) to create specific face associations for the alternating trials. Because previous fMRI studies that presented face stimuli (Amado et al., 2016; Egner et al., 2010; Summerfield et al., 2008) found the most pronounced effects of stimulus repetition and perceptual expectations in the fusiform face area (FFA; Kanwisher et al., 1997) we focused our analyses on this region.

Our design allowed us to control for effects of stimulus novelty, enabling a more accurate estimate of stimulus predictability effects in repetition designs. This also allowed us to assess whether this type of expectation may account for a portion of previously observed RS effects. To foreshadow our results, we found that predictability does modulate BOLD responses in the FFA and acts primarily upon responses to alternating stimuli. While our results support the notion of separable RS and predictability effects, they also indicate that, when predictability is confounded with stimulus repetition, as in a large number of existing studies, RS effects are likely to be inflated (or perhaps even caused) by this predictability confound.

2. Methods

2.1. *Participants*

We report how we determined our sample size, all data exclusions (if any), all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

Twenty-two volunteers participated in the study. Sample size was determined according to other studies investigating a modulation of RS from our lab (Grotheer & Kovács, 2014; Amado et al., 2016; Rostalski et al., 2019). All were informed about the procedure of the study and gave written consent for participation beforehand. One participant was excluded due to incomplete experiment while another participant's data was partially lost due to technical issues related to the MRI scanner. The remaining 20 participants (3 males; 4 left-handed) were between 19 and 28 years of age ($M = 21.9$, $SD = 2.53$). All had normal or corrected-to-normal vision. The experiment was conducted in accordance with the guidelines of the Declaration of Helsinki, and with the approval of the ethics committee of the University of Jena.

2.2. *Stimuli*

We presented 12 images of upright female faces as stimuli. Pictures were cropped to show faces without hair or clothes, resized to 440 × 400 pixels, converted to greyscale and equated in average luminance (Fig. 1A). Stimuli were presented against a black background using Psychtoolbox v.3.0.14 (Brainard, 1997; Kleiner et al., 2007) in MATLAB 2014a (The Mathworks). For each participant six stimuli were allocated randomly to be presented in the AB and the remaining six in the AX conditions. Please note that we did not have permission to publish the face images used in this experiment. As these images are protected by a license they cannot be shared.

2.3. *Experiment design*

Participants first completed a series of behavioural training sessions, followed by an fMRI session (see Fig. 1C). The experimental design, including the stimuli and task, was identical across training and fMRI data acquisition sessions, except where specified otherwise.

In each trial (Fig. 1A) an adapter (S1) and test stimulus (S2) were presented for 250 msec, separated by an inter-stimulus interval (ISI) of 400–600 msec (randomised across trials). The image size of S2 was 20% smaller than that of S1 to avoid low-level adaptation processes. Trials were separated by an inter-trial interval (ITI): for the training sessions, the ITI was 1800, 2000 or 2200 msec, randomly distributed across trials, and for the fMRI sessions it was 6, 8 or 10 sec.

In each trial, S1 and S2 could either be identical (repetition trials; Rep) or depicting different identities (alternation trials; Alt). These trial types were presented in two different contexts (Fig. 1B), labelled as “AB” and “AX”. In the AB context, the S2 face could either be a repetition of the S1 face (Rep trials), or a specific face identity that had previously been repeatedly paired and associated with the S1 identity during the training sessions (Alt trials). In these Alt trials of the AB context, each S1 face identity was consistently paired with one of the five other face identities that were allocated to the AB context. Each S1 identity in the AB stimulus set was paired with a different S2 face identity, ensuring that each face image would be presented an equal number of times throughout the experiment. In other words, once the participant has seen a given S1 face “A”, they could form expectations regarding the S2 to be a repetition of face “A” or a different, specific identity “B”. In the AX context S2 could either be the repetition of the S1 image, or a different identity, pseudo-randomly selected from the set of 5 other face identities. Therefore, in the AX context, there were no consistent pairings between S1 and S2 face identities for the Alt trials: S2 could be any of the five other faces, allocated to AX, ensuring that each face appeared the same number of times throughout this condition. This procedure ensured further that each AB and AX face identity was presented the same number of times across the experiment. Thus altogether, we had two independent factors: trial type (Rep or Alt) and context, reflecting prior associations formed for Alt trials (AB) or not having such transitional rules (AX). The proportion of Rep and Alt trials (i.e., the probability of stimulus repetition) was 50% in both AB and AX contexts.

2.4. *Procedure*

MATLAB code used to run the experiment will be available at https://osf.io/akygb/ at the time of publication. Study methods and analyses were not pre-registered prior to the research being conducted. Participants completed four training sessions across four consecutive days prior to the fMRI measurements. The fMRI session followed the last training session immediately on the fourth day. Each training session was composed of twelve blocks (60 trials per block, 720 trials per session) and lasted approximately 40 min. All sessions took place at approximately the same time of the day, in the afternoon hours to control for potential changes of attention that occur across the circadian cycle (Valdez et al., 2010).

During the training sessions participants learned the S1–S2 transition probabilities associated with face identities in the AB and AX contexts. Trials of AB and AX context were presented randomly interleaved within the same blocks of trials and with equal probability. For the AB context, each face image pairing (in Alt trials) was presented 120 times throughout the training sessions while for the AX context each of the possible Alt S1–S2 combinations was shown 24 times.

During the training and fMRI sessions the participants' task was to decide whether the S1 and S2 were the same or different face images by pressing one of two keys on a keyboard (training session) or MRI-compatible button box (fMRI session). The spatial layout of response keys/buttons and associated response fingers were kept constant across behavioural and fMRI sessions. Instructions were presented in the centre of the screen prior to each run. Participants took a self-paced break between each run. The entire fMRI session lasted approximately 60 min.

2.5. *Image acquisition*

Four experimental runs were completed, with each lasting for about 10 min and including 60 trials. A total of 240 trials were presented during the fMRI session. An additional localizer sequence was included to define the location of the FFA bilaterally (blocks of 40 images, size: 600 × 600 pixels on a grey background; exposition time: 300 msec, ISI: 200 msec; presenting faces, objects and Fourier-randomized noise patterns lasting for 20 sec each). To identify FFA and occipital face area (OFA) we contrasted face blocks with object and noise blocks (for $n = 14$) or face blocks with noise blocks (for $n = 6$) using a threshold of either $p < .05$ family-wise error

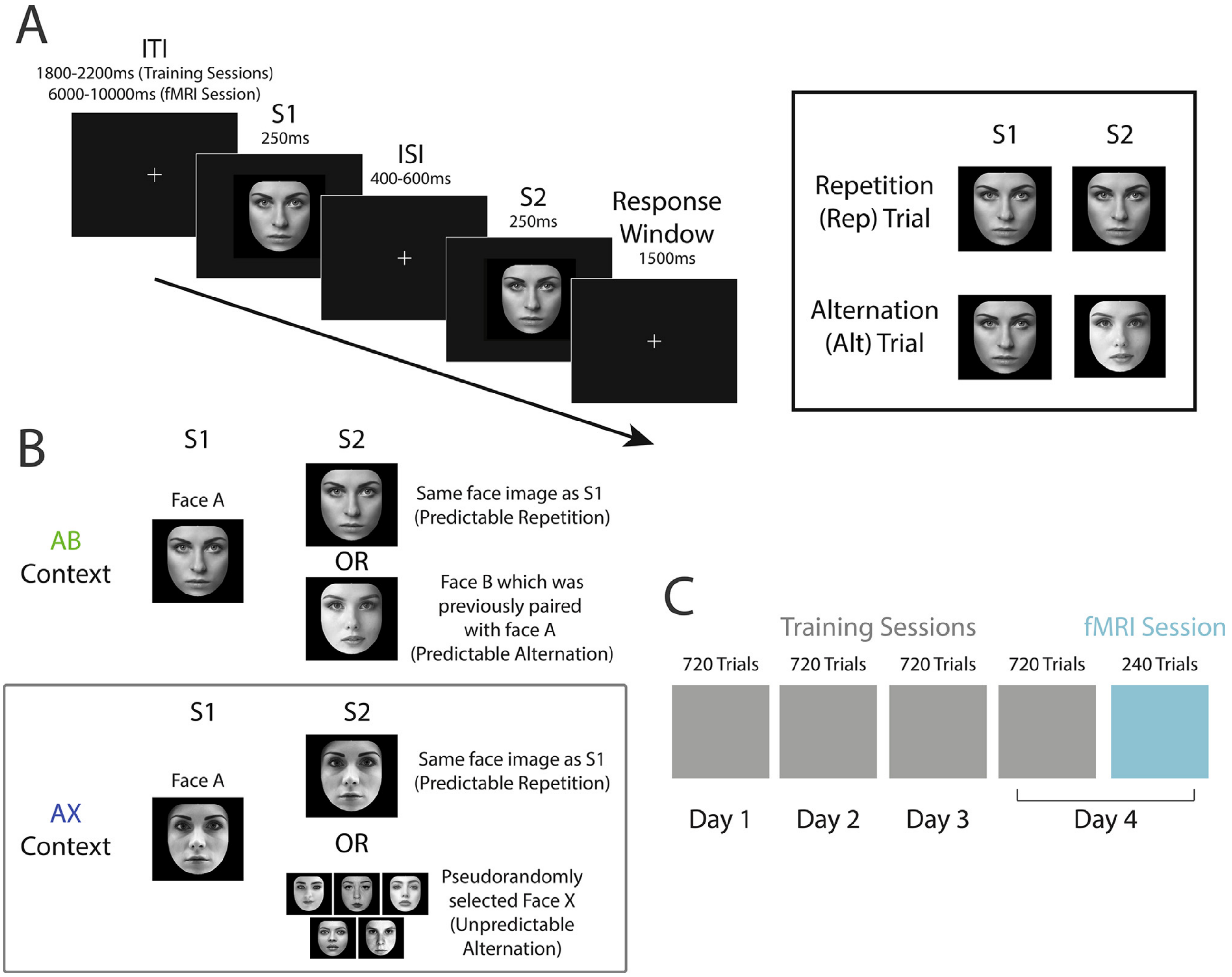


Fig. 1 – Trial structure and predictability cueing manipulation. A) In each trial S1 and S2 face stimuli were presented, separated by a 400–600 msec inter-stimulus interval (ISI). The S2 stimulus could either be the same face image as S1 (repetition trials) or a different female face (alternation trials). B) For alternation trials, the S2 face could either be a particular face "B" that was repeatedly paired with a specific S1 face "A" during the training sessions (AB context) or pseudorandomly-chosen from a set of 5 different faces (AX context). The probability of stimulus repetition was fixed at 50% across both contexts. C) Participants completed 4 training sessions over consecutive days. Trial structure, task (same-different forced choice), stimuli and AB/AX contexts were the same as in the fMRI scanning session but with a shorter ITI duration. Following the fourth training session participants then completed the fMRI session on the same day. All face images shown here are subject to a Pixabay license (https://pixabay.com/hu/service/license/). These images were not part of the actual stimulus set, as we do not have permission to publish the original images. All images in this figure have been processed in the same way as the original stimuli.

(FWE) corrected ($n = 15$) or $p < .0001$ uncorrected ($n = 5$). Lateral occipital cortex (LO) was defined by contrasting object and noise blocks. The threshold parameters used to identify FFA, OFA and LO were always the same within one participant. We could identify the right FFA in 19 out of 20 participants (average MNI coordinates (±SE): 41 (1), −48(1), −21 (1)). We could also define the left FFA in a subset of 15 participants (average MNI coordinates: −40 (1), −51 (2), −21 (1)) and included this region of interest (ROI) in a separate analysis. We could also identify the left and right OFA and left and right LO using our localiser sequences and have included further exploratory analyses of these data. All ROI coordinates are listed in our open dataset. From these coordinates the BOLD signals evoked during the experimental conditions was extracted using a 2 mm radius sphere, and the peak values were entered into the statistical models.

Magnetic Resonance Images were acquired using a 3-T magnetic resonance (MR) scanner from Siemens. For functional images, a standard T_2-weighted echo-planar imaging (EPI) sequence (35 slices, 10° tilted relative to axial, TR = 2000 msec, echo time (TE) = 30 msec, flip angle 90°, 64 × 64 matrices, in plane resolution 3 mm isotopic voxel size) was used. A high resolution T_1-weighted structural 3D scan was generated using a magnetization-prepared rapid gradient-echo (MP-RAGE; TR = 2300 msec; TE = 3.03 msec; 1 mm isotropic voxel size). For details of pre-processing and statistical analysis see Cziraki et al. (2010). Briefly, the functional images were realigned, normalized to the MNI-152 space, resampled to 2 × 2 × 2 mm resolution and spatially smoothed with a Gaussian kernel of 8 mm FWHM (SPM12, Welcome Department of Imaging Neuroscience, London, UK). A general linear model was specified, using the different conditions, as well as six movement parameters.

2.6. *Statistical analyses*

Data and code required to reproduce all analyses will be available at https://osf.io/akygb/ at the time of publication. The conditions of our ethics approval do not permit sharing of the raw MRI data supporting this study with any individual outside the author team under any circumstances. Statistical analyses were performed using Statistica (StatSoft) and JASP v0.9.1 (JASP Team). Accuracy rates and mean response times for trials with correct responses during the training sessions were analysed using 4 × 2 × 2 repeated measures ANOVAs with the factors of session (1, 2, 3, 4), context (AB, AX) and trial type (Rep, Alt). Additionally, we analysed the mean response times and accuracy rates for the fMRI session using a 2 × 2 repeated measures ANOVA with the factors of context (AB, AX) and trial type (Rep, Alt). Peak BOLD signal values were analysed using a 2 × 2 repeated measures ANOVA with the factors context (AB, AX) and trial type (Rep, Alt). For all ANOVA models, Greenhouse–Geisser corrections were applied in cases where Mauchly tests indicated violations of sphericity. Prior to analyses, one outlier dataset in the right FFA data that had a very noisy hemodynamic response function and an unusually large repetition enhancement effect (VP17 in the openly available dataset) was removed. The Holm–Bonferroni method was used to correct for multiple comparisons for follow-up tests after findings of statistically significant main effects and interactions.

We additionally performed exploratory whole-brain analyses corresponding to each main effect as well as the interaction specified in the ANOVA models. The analysis methods and results are detailed in the Supplementary Material.

3. Results

3.1. *Behavioral results*

A significant main effect of session was found for response times ($F_{(1.37,\ 21.96)} = 28.21$, $p < .001$, $\eta^2_p = .64$). Because the Mauchly test of sphericity revealed unequal variances of differences in the four-level factor session ($\chi^2(5) = 29.85$, $p < .001$), Greenhouse–Geisser corrected values are reported. Participants gradually became faster at responding across sessions, with significant differences between session 1 ($M = 647$ msec, $SE = 57$ msec) and session 2 ($M = 583$ msec, $SE = 48$ msec; $p < .001$), session 2 and session 3 ($M = 552$ msec, $SE = 48$ msec; $p = .006$) as well as session 3 and session 4 ($M = 533$ msec, $SE = 44$ msec; $p = .017$). Please note that three participants were only able to complete three of the four scheduled training sessions prior to the fMRI session. There was also a main effect of trial type ($F_{(1,16)} = 27.46$, $p < .001$, $\eta^2_p = .63$). Participants responded faster in Rep trials ($M = 560$ msec, $SE = 68$ msec) as compared to Alt trials ($M = 598$ msec, $SE = 66$ msec), showing a behavioural priming effect (Olkkonen et al., 2017). For response times no other main effects or interactions were statistically significant (main effect of context: $F_{(1,16)} = 2.53$, $p = .131$, $\eta^2_p = .14$, context by trial type interaction: $F_{(1,16)} < .001$, $p = .986$, $\eta^2_p < .001$, context by trial type by session interaction: $F_{(3,\ 48)} = .40$, $p = .755$, $\eta^2_p = .02$).

The analysis of response times during the fMRI scanning session revealed no significant effects. Only a tendency for a faster responses to repetition trials ($M = 545$ msec, $SE = 24$ msec) compared to alternation trials ($M = 562$ msec, $SE = 19$ msec; $F_{(1,19)} = 3.47$, $p = .078$, $n^2_p = .15$) could be found. There was no significant effect of context ($F_{(1,19)} = .15$, $p = .706$, $n^2_p = .01$) nor a context by trial type interaction ($F_{(1,19)} = .58$, $p = .457$, $n^2_p = .03$). Descriptive data showed that response times during the scanning session ($M = 553$ msec, $SE = 23$ msec) were comparable to those from later training sessions.

Analyses of accuracy rates during the training sessions revealed a main effect of trial type ($F_{(1,16)} = 4.52$, $p = .049$, $\eta^2_p = .22$) with a small performance advantage for repetition ($M = 95.9\%$, $SE = 2.3\%$) than for alternation trials ($M = 93.8\%$, $SE = 2.9\%$). No other main effects or interactions were statistically significant (main effect of session: $F_{(1.50,\ 24.06)} = 1.28$, $p = .287$, $\eta^2_p = .07$, main effect of context: $F_{(1,\ 16)} = .15$, $p = .708$, $\eta^2_p = .01$, trial type by context interaction: $F_{(1,\ 16)} = 2.09$, $p = .167$, $\eta^2_p = .12$, session by trial type by context interaction: $F_{(3,\ 48)} = .58$, $p = .634$, $\eta^2_p = .04$). The results of the scanning session did not show any significant effects (all p's > .5). Still, the overall performance ($M = 95.2\%$, $SE = 1.3\%$) showed that participants performed the task correctly.

3.2. *Neuroimaging results*

We performed a two-by-two repeated measures ANOVA with factors context (AB, AX) and trial type (Rep, Alt) on peak BOLD signals in the bilateral FFA (data shown in Fig. 2), bilateral OFA (Fig. 3) and bilateral LO (Fig. 4). In the right FFA (Fig. 2A and B), there was a significant RS effect; stimuli in Rep trials evoked smaller BOLD signals ($M = .65$ percent signal change, $SE = .08$) compared to those in Alt trials ($M = .71$, $SE = .08$; main effect of trial type, $F_{(1,17)} = 17.65$, $p < .001$, $\eta^2_p = .51$).

We also found an interaction between context and trial type in the rFFA ($F_{(1,17)} = 5.49$, $p = .032$, $\eta^2_p = .24$). Plotting this interaction effect revealed larger RS magnitude in the AX context (mean repetition − alternation difference = −.095, $SE = .022$) as compared to the AB context ($M = -.022$, $SE = .020$, shown in Fig. 2B). Additionally, there appeared to be a larger magnitude effect of context on Alt trials, with AX Alt trials evoking larger BOLD signals than AB trials (mean AB − AX context effect = −.046, $SE = .018$). In contrast, BOLD signals for AB and AX Rep trial differences did not differ as much ($M = .027$, $SE = .020$). Altogether, these results suggest that the extent of the observed RS largely depends on the signal magnitude of the Alt trials and this, in turn, is reduced by prior associations of S1 and S2.

As we could identify the left FFA using our localiser sequences in a subset of participants, we included this ROI in an exploratory analysis (data shown in Fig. 2C and D). We found an RS effect (main effect of trial type, $F_{(1,14)} = 12.16$, $p = .004$,

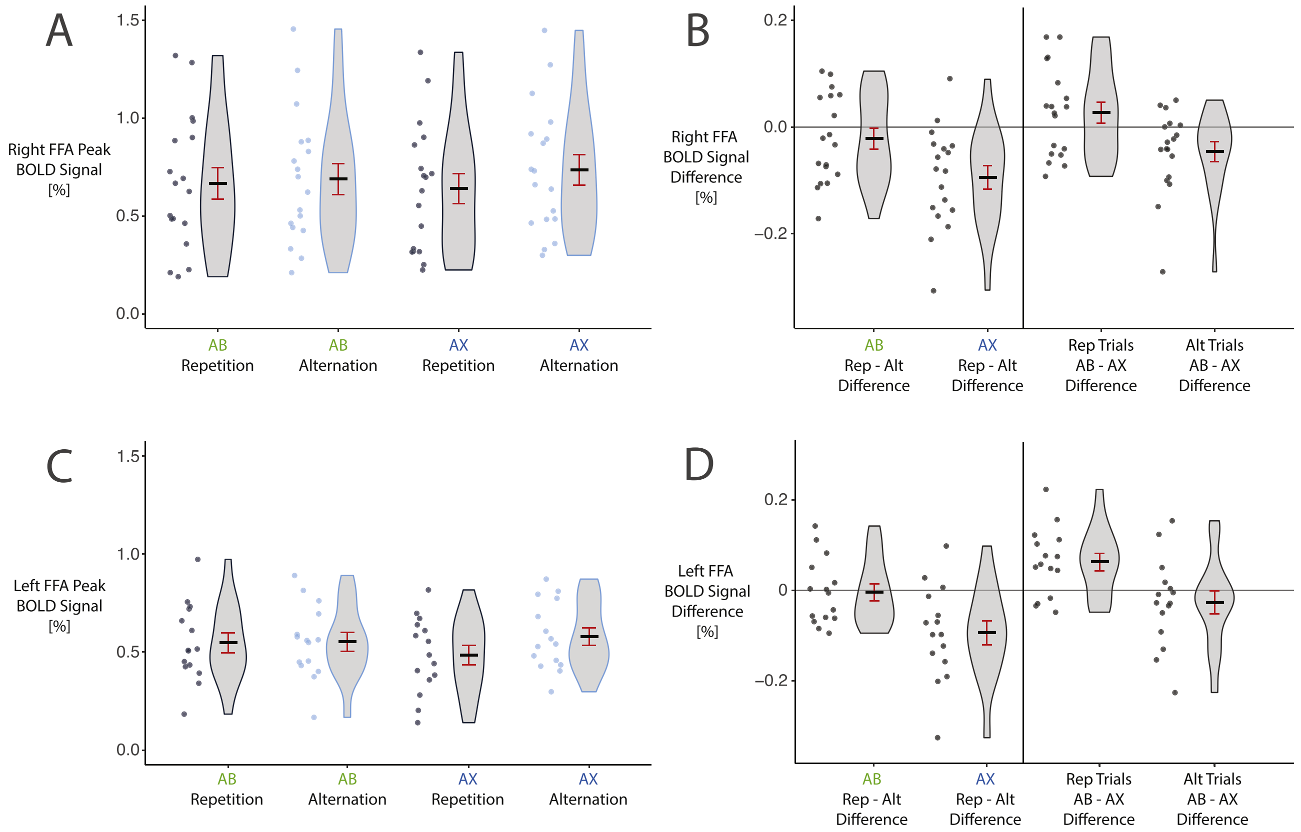


Fig. 2 – BOLD signal results for the right FFA (top row) and left FFA (bottom row). A, C) Peak BOLD signals for each Rep/Alt and AB/AX condition. Dots represent individual data points. Black lines represent group means. Error bars depict standard errors of the mean. Shaded areas depict the distributions of data for each condition. B, D) Repetition and context effects. Repetition effect (Rep − Alt) magnitudes for each context are shown in the left panels. Differences in BOLD signals by AB/AX context are displayed for repetition and alternation trials in the right panels.

$\eta_p^2 = .47$), but no main effect of context ($F_{(1,14)} = 1.87$, $p = .194$, $\eta_p^2 = .12$). There was also a statistically significant interaction, with larger repetition effects observed in AX contexts ($F_{(1,14)} = 5.86$, $p = .030$, $\eta_p^2 = .30$).

Notably, we did not observe statistically significant RS effects in our sample in the AB context, for the right FFA ($t(17) = -1.10$, uncorrected $p = .287$) or left FFA ($t(14) = -.23$, uncorrected $p = .819$, mean Rep − Alt difference $= -.004$, $SE = .019$), whereas we did find significant RS effects in the AX context (right FFA: -4.31, uncorrected $p < .001$; left FFA: $t(14) = -3.49$, uncorrected $p = .004$, $M = -.094$, $SE = .027$).

We conducted additional exploratory analyses using data from bilateral OFA and bilateral LO ROIs. We could identify the right OFA in $n = 19$ participants and the left OFA in $n = 18$ participants. We first performed a two-by-two repeated measures ANOVA with factors context (AB, AX) and trial type (Rep, Alt) on peak BOLD signals in the right OFA (data shown in Fig. 3A and B). There was a significant RS effect whereby stimuli in Rep trials evoked smaller BOLD signals compared to those in Alt trials ($F_{(1,\ 18)} = 6.14$, $p = .023$, $\eta_p^2 = .25$). There was no significant effect of context ($F_{(1,\ 18)} = .42$, $p = .528$, $\eta_p^2 = .02$), nor a significant context by trial type interaction ($F_{(1,\ 18)} = .18$, $p = .676$, $\eta_p^2 = .01$).

For analyses of the left OFA (data shown in Fig. 3C and D) there were no statistically significant effects (main effect of trial type: $F_{(1,\ 17)} = 3.91$, $p = .064$, $\eta_p^2 = .19$; main effect of context, $F_{(1,\ 17)} = 1.27$, $p = .275$, $\eta_p^2 = .07$; trial type by context interaction, $F_{(1,\ 17)} = .67$, $p = .424$, $\eta_p^2 = .04$).

We could identify the right LO in $n = 20$ participants and the left LO in $n = 20$ participants. We first performed a two-by-two repeated measures ANOVA with factors context (AB, AX) and trial type (Rep, Alt) on peak BOLD signals in the right LO (data shown in Fig. 4A and B). There was no significant RS effect ($F_{(1,\ 19)} = .18$, $p = .678$, $\eta_p^2 = .01$). There was no significant effect of context ($F_{(1,\ 19)} = .63$, $p = .429$, $\eta_p^2 = .03$), nor a significant context by trial type interaction ($F_{(1,\ 19)} = .65$, $p = .431$, $\eta_p^2 = .03$).

As we found for the right FFA results, there was an unusually large repetition enhancement effect and a noisy hemodynamic response function for a certain participant (see the very large positive AX context rep-alt difference in Fig. 4B). We treated this participant as an outlier and repeated our analyses excluding this participant and found similar results. The main effect of repetition still did not quite reach our p-value threshold for significance ($F_{(1,\ 18)} = 4.34$, $p = .052$, $\eta_p^2 = .19$). There were also no significant effects of context ($F_{(1,}$

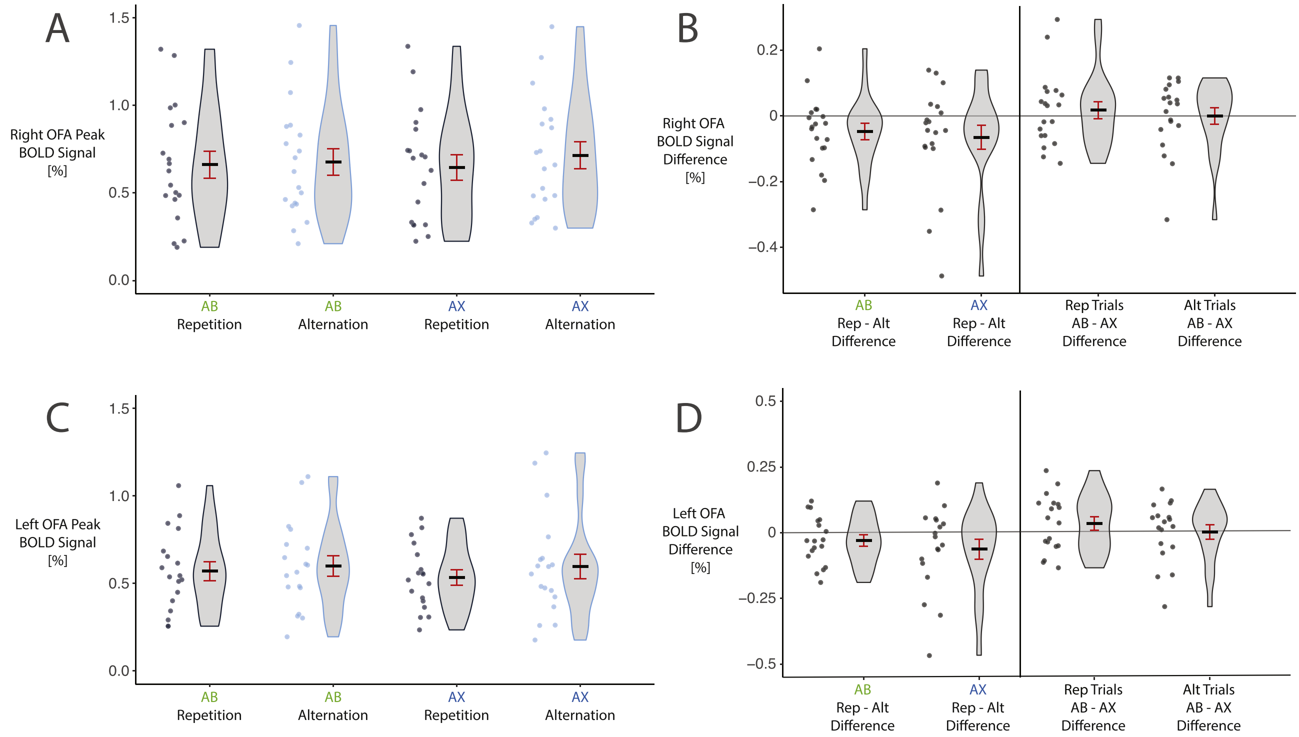


Fig. 3 – BOLD signal results for the right OFA (top row) and left OFA (bottom row). A, C) Peak BOLD signals for each Rep/Alt and AB/AX condition. Dots represent individual data points. Black lines represent group means. Error bars depict standard errors of the mean. Shaded areas depict the distributions of data for each condition. B, D) Repetition and context effects. Repetition effect (Rep – Alt) magnitudes for each context are shown in the left panel. Differences in BOLD signals by AB/AX context are displayed for repetition and alternation trials in the right panel.

$_{18)} = 1.42, p = .249, \eta^2_p = .07$) nor a trial type by context interaction ($F_{(1,\ 18)} = .23, p = .640, \eta^2_p = .01$).

For analyses of the left LO (data shown in Fig. 4C and D) there were no statistically significant effects (main effect of trial type: $F_{(1,\ 19)} = 1.80, p = .195, \eta^2_p = .09$; main effect of context, $F_{(1,\ 19)} = 1.30, p = .269, \eta^2_p = .06$; trial type by context interaction, $F_{(1,\ 19)} = .83, p = .375, \eta^2_p = .04$).

Additionally, we performed an exploratory second-level analysis with all 20 subjects. A summary of the whole-brain analysis results can be found in the supplementary material.

4. Discussion

To investigate the interplay between repetition and expectation effects, we presented pairs of faces which could either repeat or alternate within a trial, in two different contexts. In one context, the alternating faces were chosen randomly and were therefore unpredictable, while in the other context the second face in alternating trials could be predicted after seeing the first, due to previously learned transitional rules and contingencies. In both contexts, the repeated stimuli were predictable. We found face repetition-related reductions of BOLD signals in the left and right FFA and in the right OFA, consistent with a large body of work (for a review see Grill-Spector et al., 2006). We also report that responses to alternating stimuli differed markedly, depending on the context; unpredictable stimulus pairs (in the AX context) evoked larger BOLD signals than those which were predictable (in the AB context) in the FFA. This in turn modulated the measured repetition-alternation signal differences that typically defines the measurement of RS, and even determined whether or not we found statistically-significant RS effects in our sample. Here, the point estimate of RS magnitude in the right FFA was over four times as large in the AX ($M = .095$) compared to AB ($M = .022$) contexts. Our results demonstrate that stimulus predictability effects can substantially inflate conventional measures of RS, or even mimic the effect of stimulus repetition, when predictability is not equated between repeated and alternating stimuli. While it seems unlikely that all prior reports of RS could be fully explained by differences of stimulus predictability, this effect has likely inflated repetition effect sizes in a large number of existing studies.

Our results are in line with those of Pajani et al. (2017) and Feuerriegel, Churches, et al. (2018), who found similar effects of stimulus predictability using BOLD and ERP measures, respectively. Importantly, our design also controlled for effects of stimulus novelty across predictable and unpredictable contexts, which could have produced the patterns of effects seen in their experiments (Mur et al., 2010; Xiang & Brown, 1998). In their studies, the alternating stimuli in AB-type conditions were presented many times to the participants, yet the alternating stimuli in AX-type conditions were presented much more rarely (Feuerriegel, Churches, et al., 2018)

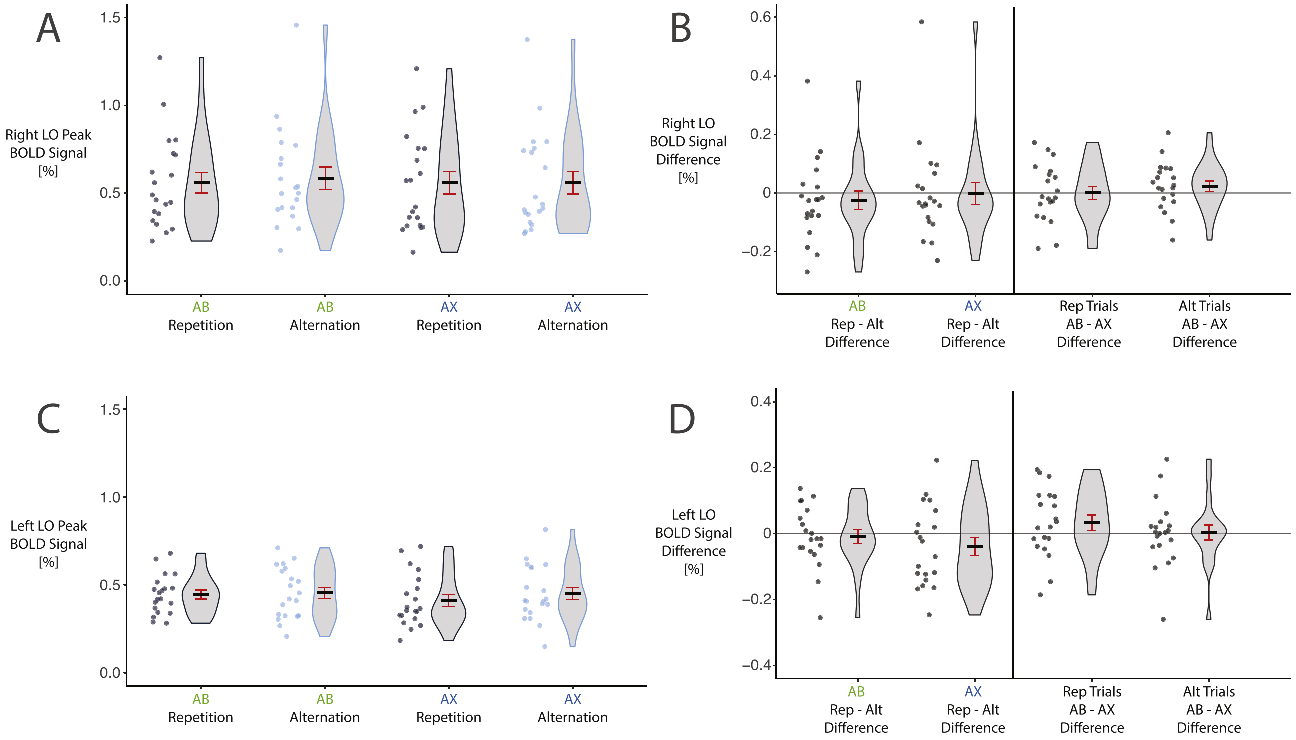


Fig. 4 – BOLD signal results for the right LO (top row) and left LO (bottom row). A, C) Peak BOLD signals for each Rep/Alt and AB/AX condition. Dots represent individual data points. Black lines represent group means. Error bars depict standard errors of the mean. Shaded areas depict the distributions of data for each condition. B, D) Repetition and context effects. Repetition effect (Rep – Alt) magnitudes for each context are shown in the left panel. Differences in BOLD signals by AB/AX context are displayed for repetition and alternation trials in the right panel.

or only once in the experiment (Pajani et al., 2017). By contrast, we presented each face image in the AB and AX contexts an equal number of times, thereby replicating their findings while also controlling for effects of novelty.

Notably, effects of stimulus predictability seem to be consistent across hemodynamic and electrophysiological measures, in contrast to effects of repetition probability manipulations as used in Summerfield et al. (2008) and subsequent replications. The effects of stimulus predictability seen here resemble expectations derived through statistical learning of image transition probabilities (as seen in single-cell recording measurements by Meyer & Olson, 2011), produced by the pairing of specific images, rather than more abstract expectations about whether a stimulus will repeat or not. These types of expectations appear to be qualitatively different to expectations pertaining to more abstract sequences of stimuli, and there is some evidence that these two have interacting effects on neural responses (Costa-Faidella et al., 2011; Feuerriegel, Keage, et al., 2018; Mittag et al., 2016).

Similar to the EEG study of Feuerriegel, Churches, et al. (2018), we observed that these context effects predominantly acted upon responses to alternating rather than repeated stimuli. This indicates that stimulus predictability selectively influenced responses to alternating stimuli, which does not modulate the underlying mechanisms of RS per se, but does influence how it is measured in commonly-used immediate repetition designs (Grill-Spector et al., 2006). This pattern of results also suggests that the violation of image-specific expectations (i.e., surprise) may underlie the observed predictability effects, and be responsible for BOLD signal increases in AX alternating trials. In our design, the likelihood of each trial type (AB-Rep, AB-Alt, AX-Rep and AX-Alt) was equated throughout the experiment. However, the relative likelihoods of the appearance of specific face images in each context were not. For example, in AB trials the S2 face could either be a repetition of S1, or a specific different face identity, with a probability ratio of 1:1. In contrast, after seeing S1 in the AX trials, an image repetition would occur 50% of the time, yet each of the 5 possible alternating face images could each appear with a probability of 10%, leading to a probability ratio of 5:1. If participants' expectations depended on the relative appearance probabilities of specific images, then this would lead to expectations more strongly weighted toward repetitions in AX contexts, and larger surprise-related BOLD increases following AX alternating stimuli. According to this interpretation, one might also expect to see similar magnitude suppression of BOLD signals for AX repetition trials, reflecting ES, whereas we observed larger context effects for alternating trials. This may be because surprise seems to have a larger effect on neural responses than fulfilled expectations (Amado et al., 2016; Kovács & Vogels, 2014). In addition, there is evidence that effects of fulfilled expectations and surprise are diminished for repeated stimuli (reviewed in Feuerriegel, Keage, et al., 2018). So, it appears that surprise-related response enhancement in AX alternating trials may have played an important role in inflating measures of RS.

We caution that our findings should not be interpreted as that repetition effects in general are simply due to a stimulus predictability effect. Previous experiments using AB-type designs and stimulus associations have reported repetition effects (Feuerriegel, Churches, et al., 2018, 2019; Pajani et al., 2017; Todorovic & de Lange, 2012). In fact, one of the earliest mentions of RS in macaques was from the seminal study of Gross et al. (1979), using an AB-type design, with associated stimuli and an S1–S2 matching task. In addition, we note that the RS effects in our study may not be strictly localized to the FFA, and may partly index inherited effects due to RS in regions early in the visual stream, such as V1, providing altered input to higher-level regions. Such 'inherited adaptation' effects (Kohn, 2007) have been widely documented (reviewed in Feuerriegel, 2016; Larsson et al., 2016) and small size changes between S1 and S2 would not fully control for such effects, given the large receptive field sizes that are present in areas earlier than the FFA in the visual hierarchy. A recent optogenetic study has cast doubt on the notion that RS is locally generated in IT (Fabbrini et al., 2019), and so it remains to be seen what the magnitude of RS effects would be when controlling for both inherited adaptation and stimulus predictability. An investigation of RS in this context should aim for higher precision (i.e., more trials per participant, or a larger sample size) than in the current study and most previous studies of RS. This is because RS, which is usually a very robust effect, was not even statistically significant in the AB context in our sample, suggesting that the true magnitudes of 'true' RS effects may be smaller than previously assumed.

While our findings do not provide strong evidence for or against predictive coding models that incorporate the notion of sensory precision (e.g., Auksztulewicz & Friston, 2016), it does appear that expectations can account for a proportion of repetition effects observed in many experiments. Results of recent experiments have not provided clear support for precision-based predictive coding models of RS (e.g., Amado et al., 2016; Rostalski et al., 2019; Vinken et al., 2018) and further tests of key model predictions are needed. While RS can be conceptualized as reflecting a strong prior belief towards stimuli encountered in the immediate past, it is still unclear exactly how RS fits within the broader taxonomy of expectation-related phenomena.

RS, ES and predictability effects are often difficult to disentangle in experimental designs. Convergent evidence across recording modalities and experimental manipulations will be critical for separating and characterizing each contribution to repetition effects. For example, electrophysiological recordings with high temporal resolution have been useful for identifying distinct effects of repetition and expectation that occur over multiple time windows following stimulus onset (e.g., Feuerriegel, Churches, et al., 2018), which are likely to be conflated when measuring BOLD signals. Methods targeting specific aspects of local neural circuit activity may also be useful, such as those employing optogenetic techniques (e.g., Fabbrini et al., 2019). Given that ES appears to be highly dependent on participants attending to the stimuli of interest (e.g., Larsson & Smith, 2012; Smout et al., 2019) assessing the time windows and measures of neural activity that are (and are not) sensitive to attention may also help partition the different phenomena that contribute to repetition effects as measured in neuroimaging experiments.

Our results should be interpreted with the following caveats in mind. First of all, our study used an immediate repetition design, and our results may not be generalizable to RS as measured using delayed repetition paradigms, in which a number of different intervening stimuli are presented between the first and repeated presentations of a given image (reviewed in Henson, 2016). Although predictive coding models encompass both types of repetition effects (e.g., Auksztulewicz & Friston, 2016) it is likely that these rely on different sets of neural mechanisms (Epstein et al., 2008; Weiner et al., 2010). It remains unclear whether these should be captured within a unifying framework, or if different sets of underlying mechanisms produce similar effects in each type of repetition design.

Second, we emphasize that the repetition effects observed in the current study likely index effects of adaptation of low- and mid-level features that may be inherited by higher level regions in the visual system (Kohn, 2007). Consequently, our results do not provide insights into the tuning or organization of neurons within ventral temporal cortex that may encode face identities. Several studies have yielded insights into the organization of such tuning schemes using stimulus repetition and adaptation designs (e.g., Drucker & Aguirre, 2009; Leopold et al., 2006; Loffler et al., 2005). However, our results do warrant caution in interpreting results of experiments in which the predictability of stimulus images systematically differed across conditions (e.g., Loffler et al., 2005).

Also, in this experiment we required participants to indicate whether each trial contained a repetition or alternation. This may have led to differences in response strategies for repeated and alternating stimuli. For example, some participants may have made their judgements by first determining whether the first and second stimuli match, leading to faster RTs in repetition trials. It is unclear whether this potential difference in strategies also led to a portion of the observed RS effects in our study. However, we believe that task-related effects are not responsible for the observed predictability effects on RS, given that we did not find a difference in RTs across predictability conditions, and that previous studies have found similar interactions using different tasks (Feuerriegel, Churches, et al., 2018; Pajani et al., 2017).

In addition, we did not find differences in mean RTs and accuracy scores across AB and AX conditions, despite extensive training and exposure to the stimulus pairings. This is despite our findings of stimulus repetition effects on the speed of responding. Because of this, it is unclear whether the predictability effects found in our neuroimaging results were actually used for decision making during the task. Validly-cued expectancies for certain stimuli have led to faster responses in previous studies (Hall et al., 2018; Mulder et al., 2012). However, these experiments have typically conflated expectations to see a certain stimulus with expectations to make the motor action required to report the perceptual decision corresponding to that stimulus (Gold & Stocker, 2017). For example, in a left/right motion discrimination task, a cued expectation to see a leftward-moving stimulus co-occurs with an expectation to press the response button associated with leftward motion. This bias in motor action preparation, which can even precede the onset of a decision-relevant stimulus, is likely to account for a substantial portion of expectation-related effects on RTs (e.g., Bogacz et al., 2010; de Lange et al., 2013). In our task there were no cued biases toward a particular button response, and participants' expectations for how to respond were balanced across AB and AX contexts. In other words, participants' expectations to observe a repetition or alternation was not biased as the probability of each trial type was kept constant at 50%. This may be why we did not observe predictability effects on behaviour.

More generally, the relationships between decision-making performance and expectation and repetition effects have not been clearly defined. Recent studies have cast doubt on the idea that contextual expectations affect sensory representations that are used for perceptual decision making, particularly when controlling for effects of feature-selective attention (Bang & Rahnev, 2017; Rungratsameetaweemana et al., 2018). Repetition priming effects on RTs also tend to be more consistently associated with RS that occurs in frontal regions rather than in the visual system (reviewed in Horner, 2012; Wig, 2012; Schacter et al., 2012). It is possible that the RS and predictability effects observed in our study are related to particular types of learning and behaviour, but are not reflected in accuracy or mean RT measures in discrete choice perceptual decision tasks (Wig, 2012).

An additional point is that our study included multiple training sessions before fMRI scanning. These sessions were included so that participants had ample opportunity to learn the stimulus image transition probabilities for alternating stimuli in the AB and AX blocks. Without a training period, the participants could not learn to form image-specific expectations for alternating face images in our experiment. The face images would have become highly familiar to participants during the training sessions, in contrast to relatively novel or trial-unique face stimuli used in many previous studies (e.g., Amado et al., 2016; Summerfield et al., 2008). Face familiarity does influence how faces are encoded (Johnston & Edmonds, 2009; Ramon & Gobbini, 2018; Young & Burton, 2017) which may consequently alter the magnitude of RS (e.g., see Henson, 2016 for strong evidence of familiarity effects in delayed repetition designs). For example, it is possible that neural population level representations of each face identity became more distinct with familiarization (Freedman et al., 2006; Meyer et al., 2014) leading to larger observed RS (Verhoef et al., 2008; De Baene & Vogels, 2010). It is unclear whether this number of training sessions is required for developing image-specific expectations, and the study of Pajani et al. (2017) used a much shorter training protocol. Future work could investigate whether transition probabilities could be learned more rapidly than the familiarization process, and whether similar predictability effects are observed for relatively novel stimuli.

5. Conclusion

We have shown that, in immediate repetition designs, an observer's capacity to predict the image of repeated compared to unrepeated stimuli has a substantial effect on the observed magnitude of RS. While this does not necessarily mean that RS is best accounted for by predictive coding models, it does indicate that measures of repetition effects have likely been inflated due to this confound in a very large number of previous studies, including those run within our own labs. We also highlight stimulus predictability as an important, yet commonly overlooked, factor to consider when investigating the hierarchy of expectation effects implemented within the visual system.

Author contributions

Sophie-Marie Rostalski: investigation, project administration, formal analysis, writing – original draft. **Catarina Amado**: conceptualization, methodology, software, investigation. **Gyula Kovács**: conceptualization, methodology, resources, writing – review & editing, supervision, project administration, funding acquisition. **Daniel Feuerriegel**: Conceptualization, methodology, software, data curation, formal analysis, writing – review & editing, visualization, supervision.

Declaration of competing interest

The authors declare no competing financial interests.

Acknowledgements

The authors would like to thank Carolin Altmann for providing the stimulus set. This work was supported by a grant from the Deutsche Forschungsgemeinschaft (KO3918/5-1). We also thank the two Reviewers and the Editor for their helpful suggestions and commentary.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2020.07.010.

REFERENCES

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. https://doi.org/10.1016/j.neuropsychologia.2016.02.003

Auksztulewicz, R., & Friston, K. (2016). Repetition suppression and its contextual determinants in predictive coding. *Cortex, 80*, 125–140. https://doi.org/10.1016/j.cortex.2015.11.024

Bang, J. W., & Rahnev, D. (2017). Stimulus expectation alters decision criterion but not sensory signal in perceptual decision making. *Scientific Reports, 7*, 17072. https://doi.org/10.1038/s41598-017-16885-2

Bogacz, R., Wagenmakers, E.-J., Forstmann, B. U., & Nieuwenhuis, S. (2010). The neural basis of the speed-accuracy tradeoff. *Trends in Neurosciences, 33*(1), 10–16.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433–436. https://doi.org/10.1163/156856897X00357

Costa-Faidella, J., Grimm, S., Slabu, L., Díaz-Santaella, F., & Escera, C. (2011). Multiple time scales of adaptation in the auditory system as revealed by human evoked potentials. *Psychophysiology, 48*, 774–783. https://doi.org/10.1111/j.1469-8986.2010.01144.x

Cziraki, C., Greenlee, M. W., & Kovacs, G. (2010). Neural correlates of high-level adaptation-related aftereffects. *Journal of Neurophysiology, 103*, 1410–1417. https://doi.org/10.1152/jn.00582.2009

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. https://doi.org/10.1093/cercor/bhp277

De Gardelle, V., Waszczuk, M., Egner, T., & Summerfield, C. (2013). Concurrent repetition enhancement and suppression responses in extrastriate visual cortex. *Cerebral Cortex, 23*, 2235–2244. https://doi.org/10.1093/cercor/bhs211

Dhruv, N. T., Tailby, C., Sokol, S. H., & Lennie, P. (2011). Multiple adaptable mechanisms early in the primate visual pathway. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience, 31*, 15016–15025. https://doi.org/10.1523/jneurosci.0890-11.2011

Drucker, D. M., & Aguirre, G. K. (2009). Different spatial scales of shape similarity representation in lateral and ventral LOC. *Cerebral Cortex, 19*(10), 2269–2280. https://doi.org/10.1093/cercor/bhn244

de Lange, F. P., Rahnev, D. A., Donner, T. H., & Lau, H. (2013). Prestimulus oscillatory activity over motor cortex reflects perceptual expectations. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience, 33*(4), 1400–1410.

Eger, E. (2004). BOLD repetition decreases in object-responsive ventral visual areas depend on spatial attention. *Journal of Neurophysiology, 92*, 1241–1247. https://doi.org/10.1152/jn.00206.2004

Egner, T., Monti, J. M., & Summerfield, C. (2010). Expectation and surprise determine neural population responses in the ventral visual stream. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience, 30*, 16601–16608. https://doi.org/10.1523/JNEUROSCI.2770-10.2010

Epstein, R. A., Parker, W. E., & Feiler, A. M. (2008). Two kinds of fMRI repetition suppression? Evidence for dissociable neural mechanisms. *Journal of Neurophysiology, 99*, 2877–2886. https://doi.org/10.1152/jn.90376.2008

Fabbrini, F., Van den Haute, C., De Vitis, M., Baekelandt, V., Vanduffel, W., & Vogels, R. (2019). Probing the mechanisms of repetition suppression in inferior temporal cortex with optogenetics. *Current Biology, 29*, 1988–1998. https://doi.org/10.1016/J.CUB.2019.05.014. e4.

Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience, 4*, 1–23. https://doi.org/10.3389/fnhum.2010.00215

Feuerriegel, D. (2016). Selecting appropriate designs and comparison conditions in repetition paradigms. *Cortex, 80*, 196–205. https://doi.org/10.1016/j.cortex.2015.10.022

Feuerriegel, D., Churches, O., Coussens, S., & Keage, H. A. D. (2018). Evidence for spatiotemporally distinct effects of image repetition and perceptual expectations as measured by event-related potentials. *NeuroImage, 169*, 94–105. https://doi.org/10.1016/j.neuroimage.2017.12.029

Feuerriegel, D., Churches, O., Coussens, S., & Keage, H. A. D. (2019). Temporal expectations modulate face image repetition suppression of early stimulus evoked event-related potentials. *Neuropsychologia, 122*, 76–87. https://doi.org/10.1016/j.neuropsychologia.2018.11.010

Feuerriegel, D., Keage, H. A. D., Rossion, B., & Quek, G. L. (2018). Immediate stimulus repetition abolishes stimulus expectation and surprise effects in fast periodic visual oddball designs. *Biological Psychology, 138*, 110–125. https://doi.org/10.1016/j.biopsycho.2018.09.002

Fiser, J., & Aslin, R. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 15822–15826. https://doi.org/10.1073/pnas.232472899

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*(11), 1631–1644.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*, 815–836. https://doi.org/10.1098/rstb.2005.1622

Gold, J. I., & Stocker, A. A. (2017). Visual decision-making in an uncertain and dynamic world. *Annual Review of Vision Science, 3*, 227–250. https://doi.org/10.1146/annurev-vision-111815-114511

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Science, 10*, 14–23. https://doi.org/10.1016/j.tics.2005.11.006

Gross, C. G., Bender, D. B., & Gerstein, G. L. (1979). Activity of inferior temporal neurons in behaving monkeys. *Neuropsychologia, 17*, 215–229. https://doi.org/10.1016/0028-3932(79)90012-5

Grotheer, M., Hermann, P., Vidnyánszky, Z., & Kovács, G. (2014). Repetition probability effects for inverted faces. *NeuroImage, 102*, 416–423. https://doi.org/10.1016/j.neuroimage.2014.08.006

Grotheer, M., & Kovács, G. (2014). Repetition probability effects depend on prior experiences. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience, 34*, 6640–6646. https://doi.org/10.1523/JNEUROSCI.5326-13.2014

Grotheer, M., & Kovács, G. (2015). The relationship between stimulus repetitions and fulfilled expectations. *Neuropsychologia, 67*, 175–182. https://doi.org/10.1016/j.neuropsychologia.2014.12.017

Hall, M. G., Mattingley, J. B., & Dux, P. E. (2018). Electrophysiological correlates of incidentally learned expectations in human vision. *Journal of Neurophysiology, 119*, 1461–1470. https://doi.org/10.1152/jn.00733.2017

Henson, R. N. (2016). Repetition suppression to faces in the fusiform face area: A personal and dynamic journey. *Cortex, 80*, 174–184. https://doi.org/10.1016/j.cortex.2015.09.012

Horner, A. (2012). Focusing on the frontal cortex. *Cognitive Neuroscience, 3*(3–4), 246–247. https://doi.org/10.1080/17588928.2012.689959

Johnston, R. J., & Edmonds, A. J. (2009). Familiar and unfamiliar face recognition: A review. *Memory, 17*(5), 577–596. https://doi.org/10.1080/09658210902976969

Kaliukhovich, D. A., & Vogels, R. (2011). Stimulus repetition probability does not affect repetition suppression in macaque inferior temporal cortex. *Cerebral Cortex, 21*, 1547–1558. https://doi.org/10.1093/cercor/bhq207

Kaliukhovich, D. A., & Vogels, R. (2014). Neurons in macaque inferior temporal cortex show no surprise response to deviants in visual oddball sequences. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience, 34*, 12801–12815. https://doi.org/10.1523/jneurosci.2154-14.2014

Kaliukhovich, D. A., & Vogels, R. (2016). Divisive normalization predicts adaptation-induced response changes in macaque inferior temporal cortex. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience, 36*, 6116–6128. https://doi.org/10.1523/jneurosci.2011-15.2016

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience, 17*, 4302–4311. https://doi.org/10.1098/Rstb.2006.1934

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception, 36*(14), 1–16.

Kohn, A. (2007). Visual adaptation: Physiology, mechanisms, and functional benefits. *Journal of Neurophysiology, 97*(5), 3155–3164. https://doi.org/10.1152/jn.00086.2007

Kovács, G., & Vogels, R. (2014). When does repetition suppression depend on repetition probability? *Frontiers in Human Neuroscience, 8*(September), 685. https://doi.org/10.3389/fnhum.2014.00685

Kronbichler, L., Said-yürekli, S., & Kronbichler, M. (2018). Perceptual expectations of object stimuli modulate repetition suppression in a delayed repetition design, 1–8. https://doi.org/10.1038/s41598-018-31091-4.

Larsson, J., & Smith, A. T. (2012). fMRI repetition suppression: Neuronal adaptation or stimulus expectation? *Cerebral Cortex, 22*, 567–576. https://doi.org/10.1093/cercor/bhr119

Larsson, J., Solomon, S. G., & Kohn, A. (2016). fMRI adaptation revisited. *Cortex, 80*, 154–160. https://doi.org/10.1016/j.cortex.2015.10.026

Leopold, D. A., Bondar, I. V., & Giese, M. A. (2006). Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature, 442*(7102), 572–575. https://doi.org/10.1038/nature04951

Loffler, G., Yourganov, G., Wilkinson, F., & Wilson, H. R. (2005). fMRI evidence for the neural representation of faces. *Nature Neuroscience, 8*(10), 1386–1390. https://doi.org/10.1038/nn1538

Mayrhauser, L., Bergmann, J., Crone, J., & Kronbichler, M. (2014). Neural repetition suppression: Evidence for perceptual expectation in object-selective regions. *Frontiers in Human Neuroscience, 8*, 1–8. https://doi.org/10.3389/fnhum.2014.00225

Meyer, T., & Olson, C. R. (2011). Statistical learning of visual transitions in monkey inferotemporal cortex. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 19401–19406. https://doi.org/10.1073/pnas.1112895108

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.

Mittag, M., Takegata, R., & Winkler, I. (2016). Transitional probabilities are prioritized over stimulus/pattern probabilities in auditory deviance detection: Memory basis for predictive sound processing. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience, 36*, 9572–9579. https://doi.org/10.1523/JNEUROSCI.1041-16.2016

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. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience, 32*, 2335–2343. https://doi.org/10.1523/JNEUROSCI.4156-11.2012

Murray, S. O., & Wojciulik, E. (2004). Attention increases neural selectivity in the human lateral occipital complex. *Nature Neuroscience, 7*, 70–74. https://doi.org/10.1038/nn1161

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. https://doi.org/10.1093/cercor/bhp272

Olkkonen, M., Aguirre, G. K., & Epstein, R. A. (2017). Expectation modulates repetition priming under high stimulus variability. *17*, 1–16. https://doi.org/10.1167/17.6.10

Pajani, A., Kouider, S., Roux, P., & de Gardelle, V. (2017). Unsuppressible repetition suppression and exemplar-specific expectation suppression in the fusiform face area. *Scientific Reports, 7*, 160. https://doi.org/10.1038/s41598-017-00243-3

Ramon, M., & Gobbini, M. I. (2018). Familiarity matters: A review on prioritized processing of personally familiar faces. *Visual Cognition, 26*(3), 179–195. https://doi.org/10.1080/13506285.2017.1405134

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. https://doi.org/10.1038/4580

Romberg, A. R., & Saffran, J. R. (2011). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*, 906–914. https://doi.org/10.1002/wcs.78.Statistical

Rostalski, S.-M., Amado, C., & Kovács, G. (2019). Repetition suppression for noisy and intact faces in the occipito-temporal cortex. *Frontiers in Psychology, 10*, 1348. https://doi.org/10.3389/fpsyg.2019.01348

Rungratsameetaweemana, N., Itthipuripat, S., Salazar, A., & Serences, J. T. (2018). Expectations do not alter early sensory processing during perceptual decision-making. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience, 38*, 5632–5648. https://doi.org/10.1523/JNEUROSCI.3638-17.2018

Schacter, D. L., Wig, G. S., & Stevens, W. D. (2012). Reductions in cortical activity during priming. *Current Opinion in Neurobiology, 17*, 171–176.

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), Article e3000368.

Solomon, S. G., & Kohn, A. (2014). Moving sensory adaptation beyond suppressive effects in single neurons. *Current Biology, 24*, R1012–R1022. https://doi.org/10.1016/j.cub.2014.09.001

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Science, 13*, 403–409. https://doi.org/10.1016/j.tics.2009.06.003

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. https://doi.org/10.1038/nn.2163

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. https://doi.org/10.3389/fnhum.2011.00067

Todorovic, A., & de Lange, F. P. (2012). Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience, 32*, 13389–13395. https://doi.org/10.1523/JNEUROSCI.2227-12.2012

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*, 1934–1945. https://doi.org/10.1162/jocn.2009.21131

Valdez, P., Ramírez, C., García, A., Talamantes, J., & Cortez, J. (2010). Circadian and homeostatic variation in sustained attention. *Chronobiology International, 27*, 393–416. https://doi.org/10.3109/07420521003765861

Verhoef, B. E., Kayaert, G., Franko, E., Vangeneugden, J., & Vogels, R. (2008). Stimulus similarity-contingent neural adaptation can be time and cortical area dependent. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience, 28*(42), 10631–10640. https://doi.org/10.1523/JNEUROSCI.3333-08.2008

Vinken, K., Op de Beeck, H., & Vogels, R. (2018). Face repetition probability does not affect repetition suppression in macaque inferotemporal cortex. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience, 38*(34), 7492–7504. https://doi.org/10.1523/JNEUROSCI.0462-18.2018, 0462-18.

Weiner, K. S., Sayres, R., Vinberg, J., & Grill-Spector, K. (2010). fMRI-adaptation and category selectivity in human ventral temporal cortex: Regional differences across time scales. *Journal of Neurophysiology*, *103*, 3349–3365. https://doi.org/10.1152/jn.01108.2009

Wig, G. (2012). Repetition suppression and repetition priming are processing outcomes. *Cognitive Neuroscience*, *3*(3–4), 247–248. https://doi.org/10.1080/17588928.2012.689964

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.

Young, A. W., & Burton, A. M. (2017). Recognizing faces. *Current Directions in Psychological Science*, *26*(3). https://doi.org/10.1177/0963721416688114

Measures of repetition suppression in the Fusiform Face Area are inflated by co-occurring effects of statistically learned associations: Supplementary material

Sophie-Marie Rostalski [a], Catarina Amado [b], Gyula Kovács [a,+++], Daniel Feuerriegel [c,+++]

[a] Institute of Psychology, Friedrich Schiller University Jena, Jena, Germany
[b] Experimental Cognitive Science, Eberhard Karls Universität Tübingen, Tübingen, Germany
[c] Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, Australia
+++ These two authors contributed to the publication equally as senior authors

Corresponding author: Sophie-Marie Rostalski, Institute of Psychology, Friedrich Schiller University Jena, Germany. E-mail address: sophie-marie.rostalski@uni-jena.de

1. Methods and Results For the Whole-Brain Analyses

To test whether effects of repetition or predictability were observed in other areas than the predefined ROIs, we performed a group-level whole-brain analysis testing for the main effect of context, the main effect of trial type and a context x trial type interaction. Therefore, we created contrasts for every experimental condition for all 20 participants and submitted those to a flexible design group level analysis with conditions subject, context and trial type. We used a threshold of $p < 0.05$ (FWE) with a minimum cluster size of > 50 voxels. All statistically significant clusters of effects are listed in Supplementary Table 1. Locations of statistically significant clusters with the strongest activation for each contrast, as well as an additional activation in the FFA are visually depicted in Supplementary Figure 1.

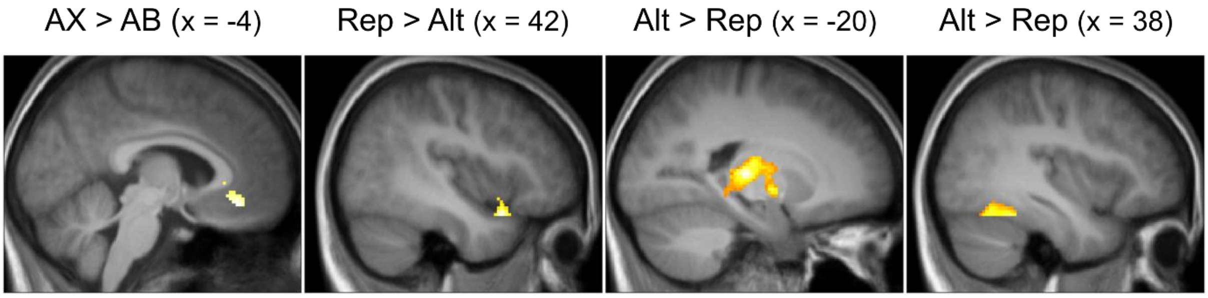
AX trials evoked larger BOLD signals than AB trials at four clusters located within Brodmann Area (BA) 11, BA 23, BA 39 and BA 32. When testing for voxels with larger BOLD signals in the AB compared to AX trials, no statistically significant clusters were found.

Repetition trials evoked larger BOLD signals compared to alternation trials (Rep > Alt) across three clusters located within BA 38, BA 6 and the Pulvinar. Three clusters of voxels exhibiting the opposite pattern (Alt > Rep) were identified within the Pulvinar and BA 21 and a cluster in the fusiform gyrus (BA 37).

No statistically significant clusters were found when testing for the interaction of context and trial type.

Supplementary Table 1. Summary of significant activations identified from the group level whole-brain analysis

CONTRAST	**BRAIN REGION**	**COORDINATES**	**CLUSTER SIZE (VOXELS)**	**THRESHOLD**
AX > AB	Medial frontal gyrus; Brodmann area 11	-4, 36, -14	89	P < 0.05, FWE
	Posterior cingulate; Brodmann area 23	4, -58, 14	217	P < 0.05, FWE
	Angular gyrus; Brodmann area 39	48, -72, 30	63	P < 0.05, FWE
	Anterior cingulate; Brodmann area 32	14, 48, -4	107	P < 0.05, FWE
REP > ALT	Superior temporal gyrus; Brodmann area 38	42, 12, -20	60	P < 0.05, FWE
	Superior frontal gyrus; Brodmann area 6	24, 10, 54	403	P < 0.05, FWE
	Pulvinar	-12, -28, 16	86	P < 0.05, FWE
ALT > REP	Pulvinar	-20, -24, 6	1611	P < 0.05, FWE
	Fusiform Gyrus, Brodmann area 37	38, -48, -20	271	P < 0.05, FWE
	Middle temporal gyrus, Brodmann area 21	-54, 8, -20	446	P < 0.05, FWE



Supplementary Figure 1. Selected clusters of statistically significant voxels overlaid by the average anatomical image of our subject group. Strongest activation for the AB > AX contrast (leftmost panel) in the medial frontal gyrus and Rep > Alt contrast (middle left panel) in the superior temporal gyrus and two clusters for the Alt > Rep contrast (right panels) in the Pulvinar and the fusiform gyrus are displayed.

3.3 Study III. Person identity-specific adaptation effects in the ventral occipito-temporal cortex

This version of the manuscript has been submitted to European Journal of Neuroscience.

Sophie-Marie Rostalski[1], Jonathan Robinson[2,3], Géza GergelyAmbrus[1], Patrick Johnston[3]*, Gyula Kovács[1]*

[1]Department of Biological Psychology and Cognitive Neurosciences, Institute of Psychology, Friedrich Schiller University, Jena, Germany,

[2]Department of Philosophy, Monash University, Melbourne, Australia

[3]School of Psychology & Counselling, Faculty of Health, Queensland University of Technology, Brisbane, Australia

* These two authors contributed to the publication equally as senior authors.

Main research question:

Can image-independent identity changes alter responses in ventral occipito-temporal regions?

Person identity-specific adaptation effects in the ventral occipito-temporal cortex

Sophie-Marie Rostalski[1], Jonathan Robinson[2,3], Géza Gergely Ambrus[1], Patrick Johnston[3*], Gyula Kovács[1*]

[1] Department of Biological Psychology and Cognitive Neurosciences, Institute of Psychology, Friedrich Schiller University, Jena, Germany,

[2] Department of Philosophy, Monash University, Melbourne, Australia

[3] School of Psychology & Counselling, Faculty of Health, Queensland University of Technology, Brisbane, Australia

*** These two authors contributed to the publication equally as senior authors.**

ABSTRACT

Identifying the faces of familiar persons requires the ability to assign several different images of a face to a common identity. Previous research showed that the occipito-temporal cortex, including the fusiform and the occipital face areas, is sensitive to personal identity. Still, the viewpoint, facial expression, and image-independence of this information are currently under heavy debate. Here we adapted a ambient-face rapid serial visual stimulation paradigm (Johnston et al., 2016) and presented highly variable images of famous persons to measure fMRI adaptation. FMRI adaptation is considered as the neuroimaging manifestation of repetition suppression, a neural phenomenon currently explained as a correlate of reduced predictive error responses for expected stimuli. We revisited the question of image-invariant identity-specific encoding mechanisms of the occipito-temporal cortex, using fMRI adaptation with a particular interest in predictive mechanisms. Participants were presented with trials containing eight different images of a famous person, images of eight different famous persons, or seven different images of a particular famous person followed by an identity change to violate potential expectation effects about person identity. We found an image-independent adaptation effect of identity for famous faces in the fusiform face area. However, in contrast to previous electrophysiological studies using similar paradigms, no release of the adaptation effect was observed when identity-specific expectations were violated. Our results support recent multivariate pattern analysis studies, showing image-independent identity encoding in the core face-processing areas of the occipito-temporal cortex. These results are discussed in the frame of recent identity-processing models and predictive mechanisms.

KEY WORDS

Identity recognition, image invariance, fMRI, occipito-temporal cortex

ACKNOWLEDGEMENTS

We would like to thank Philipp Böhm, for his help with recruiting participants and collecting data. This work was partly supported by a grant from the Deutsche Forschungsgemeinschaft (KO3918/5-1) and by the German Academic Exchange Service (DAAD-57446347).

INTRODUCTION

The efficient processing of human faces is an important aspect of social perception. In addition to detecting or recognizing faces in general, the identification of specific familiar faces across a variety of changes in low-level features, such as viewing angle, expression, illumination, or image contrast, is an important ability (Jenkins and Burton, 2011), giving us a great advantage in social situations. Although we are able to perceive a great deal of information even from unfamiliar faces (e.g., sex, age, emotional state, etc.), the identification of a familiar person requires a pre-existing internal representation of that particular person. Tasks involving face recognition or identification (e.g., face matching) that are easy for familiar faces can be very challenging for faces of persons we do not know (Bruce et al., 2001). Thus, it seems more and more likely that familiar and unfamiliar faces are represented differently (Kovács, 2020). Although forming a differential representation of an unfamiliar identity (ID) by performing perceptual tasks, such as sorting faces has been found to be stable enough to elicit a differential electrophysiological response in later event-related potential (ERP) components (Andrews et al., 2017), most studies found that this type of familiarization is not sufficient for creating a stable and robust ID representation in the brain (Dubois et al., 1999; for a review see Natu and O'Toole, 2011). Hence, the exact process of forming a stable identity representation is still unknown (for a recently proposed model, see Kovács, 2020). However, an image-invariant representation is essential for the identification of a person across a wide variety of possible situations and stimulus material.

In an influential model of face processing (Haxby et al., 2000 and Gobbini and Haxby, 2007), areas in the fusiform gyrus represent the ID of a perceived face. Although we now know that a broader network of regions is involved in ID-specific information processing (Duchaine and Yovel, 2015), the fusiform gyrus is still considered to be part of the so-called core face network, supporting higher, more holistic, level of face representation compared to the feature-based representations supported by the inferior occipital gyrus (for a recent multivariate pattern analysis paper, supporting this conclusion see (Tsantani et al., 2021). Therefore, the fusiform gyrus is a strong candidate structure for focusing experiments testing image-invariant face identification processes – that is to say, theoretically, the FFA is proposed to be responsible for identifying facial identity at a conceptual level, rather than simply categorizing familiar images as belonging to different ID categories. A well-established method for testing the extent of stimulus and stimulus-attribute sensitivity of an area is fMRI- adaptation (fMR-a; Grill-Spector & Malach, 2001; Krekelberg et al., 2006). This technique has already been used extensively for twenty years to test for functional characteristics of cortical neurons in several studies (Grill-Spector et al., 1999). For example signal adaptation in the lateral occipital cortex could be observed despite changes in position or size of objects (Malach et al., 1995). If a neuronal population keeps showing reduced responses to a specific object repeatedly presented but in different orientations, these neurons are considered viewpoint-invariant (Andresen et al., 2009). A release from this stimulus-repetition-related adaptation, on the other hand, would be regarded as evidence that this neuronal population encodes the viewpoint. By similar logic, across a set of highly variable ("ambient") face images, if what remains stable with respect to repetition across those set of images is person identity at a conceptual level, such modulation of the neuronal response would be considered as evidence that the neuronal population encodes person identity.

A reduced response to repeated face images in the fusiform face area (FFA; Kanwisher, McDermott, & Chun, 1997) has been shown by many previous studies, using similar techniques (Andrews and Ewbank, 2004; Loffler et al., 2005), improving our understanding of face representations. Still, it is not fully clarified under which circumstances an adaptation effect can be found in ventro-temporal face processing regions. For example, changes in viewpoint were found to result in a release from adaptation in the FFA for unfamiliar, but not for familiar faces (Ewbank and Andrews, 2008), suggesting their differential encoding within the area. This led to the conclusion that the representation of familiar faces is rather viewpoint-independent while that of unfamiliar faces is viewpoint-specific. In other words, different viewpoints of unfamiliar IDs may be perceived as different IDs, whereas familiar ID representations in the FFA are more stable.

The fact that this difference could be found in the FFA provided further evidence for a representation of facial ID in this region. Eger et al. (2005) used familiar and unfamiliar faces differing in rotation angle and expression and found a stronger image-independent reduction in activity due to ID repetition in anterior than in the middle fusiform gyrus, especially for familiar faces. Later, Xu, Yue, Lescroart, Biederman, & Kim (2009) measured the BOLD signal in the FFA elicited by two artificially generated and, therefore, unfamiliar faces from either the same person or different identities and either with the same or with varying viewing angles. They demonstrated that both the change in ID and viewpoint elicited a larger BOLD signal in the FFA compared to presenting identical images. Notably, the ID and viewpoint changing conditions led to similar BOLD signals, indicating that both led to a release of adaptation equally. This finding puts doubt on the role of FFA in the encoding of invariant aspects of face processing as shown in other fMR-a studies (Eger, Schyns, & Kleinschmidt, 2004; Winston, Henson, Fine-Goulden, & Dolan, 2004). For example, Davies-Thompson, Gouws, & Andrews (2009) tested whether occipito-temporal face-selective regions use an image-invariant neural code for familiar face representation in a block design. They contrasted blocks of repetitions of identical images of the same ID with blocks of different images of the same ID and blocks containing different images of different identities – for familiar and unfamiliar identities, separately. A reduced response to the same image - same ID condition, compared to different images of different identities, was found both for familiar and unfamiliar faces in the FFA. Surprisingly, showing different images from the same or different identities resulted in different responses neither for unfamiliar nor familiar faces. In a subsequent study, Davies-Thompson, Newling, & Andrews (2013) replicated these results with a slightly different design: presenting blocks of eight different images of the same ID in contrast to blocks with eight images of different identities did not result in a significantly reduced response in the FFA nor occipital face area (OFA). Altogether, these neuroimaging results suggest the existence of a relatively low-level, image-dependent representation of ID within the core network areas. This conclusion, however, is at odds with more recent studies, which used machine learning techniques to perform multivariate pattern analyses (MVPA) on the FFA. Axelrod and Yovel (2015) were able to discriminate between the response patterns obtained for the different images of two highly familiar identities reliably. Recently, Tsantani et al. (2021) used short video clips and tested the available information in FFA and OFA. They found that both the OFA and the FFA contain ID-specific information and that the FFA reflects higher-level and more image-independent information than the OFA. Thus, so far, no consensus exists in the literature regarding the nature of ID representation of the core face-processing network areas.

Notwithstanding, there is clear evidence for the existence of an early-mid latency ID-specific adaptation effect from recent EEG and MEG studies (e.g., Simpson et al., 2015). Simpson and colleagues (2015) used an MEG adaptation design to show that faces but not objects showed clear adaptation effects localized to the FFA at around 170ms post-stimulus onset. This region also showed a release from adaptation to different identity faces at a latency of around 250-300ms post-stimulus onset. This implies that the FFA is engaged, at different latencies, both in the holistic processing and individuation of face stimuli and, at a slightly later latency, of attaching these holistic face images to particular person identities.

Additional evidence comes from Johnston and colleagues (2016). They found a modulation of the face-sensitive N170 amplitude by introducing expectations about the appearance of a given ID, using several highly variable, ambient (Jenkins and Burton, 2011) face images of the same person in the EEG. A rarely occurring similar image of a different ID in a stream of various face images of the same person was perceived as a deviant. Accordingly, the release from adapting to one ID by encountering a different one manifested in a higher N170 amplitude. The authors related the underlying processes of their observation to predictive coding theories (Friston, 2005; Rao and Ballard, 1999). Those could explain this increase in activity since it is assumed that perception is based on expectations that rely on the prior statistical probability of events. These rely on experiences and are called “prior beliefs,” which are continuously updated in the course of our everyday life. For example, a stream of images showing the same ID would therefore strengthen our expectations about the future appearance of another image of the same person. Thus, the sudden occurrence of a different ID is unexpected and manifests in a greater prediction error, which is measurable as an enhanced neural response in the prediction estimator areas. The key idea here is that when a particular identity is expected to occur, perceptual “templates” corresponding to that identity are preactivated to prioritize the rapid confirmation of that identity (Parr et al., 2020).

More recently, in a paper describing several EEG and one MEG experiment, Johnston and colleagues (Johnston et al., 2017) deployed a “contextual trajectory paradigm,” wherein a series of trials consisting of 5 images depicted a specific contextual trajectory with the final stimulus transition either confirming that trajectory or violating the expectation. They tested trajectories for facial expressions (e.g., from neutral to happy), body rotation (e.g., turning from left to right), and locations of stimuli on the screen (e.g., clockwise or anticlockwise motion). For each experiment, a robust pattern was found, showing a modulatory effect of predictability of the last image in the N170. Irrespective of stimulus type, an enhanced N170 amplitude was found for unpredictable versus predictable stimulus transitions. These results show that a contextual modulation of early ERP components can be found after only four informative images (priors), offering a basis for expectations. In line with this idea, the effects of expectation violations were found to be more pronounced both after five compared to three priors (Robinson et al., 2018) corresponding to the higher precision of the prior belief (Friston and Kiebel, 2009), and where the size of the perceptual distance between the violation event and the expected event was greater (Robinson et al., 2018).

FMR-a is considered as the neuroimaging manifestation of repetition suppression (RS), a phenomenon which is explained currently by many as the correlate of predictive error reduction of repeated or frequent, thereby expected stimuli (for a review, see Kovács and Schweinberger, 2016).

There is also evidence from a recent MEG study showing that prediction error-signals to violations of expected head orientation and facial ID could be spatially dissociated. Whereas prediction-error responses to stimulus orientation were localized to the dorsal visual processing stream, error signals to facial ID were localized to the right fusiform gyrus, among other locations (Robinson et al., 2020). Therefore, adopting the logic of the Johnston et al. (2016) ERP experiment to fMRI, we implemented an fMR-a design to measure the magnitude of release from adaptation in violated ID-specific expectation in key face processing regions. Additionally, this design allows us to revisit the issue of image-invariant ID-specific encoding mechanisms of the FFA by using fMR-a.

Specifically, this work aims at investigating ID-specific processing in the fusiform and occipital face areas, using ambient images of celebrities, very well known to our participants. Furthermore, by generating and violating expectations about person ID, we aimed at testing if prediction error responses are manifest in the neuroimaging signal.

MATERIALS AND METHODS

Participants

Thirty healthy participants took part in this experiment. They gave their informed consent for participation in accordance with the guidelines of the Declaration of Helsinki and with the approval of the ethics committee of the University of Jena. No participant had any history of neurological or psychiatric illness, and all had normal or corrected to normal vision. Three participants had to be excluded from the analysis due to excessive head movements during the scanning session. Altogether, 27 right-handed subjects (16 female/11 male; mean age 27(± 5,7) years) were included in the current analysis. Please note, that some regions of interest (ROI) could not be localized in every participant. Therefore, the number of participants can slightly differ for the different areas.

Stimuli

Colorful images of 16 celebrities (8 males: Chris Hemsworth, Chris Pratt, David Beckham, Ewan McGregor, Gerard Butler, Jude Law, Matt Damon and Tom Hardy, 8 females: Cameron Diaz, Charlize Theron, Gwyneth Paltrow, Jennifer Lawrence, Kate Hudson, Kirsten Dunst, Reese Witherspoon and Scarlett Johansson) that were freely available on the internet were used for this experiment (for examples see Figure 1). The photographs vary in their physical properties (e.g., expression, head-position, eye-gaze, hairstyle, camera-angle, camera exposure, luminance). These types of images are also known from the literature as “ambient images” as they contain natural day-to-day variations under different conditions and can be compared to situations during daily life face recognition (Bortolon et al., 2018; Jenkins et al., 2011). Apart from the fact that we did not use duplicates or flipped image versions, the only other image selection criterion was that the viewing direction of the faces was at least roughly directed towards the camera. Eighteen such ambient images per ID were selected, aligned, and scaled to a resolution of 250*250 pixels (3.3° in radius). Thus, the stimulus set contained 288 different images of 16 different identities. By implementing ambient exemplar images, we ensured that any observed effect reflects higher-level ID processing, independently of the physical features of the images.

In order to functionally localize specific areas of interest (FFA and OFA), a sequence of blocks with images depicting faces, objects, and Fourier noise images were used. Each stimulus category consisted of 40 different stimuli. Faces were randomly occurring colored images of different famous and unknown persons. Identities in the localizer were different from those from the experimental task. Image blocks of objects encompassed a mixture of various items (e.g., food, clothing, etc.). All stimuli were scaled to a resolution of 600x600 pixels (8.5° in radius). The Fourier noise images were created by an algorithm described in Dakin et al. (2002).

Experimental design

The experiment was presented using MATLAB 2013a (The Mathworks) and Psychtoolbox v.3.0.14 (Brainard, 1997). A trial was composed of eight subsequently presented face images of same-sex identities (female and male face trials were presented 50%). Each image was presented for 500ms without any inter-stimulus-interval (ISI) and was slightly jittered spatially around the screen center to avoid low-level adaptation processes. Thus, a trial lasted 4 seconds, and it ended with a fixation cross. The intertrial interval (ITI) was randomized to 2,4 or 6 seconds. Four different conditions of such trials were created:

First, *alternation* trials (ALT) consisted of the ambient images of eight different identities.

Second, in *adaptation* trials (ADA), eight different images of the same ID were presented to test image-independent ID-specific adaptation effects.

Third, in the e*xpectation violation* condition (EV), a series of seven different images of the same ID was followed by the face of another ID. We reasoned that if ID-specific expectation modulates the observed adaptation effects, then the unexpected change of facial ID should lead to a release of adaptation.

Theoretically, after seeing the second image in the row, participants could expect the appearance of images of the same ID both in ADA and EV and the appearance of different identities in ALT. Therefore, we created a fourth condition which was similar to alternation trials but ended with the repetition of one ID. In this condition, the first six images depicted six different identities, followed by two images of the same ID. These *Alternation with final repetition* (ALT-II) trials were not subject to any specific hypotheses but ensured us that the participants had equal number of trials where the first seven images were depicting the same ID (ADA, EV) or different identities (ALT, ALT-II). Examples for all four trial types are shown in Figure 1.

The trials of these four conditions appeared with the same probability randomly, with the only constraint being that a maximum of three subsequent trials could depict same sex identities and a maximum of two trials could come from the same condition.

Since attention is known to modulate response suppression and expectation violation (Larsson and Smith, 2012), participants had to perform a task, unrelated to the above-described manipulations. They had to respond to images with reduced size (1.98°), which could occur at any position within a trial sequence (detection rate of these target trials: 68,2% plus minus 21,2%). This task was set to ensure that participants focused their attention on the stimuli without diverting their attention to the different conditions. To avoid potential effects of attention these target trials were removed from all further analyses. The main experimental procedure comprised 4 runs with one run including 80 trials (20 trials per condition and 10% target-detection-trials).

We hypothesized that in areas encoding facial ID in an image-independent manner, ADA should lead to a reduced average BOLD signal, when compared to ALT trials. Specifically, if an area is involved in image-independent facial ID processing, a lower BOLD response should be observable when different images of the same famous ID are repeated compared to when images of different famous identities are presented. In addition, previous studies suggested that the electrophysiological measures of face processing reflect the expectation of the occurrence of the same or different identities and the violation of these expectations (Johnston et al., 2016). We reasoned that if predictive mechanisms explain ID-specific signal reductions, then the violation of such expectations in the EV condition should manifest in a release of adaptation as well.

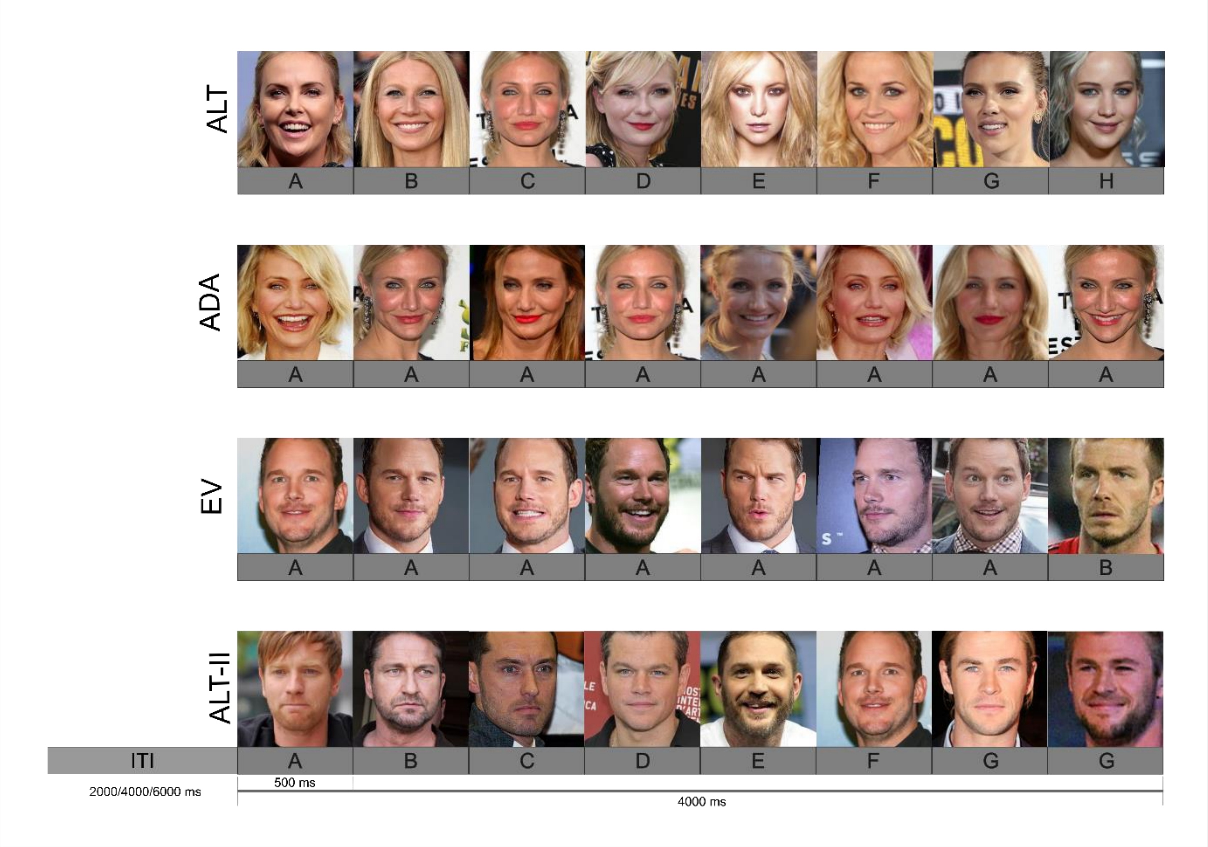


Figure 1. Conditions and trial structure in the four applied conditions. Each condition was composed of 8 faces. ALT: different same-sex identities. ADA: 8 different images of the same ID. EV: 7 different images of the same

ID, followed by another same-sex ID. ALT-II: 6 images of different same-sex identities, followed by two different images of another, same-sex ID. Please note that these images are just examples and might not have been part of the actual stimulus set, as we do not have permission for publishing all exemplars we used. All images shown here are subject to either Creative Commons Attribution-Share Alike license or have no copyrights (public domain).

Procedure and imaging parameters

Participants were introduced to the MRI center, and a medical briefing was conducted. Next, they were asked to make familiarity judgments about the female and male identities used in the main experiment. For this, sample images of each ID (which were not used in the main experiment) were presented first alone and then together with their names and professions. Participants had to indicate whether they are familiar with them. Only if they reported to know the face and the name, the ID was evaluated as being familiar to the participant.

The scanning was conducted with a 3 Tesla MR Scanner (Siemens Prisma fit). All functional data was obtained using an Echo Planar Imaging (EPI) Sequence (35 slices; TR = 2000ms; TE = 30ms; flip angle = 90°; 64*64 matrices; in-plane resolution: 3x3 mm^2; slice thickness: 3mm). A magnetization-prepared rapid gradient-echo sequence (MP-RAGE; TR=2300ms; TE=3,03, 1mm isotropic voxel size) was used to acquire high-resolution T1-weighted sagittal images to generate 3D structural scans. All images were acquired using a 20-channel head coil.

Behavioral data were recorded by a button box. There was only one button to signal the detection of target stimuli. First, within each scanning session, two experimental runs were administered, followed by the anatomical scan and another two experimental runs. Finally, the localizer scan completed the session of approximately one hour.

We implemented a localizer sequence to determine the relevant regions of interest. Here, blocks of images (presented for 250 ms) showing faces, objects, and Fourier noise were used. Each block was repeated five times, interleaved with blank periods of 12 seconds. Stimuli occurred randomly within one block.

Data processing and statistical analyses

Data and code required to reproduce all analyses will be available at OSF (https://osf.io/m3pwt/) at the time of publication. The conditions of our ethics approval do not permit the publishing of the raw MRI data. We will therefore provide extracted fMRI data from individual coordinates, which will be made available as well.

Neuroimaging data were preprocessed using SPM12 (Wellcome Department of Imaging Neuroscience, London, UK). In brief, the functional data were corrected for shifts in acquisition time of slices, realigned to correct for movement, co-registered to the anatomical images, normalized to the MNI-152 space, resampled to 2 mm isotropic voxel size, and finally spatially smoothed with an 8mm FWHM Gaussian kernel. A general linear model was specified, using the onsets of the trials of the four different conditions and the six movement parameters as regressors. For the experimental functional data, hemodynamic derivatives were added to the model.

To identify the location of FFA and OFA, we contrasted face blocks with blocks of objects and Fourier-randomized noise from the localizer sequence with a threshold of either $p < .05$ family-wise error (FWE) corrected ($n = 19$) or $p < .0001$ uncorrected ($n = 8$). The right FFA could be localized in all 27 participants (average MNI coordinates (±SE): 41 (0), -53 (2), -19 (1)) and in 26 participants in the left hemisphere (average MNI coordinates (±SE): -40 (1), -52 (2), -20 (1)). For every subject, we used the same contrast and threshold to identify OFA. The Right OFA could be localized in 25 participants (average MNI coordinates (±SE): 42 (1), -75 (2), -12 (1)). In the left hemisphere, OFA was also localized in 25 subjects (average MNI coordinates (±SE): -40 (1), -77 (2), -12 (1)). Individual coordinates can be found in the supplementary material Table 1.

The BOLD signals evoked during experimental conditions were extracted from each individual ROI, defined as the peak face responding voxel, using a 2 mm radius sphere (Brett, 2011). Hemodynamic response functions (HRF) were inspected to assure that the ROIs were identified reliably, and the extracted signal was evaluated. The peak HRF values were entered into the statistical models. We performed a two-way repeated-measures ANOVA with the within-subject factors of hemisphere (*right*, *left*) and condition (*adaptation*, *expectation violation*. *alternation*. *alternation-II*). Finally, we conducted an exploratory, second-level whole-brain analysis.

RESULTS

To assess whether participants were familiar with the presented identities, they filled out a questionnaire prior to the experiment. Mean familiarity ratings are 73,2% for male and 80,6% for female IDs and show that our participants were familiar with the stimuli.

FMRI Results

The mean BOLD signal in the FFA for all conditions is presented in Figure 2A. The repeated measures ANOVA revealed a main effect of condition ($F_{(3,60)} = 2.962$, $p = 0.039$, $\eta_p^2 = 0.129$). Post-hoc tests showed a significant difference between *ALT* and *ADA* trials, $t(20) = -2.971$, $p_{\text{holm}} = 0.045$. This shows that the presentation of different, highly variable natural images of the same ID leads to response reduction in the FFA, suggesting that it plays a role in the encoding of ID in an image-independent manner. All other comparisons remained nonsignificant. No interaction of hemisphere and condition was found ($F_{(3,60)}$ = 1.324, p = 0.275, η_p^2 = 0.062), suggesting similar effects over the left and right FFA.

The repeated measures ANOVA showed a significant main effect of condition in the bilateral OFA as well ($F_{(1.978,37.581)}$ = 4.693, p = 0.015, η_p^2 = 0.198; Greenhouse-Geisser corrected). Post-hoc tests revealed that the only significant difference is between ALT and EV t(19) = 3.034, p_{holm} = 0.022. Unlike in the FFA, the *ALT-ADA* comparison remained nonsignificant for the OFA t(19) = -2.610, p_{holm} = 0.055. We did not find an interaction of hemisphere and condition for the OFA either ($F_{(3,57)}$ = 1.289, p = 0.287, η_p^2 = 0.064).

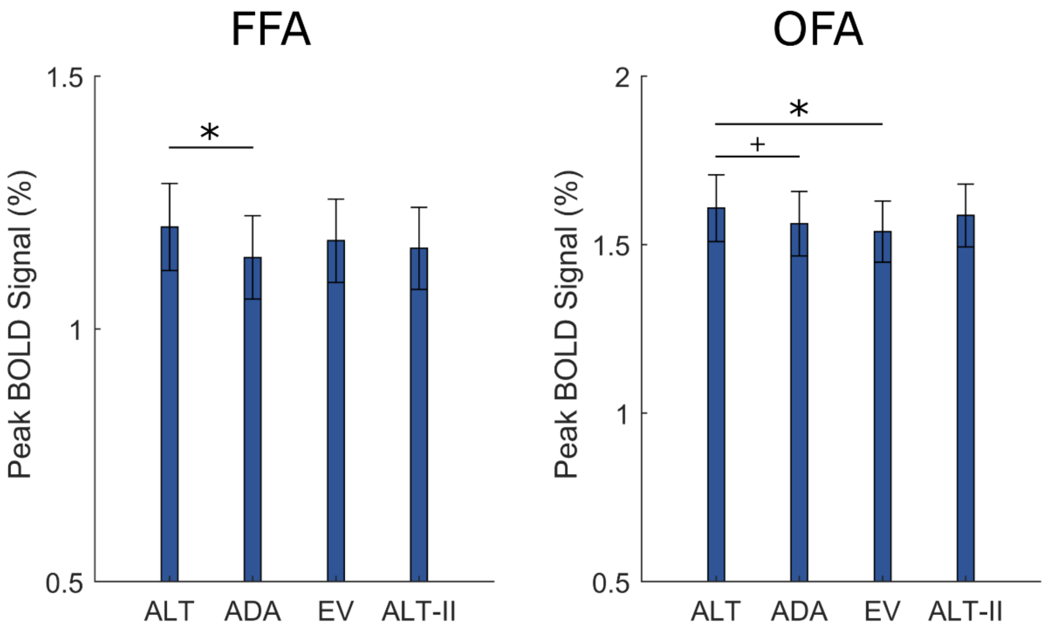


*Figure 2. Peak BOLD signal to the different conditions for the bilateral FFA (left panel) and OFA (right panel). * represents p<0.05; + represents p=0.055.*

To better evaluate the evidence for differences of BOLD responses to our conditions, we conducted a Bayesian repeated-measures ANOVA, including post hoc tests. We report Bayes factor BF_{10} for both analyses, reflecting how much more likely our data occur under alternative hypotheses than the null hypothesis.

In the FFA, the Bayes factor for a main effect of condition was 1.353, signaling that the effect of condition is more likely than the null hypothesis. More interestingly, post hoc tests revealed a Bayes factor for the *ADA* versus *ALT* comparison of 6.391, confirming our previous analysis and a Bayes factor of 1.463 for the difference between *ADA* and *EV* trials. All other comparisons revealed Bayes factors < 1, suggesting that the evidence favors the null hypothesis.

The Bayesian repeated-measures ANOVA of OFA data revealed a Bayes factor of 7.390, favoring the alternative hypothesis over the null hypothesis by a factor of 7. Post hoc tests showed Bayes factors for the *EV* trial comparison with *ALT-II* of 25.563 and with *ALT* trials of 1.930. The Bayes factor for the comparison of *ADA* and *ALT* trials was 1.904, which is similar to the EV and ALT comparison. Additionally, the comparison of ADA and ALT-II revealed a Bayes factor of 2.319. All other Bayes factors remained < 1.

Whole-brain Analyses

In order not to overlook any area that might show activation differences to the different conditions outside the pre-defined ROIs, we computed a whole-brain random design analysis contrasting adaptation > alternation on the group level. Applying a threshold of $p < 0.0001_{\text{uncorrected}}$ revealed only one cluster of activation (k = 5) in the right inferior frontal gyrus (MNI[x,y,z]: 62, 8, 24) (Figure 3), an area close to the inferior frontal face area (iFFA), which is part of the extended face-processing network and is known to play a role in the processing of eye-gaze and the semantic aspects of faces (Chan and Downing, 2011; Duchaine and Yovel, 2015; Ishai, 2008). No other contrast revealed significant clusters.

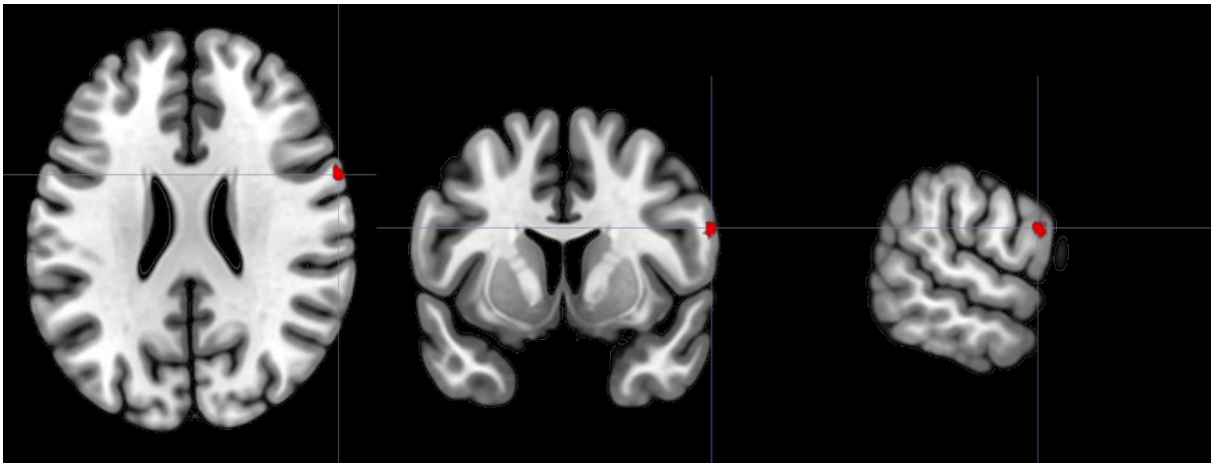


Figure 3. Results of the whole-brain analyses. Significant cluster for the contrast ALT > ADA.

DISCUSSION

In the present study, we investigated ID-specific adaptation effects within the occipito-temporal face processing areas. We found an image-independent adaptation effect of identity for famous faces in the fusiform face area. This difference of presenting highly variable, ambient images of the same versus different identities was only significant for the FFA. In contrast, the OFA showed significantly lower activation for a condition where expectations are violated compared to alternating identities and a strong tendency for ID specific adaptation.

Previously, Ewbank and Andrews (2008) found fMR-a across different viewpoints in the FFA to familiar, but not to unfamiliar faces. However, their adaptation condition contained images, although varying in viewpoint, that still came from the same original images. Also, the implemented range of viewpoint change of this study was relatively small (12°). Therefore, these images were very similar in low-level features. Their interpretation of an ID-specific adaptation effect for familiar faces in the FFA is consequently only partly justified. Still, this conclusion is confirmed by our current study by the application of highly variable, ambient images (Jenkins and Burton, 2011).

Other studies in which blocks with different images from the same ID and blocks with different images of different IDs were contrasted failed to find a difference in FFA responses to those conditions (Davies-Thompson et al., 2009). Thus, the current study shows that it is possible to discriminate familiar identities in FFA activity, providing evidence for a stable and image-independent ID representation in the area. One explanation for the discrepant results of the current and previous studies could come from small but significant differences in the applied designs. While we used colorful ambient images, Davies-Thompson et al.(2009) used gray-scale frontal faces. Also, they presented ten images per block and chose stimuli from a larger stimulus set. However, the latter differences should have made it easier to find an ID-specific adaptation effect in their study. In a follow-up experiment, they specifically investigated the responses to blocks of 8 different images showing either the same or different IDs (Davies-Thompson et al., 2013). Again, there was no significant fMR-a effect for ID present. In this study, responses to familiar and unfamiliar faces were compared and the same results were found for both stimulus types.

More recent studies, which applied multivariate pattern analyses (MVPA) to fMRI data have been able to discriminate between identities even when they were unknown (Anzellotti et al., 2014; Nestor et al., 2011). Our current results confirm these findings and support the idea that the FFA discriminates between identities independently of images and is therefore subject to ID-specific adaptation effects. Our results also fit those of recent MVPA studies of highly familiar faces which are more comparable to the stimulus material of the current study (Axelrod and Yovel, 2015; Tsantani et al, 2020). Both studies found identity-specific information in the bilateral FFA.

Although the adaptation versus alternation comparison did not reach significance in the OFA, we could show that the expectation violation was significantly different to the alternation condition. More interestingly the neural response to trials in which expectations were violated was almost the same as for the adaptation condition. This marginally significant effect of ID-specific adaptation hints towards an effect similar to that of the FFA but requires some more detailed analyses, such as previously mentioned multivariate ones.

Unlike the previous ERP studies which used similar paradigms (Johnston et al., 2016) we were unable to find effects of ID-specific expectation-suppression and expectation-violation related response enhancements in the occipito-temporal cortex. This may be due to the low temporal resolution of fMRI. Because of the limitation of the duration of the entire experiment, we opted for a paradigm in which the BOLD response to the entire trial is modeled. Since the images within a trial are not separated by sufficiently long ISIs, it was not possible to separate the response to the individual images (specially to the last one or two images) from the rest.

The whole-brain analysis revealed a single cluster, being more active for ALT when compared to ADA in the inferior frontal gyrus, corresponding closely to the recently described area of iFFA. This area is supposed to be part of the face-processing network and is known to play a role in the processing of dynamic face properties as well as eye-gaze information. (Chan and Downing, 2011; Duchaine and Yovel, 2015; Ishai, 2008). As identity was kept constant in ADA, but changed continuously in ALT, our results raise the possibility that this area is also specifically involved in high-level predictions about identity continuity within an image sequence. The confirmation of this hypothesis, however, will require specific future studies.

CONCLUSION

Confirming results from multivariate pattern analyses, the present study shows image-independent ID-specific adaptation effects in the fusiform face area for famous familiar faces. Especially in combination with the results of the occipital face area, our results suggest that the ID representations in occipito-temporal regions are not yet sufficiently clarified and that further research is needed. We could not replicate results from MEG studies, showing expectation violation effects related to facial identity in the fusiform gyrus.

AUTHOR CONTRIBUTIONS

PJ and GK devised and supervised the project, SMR, JR, GA, GK, and PJ designed the experiment, GA programmed the experiment. SMR carried out the experiment and performed all analyses. SMR and GK wrote the manuscript with support from JR, GA, and PJ.

FUNDING AND CONFLICTS OF INTEREST

The work was partly funded by the German academic exchange Service DAAD (57446347) and by the Deutsche Forschungsgemeinschaft (KO3918/5-1). The authors declare no conflict of interests.

REFERENCES

Andresen, D.R., Vinberg, J., Grill-Spector, K., 2009. The representation of object viewpoint in human visual cortex. Neuroimage 45, 522–536. https://doi.org/10.1016/j.neuroimage.2008.11.009

Andrews, S., Burton, A.M., Schweinberger, S.R., Wiese, H., 2017. Event-related potentials reveal the development of stable face representations from natural variability. Q. J. Exp. Psychol. 70, 1620–1632. https://doi.org/10.1080/17470218.2016.1195851

Andrews, T.J., Ewbank, M.P., 2004. Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. Neuroimage 23, 905–913. https://doi.org/10.1016/j.neuroimage.2004.07.060

Anzellotti, S., Fairhall, S.L., Caramazza, A., 2014. Decoding representations of face identity that are tolerant to rotation. Cereb. Cortex 24, 1988–1995. https://doi.org/10.1093/cercor/bht046

Axelrod, V., Yovel, G., 2015. Successful decoding of famous faces in the fusiform face area. PLoS One 10, 19–25. https://doi.org/10.1371/journal.pone.0117126

Bortolon, C., Lorieux, S., Raffard, S., 2018. Self- or familiar-face recognition advantage? New insight using ambient images. Q. J. Exp. Psychol. (Hove). 71, 1396–1404. https://doi.org/10.1080/17470218.2017.1327982

Brainard, D.H., 1997. The Psychophysics Toolbox. Spat. Vis. 10, 433–436. https://doi.org/10.1163/156856897X00357

Brett, M., 2011. MarsBaR Documentation. Funct. Imaging 1–45.

Bruce, V., Henderson, Z., Newman, C., Burton, A.M., 2001. Matching identities of familiar and unfamiliar faces caught on CCTV images. J. Exp. Psychol. Appl. 7, 207–218. https://doi.org/10.1037/1076-898X.7.3.207

Chan, A.W.-Y., Downing, P.E., 2011. Faces and Eyes in Human Lateral Prefrontal Cortex. Front. Hum. Neurosci. 5, 51. https://doi.org/10.3389/fnhum.2011.00051

Dakin, S.C., Hess, R.F., Ledgeway, T., Achtman, R.L., 2002. What causes non-monotonic tuning of fMRI response to noisy images? Curr. Biol. 12, R476–R477. https://doi.org/10.1016/S0960-9822(02)00960-0

Davies-Thompson, J., Gouws, A., Andrews, T.J., 2009. An image-dependent representation of familiar and unfamiliar faces in the human ventral stream. Neuropsychologia 47, 1627–1635. https://doi.org/10.1016/j.neuropsychologia.2009.01.017

Davies-Thompson, J., Newling, K., Andrews, T.J., 2013. Image-invariant responses in face-selective regions do not explain the perceptual advantage for familiar face recognition. Cereb. Cortex 23, 370–377. https://doi.org/10.1093/cercor/bhs024

Dubois, S., Rossion, B., Schlitz, C., Bodart, J.M., Michel, C., Bruyer, R., Crommelinck, M., 1999. Effect of familiarity on the processing of human faces. Neuroimage 9, 278–289. https://doi.org/10.1006/nimg.1998.0409

Duchaine, B., Yovel, G., 2015. A Revised Neural Framework for Face Processing. Annu. Rev. Vis. Sci. 1, 393–416. https://doi.org/10.1146/annurev-vision-082114-035518

Eger, E., Schweinberger, S.R., Dolan, R.J., Henson, R.N., 2005. Familiarity enhances invariance of face representations in human ventral visual cortex: fMRI evidence. Neuroimage 26, 1128–1139. https://doi.org/10.1016/j.neuroimage.2005.03.010

Eger, E., Schyns, P.G., Kleinschmidt, A., 2004. Scale invariant adaptation in fusiform face-responsive regions. Neuroimage 22, 232–242. https://doi.org/10.1016/j.neuroimage.2003.12.028

Ewbank, M.P., Andrews, T.J., 2008. Differential sensitivity for viewpoint between familiar and unfamiliar faces in human visual cortex. Neuroimage 40, 1857–1870. https://doi.org/10.1016/j.neuroimage.2008.01.049

Friston, K., 2005. A theory of cortical responses. Philos. Trans. R. Soc. B Biol. Sci. 360, 815–836. https://doi.org/10.1098/rstb.2005.1622

Friston, K., Kiebel, S., 2009. Predictive coding under the free-energy principle. Philos. Trans. R. Soc. B Biol. Sci. 364, 1211–1221. https://doi.org/10.1098/rstb.2008.0300

Gobbini, M.I., Haxby, J. V., 2007. Neural systems for recognition of familiar faces. Neuropsychologia 45, 32–41. https://doi.org/10.1016/j.neuropsychologia.2006.04.015

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., Malach, R., 1999. Differential processing of objects under various viewing conditions in the human lateral occipital complex. Neuron. https://doi.org/10.1016/S0896-6273(00)80832-6

Grill-Spector, K., Malach, R., 2001. fMR-adaptation: A tool for studying the functional properties of human cortical neurons. Acta Psychol. (Amst). 107, 293–321. https://doi.org/10.1016/S0001-6918(01)00019-1

Haxby, J. V., Hoffman, E.A., Gobbini, M.I., 2000. The distributed human neural system for face perception. Trends Cogn. Sci. 4, 223–233. https://doi.org/10.1016/S1364-6613(00)01482-0

Ishai, A., 2008. Let's face it: It's a cortical network. Neuroimage.

https://doi.org/10.1016/j.neuroimage.2007.10.040

Jenkins, R., Burton, M., 2011. Stable face representations. Philos. Trans. R. Soc. B Biol. Sci. 366, 1671–1683. https://doi.org/10.1098/rstb.2010.0379

Jenkins, R., White, D., Van Montfort, X., Mike Burton, A., 2011. Variability in photos of the same face. Cognition 121, 313–323. https://doi.org/10.1016/j.cognition.2011.08.001

Johnston, P., Overell, A., Kaufman, J., Robinson, J., Young, A.W., 2016. Expectations about person identity modulate the face-sensitive N170. Cortex 85, 54–64. https://doi.org/10.1016/j.cortex.2016.10.002

Johnston, P., Robinson, J., Kokkinakis, A., Ridgeway, S., Simpson, M., Johnson, S., Kaufman, J., Young, A.W., 2017. Temporal and spatial localization of prediction-error signals in the visual brain. Biol. Psychol. 125, 45–57. https://doi.org/10.1016/j.biopsycho.2017.02.004

Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. 17, 4302–11. https://doi.org/10.1098/Rstb.2006.1934

Kovács, G., 2020. Getting to know someone: Familiarity, person recognition, and identification in the human brain. J. Cogn. Neurosci. 32, 2205–2225. https://doi.org/10.1162/jocn_a_01627

Kovács, G., Schweinberger, S.R., 2016. Repetition suppression – An integrative view. Cortex 80, 1–4. https://doi.org/10.1016/j.cortex.2016.04.022

Krekelberg, B., Boynton, G.M., van Wezel, R.J.A.A., 2006. Adaptation: from single cells to BOLD signals. Trends Neurosci. 29, 250–256. https://doi.org/10.1016/j.tins.2006.02.008

Larsson, J., Smith, A.T., 2012. fMRI Repetition Suppression : Neuronal Adaptation or Stimulus Expectation ? Cereb. Cortex 22, 567–576. https://doi.org/10.1093/cercor/bhr119

Loffler, G., Yourganov, G., Wilkinson, F., Wilson, H.R., 2005. fMRI evidence for the neural representation of faces. Nat. Neurosci. 8, 1386–1390. https://doi.org/10.1038/nn1538

Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R., Tootell, R.B., 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc. Natl. Acad. Sci. 92, 8135–8139. https://doi.org/10.1073/pnas.92.18.8135

Natu, V., O'Toole, A.J., 2011. The neural processing of familiar and unfamiliar faces: A review and synopsis. Br. J. Psychol. 102, 726–747. https://doi.org/10.1111/j.2044-8295.2011.02053.x

Nestor, A., Plaut, D.C., Behrmann, M., 2011. Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. Proc. Natl. Acad. Sci. 108, 9998–10003. https://doi.org/10.1073/pnas.1102433108

Parr, T., Rikhye, R.V., Halassa, M.M., Friston, K.J., 2020. Prefrontal Computation as Active Inference. Cereb. Cortex 30, 682–695. https://doi.org/10.1093/cercor/bhz118

Rao, R.P.N., Ballard, D.H., 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. 2, 79–87. https://doi.org/10.1038/4580

Robinson, J.E., Breakspear, M., Young, A.W., Johnston, P.J., 2018. Dose-dependent modulation of the visually evoked N1/N170 by perceptual surprise: A clear demonstration of prediction-error signalling. Eur. J. Neurosci. 1–11. https://doi.org/10.1111/ejn.13920

Robinson, J.E., Woods, W., Leung, S., Kaufman, J., Breakspear, M., Young, A.W., Johnston, P.J., 2020. Prediction-error signals to violated expectations about person identity and head orientation are doubly-dissociated across dorsal and ventral visual stream regions. Neuroimage 206, 116325. https://doi.org/10.1016/j.neuroimage.2019.116325

Simpson, M.I.G., Johnson, S.R., Prendergast, G., Kokkinakis, A. V., Johnson, E., Green, G.G.R., Johnston, P.J., 2015. MEG adaptation resolves the spatiotemporal characteristics of face-sensitive brain responses. J. Neurosci. 35, 15088–15096. https://doi.org/10.1523/JNEUROSCI.2090-15.2015

Tsantani, M., Kriegeskorte, N., Storrs, K., Williams, A.L., McGettigan, C., Garrido, L., 2021. FFA and OFA encode distinct types of face identity information. J. Neurosci. 41, JN-RM-1449-20. https://doi.org/10.1523/jneurosci.1449-20.2020

Winston, J.S., Henson, R.N.A., Fine-Goulden, M.R., Dolan, R.J., 2004. fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. J. Neurophysiol. 92, 1830–1839. https://doi.org/10.1152/jn.00155.2004

Xu, X., Yue, X., Lescroart, M.D., Biederman, I., Kim, J.G., 2009. Adaptation in the fusiform face area (FFA): Image or person? Vision Res. 49, 2800–2807. https://doi.org/10.1016/j.visres.2009.08.021

Person identity-specific adaptation effects in the ventral occipito-temporal cortex: supplementary material

Sophie-Marie Rostalski[1], Jonathan Robinson[2,3], Géza G. Ambrus[1], Patrick Johnston[3*], Gyula Kovács[1*]

[1] Department of Biological Psychology and Cognitive Neurosciences, Institute of Psychology, Friedrich Schiller University, Jena, Germany,

[2] Department of Philosophy, Monash University, Melbourne, Australia

[3] School of Psychology & Counselling, Faculty of Health, Queensland University of Technology, Brisbane, Australia

*** These two authors contributed to the publication equally as senior authors.**

Table 1

		FFA right			FFA left			OFA right			OFA left		
	contrast	**x**	**y**	**z**	**x**	**y**	**z**	**x**	**y**	**z**	**x**	**y**	**Z**
S01	Faces vs Objects and Noise; FWE	42	-50	-22	-40	-50	-18	40	-86	-20	-40	-86	-12
S02	Faces vs Noise; FWE	38	-58	-12	-36	-54	-20	46	-72	-6	-46	-72	-12
S03	Faces vs Objects and Noise; FWE	42	-66	-16	-36	-62	-14	40	-70	-18	-42	-80	-2
S04	Faces vs Objects+Noise; p<0.0001	42	-64	-20				36	-88	-20	-42	-84	-22
S05	Faces vs Objects+Noise; FWE	42	-58	-18	-38	-42	-24	44	-84	-4			
S06	Faces vs Objects+Noise; FWE	44	-54	-22	-42	-52	-24	42	-82	-8	-34	-80	-8
S07	Faces vs Noise; p<0.001;	38	-54	-20	-40	-38	-14				-48	-82	6
S08	Faces vs Objects+Noise; FWE	40	-50	-22	-36	-48	-22	46	-78	-16	-34	-82	-12
S09	excluded for movement												
S10	Faces vs Objects+ Noise; FWE	40	-50	-20	-40	-64	-14	50	-78	-4	-40	-64	-14
S11	Faces vs Objects+Noise; FWE	44	-70	-14	-40	-74	-16	44	-70	-14	-40	-72	-16
S12	Faces vs Objects+Noise; FWE	40	-56	-18	-36	-44	-18	42	-72	-14	-36	-76	-18
S13	Faces vs Objects+Noise; FWE	42	-48	-16	-38	-46	-22	40	-84	-12	-42	-72	-20
S14	Faces vs Objekt+Noise; FWE	46	-46	-20	-40	-46	-26	46	-82	-8	-42	-78	-6
S15	Faces vs Objects+Noise; FWE	40	-54	-16	-44	-52	-24	40	-78	-8	-36	-70	-12
S16	Faces vs Objects+Noise; FWE	46	-50	-24	-42	-52	-24	46	-82	-6	-38	-86	-16
S17	excluded for movement												
S18	Faces vs Objects+Noise; FWE	40	-62	-14	-42	-64	-18	38	-76	-8	-40	-74	-16
S19	Faces vs Noise p<0.0001	36	-38	-24	-38	-56	-20	40	-60	-16			
S20	Faces vs Objects+Noise; p<0.0001	40	-50	-18	-38	-54	-22	38	-60	-10			
S21	excluded for movement												
S22	Faces vs Objects+Noise; p<0.0001	42	-52	-22	-36	-50	-18	44	-72	-6	-40	-78	-16
S23	Faces vs Objects and Nosie	40	-38	-16	-42	-50	-22	40	-64	-16	-36	-62	-12
S24	Faces vs Objects+Noise; p<0.0001	40	-42	-22	-42	-54	-18	42	-66	-14	-36	-78	-14
S25	Faces vs Objects+Noise; FWE	40	-52	-18	-42	-48	-14	42	-62	-16	-40	-76	-12
S26	Faces vs Noise; p<0.0001	40	-54	-20	-40	-54	-16	38	-78	-10	-44	-82	-4
S27	Faces vs Objects+Noise; FWE	44	-50	-26	-38	-46	-20	42	-78	-16	-46	-78	-10
S28	Faces vs Objects+Noise; FWE	40	-68	-16	-44	-46	-24	40	-68	-16	-36	-66	-14
S29	Faces vs Objects+Noise; p<0.001	44	-46	-20	-40	-54	-22	40	-72	-18	-48	-78	-12
S30	Faces vs Objects+Noise; FWE	40	-50	-18	-36	-44	-22	42	-82	-8	-38	-82	-10

4 DISCUSSION

4.1 Summary and discussion of the specific studies

In this chapter, I will summarize the results of the current studies and show their contribution to our understanding of the neuronal mechanisms underlying the effects of stimulus repetitions and predictive processes. As all studies in themselves have been discussed in the related manuscripts separately, here we will focus on the overarching results and their relations to each other.

4.1.1 Sensory noise does not affect RS magnitude in Study I

The first empirical study (chapter 3.1) aimed to test the influence of sensory noise on RS. No such effect was evident in the examined brain regions. Importantly, we did not apply either probability-based or cue-based manipulations to RS (see, Figure 12). Repetition and alternation trials appeared equally throughout the experiment (Figure 28A), which refers to constant priors. Further, the first stimulus in each trial was only informative about the noise level to be expected in the upcoming stimulus. This design allowed us to test hypotheses of PC models without manipulating higher-order expectations. The precision of sensory input was manipulated by adding noise to unfamiliar face stimuli. Given that the repetition of a stimulus is the default prior in PC models, RS should be evident for clear and noisy stimuli – at least in a design where no higher-order expectations need to be considered for efficiently processing the incoming sensory data. We found a main effect signaling RS, as well as reduced responses to noisy trials in general. However, no modulation of RS due to sensory precision was observed.

With regard to unequal predictability of an upcoming stimulus in classical RS experiments as addressed in Study II (chapter 3.2), one could argue that although the first stimulus in repetition trials serves as a strong predictor for the upcoming stimulus, this prediction is more demanding and thereby imprecise for noisy trials. We could not provide evidence for this assumption in our design. The absence of RS modulation does not necessarily mean the absence of inferential processes. Still, our results can also be explained by feedforward models, such as the sharpening or facilitation models (Grill-Spector et al., 2006), or local neuronal computations that fit within PC models but can occur without dop-down influence

(Bastos et al., 2012), which all offer explanations for RS (see chapter1.2.2). However, with the present design and method, the underlying mechanisms of RS are difficult to ascertain.

4.1.2 Predictability of stimuli can mimic RS effects in Study II.

In Study II (chapter 3.2) of this thesis, participants were visually familiarized with a set of twelve stimuli depicting female faces in four behavioral training sessions. The initially novel faces underlay a specific pattern of statistical associations (see Figure 28B for a simple illustration). Although all conditions were presented in an interleaved manner, participants incidentally encoded transitional rules while performing an unrelated task. After training, participants could predict features of upcoming stimuli regardless of it being a repetition of the recently encountered one or an alternation in the AB-context as revealed by fMRI data. Stimulus-specific expectations in alternation trials were found to elicit responses similar to those measured for a stimulus repetition of a face in the fusiform gyrus. This finding is in line with previous neuroimaging results for face responses in the FFA (Pajani et al., 2017) and results from electrophysiological studies (Feuerriegel et al., 2018a). It clearly demonstrates the importance of stimulus predictability in determining the magnitude of RS. Depending on implemented designs and stimulus material, it can affect related phenomena as well (Grotheer & Kovács, 2015; Summerfield et al., 2008). In addition, in the occipital face area, RS was evident in the right hemisphere only, and no modulatory effect of stimulus predictability was observed. This finding suggests that the FFA but not OFA is sensitive to statistically learned transition rules. Our findings do not provide direct evidence for feedforward or PC models of RS. However, we show how stimulus-specific expectations can alter specific measures of RS, which is in line with PC models for perceptual expectations and relevant for all investigations that implement repetition and alternation trials to test for prediction effects.

4.1.3 Identity-specific adaptation effects but no effect of violated expectations are present in occipito-temporal face regions in Study III

In Study III, we investigated adaptation effects for higher-order information in face stimuli. Identity-related and image-independent adaptation effects were tested in a design with short blocks of eight ambient images that could show the same or different famous identities (Figure 28C). The FFA revealed an identity-specific adaptation effect, and a marginal effect was observed in the OFA. Evidence for identity processing in the OFA comes from several

TMS studies (e.g., Ambrus et al., 2017b; Eick et al., 2020). Furthermore, the OFA is connected to higher-level face processing areas (Pyles et al., 2013). Together with the assumption that famous identities come with a stable neural representation that can be activated, top-down predictions can influence the OFA's neural activity. Hence, the observed tendency for identity-related adaptation effects could result from feedback projections of higher-level areas. Identity-specific adaptation effects in the FFA support its crucial role in face processing as proposed by several models (e.g., Duchaine & Yovel, 2015; Haxby et al., 2000). However, its capability of encoding image-independent identity information has only been shown by MVPA studies so far (Axelrod & Yovel, 2015; Tsantani et al., 2021). The proposed prediction-related response enhancement due to expectation violation suggested by electrophysiological studies (Johnston et al., 2016) was not observed in our data. The usage of famous familiar faces led us to believe that a stable identity representation already exists and is activated rapidly. Testing identity-specific expectations refers to a solid prior belief in terms of predictions. That no such effect was evident, although the FFA encodes the critical information suggests that we might have to implement a paradigm that allows for higher spatial and temporal resolution analyses.

4.1.4 General discussion

In brief, in the first two studies, we examined the robustness of RS by attempting to modulate its magnitude through manipulations of precision and predictability as suggested by PC models. No effects of our precision manipulation were present in face-sensitive regions FFA and OFA for unfamiliar faces in Study I. In Study II, the diminishing influence of stimulus predictability on RS was evident in responses to visually familiar faces but only in the FFA. In Study III, we used an adaptation design to investigate identity representations of famous persons in the FFA and OFA. An image-independent adaptation effect to well-known identities was evident in the FFA, and a tendency for the same effect was observed in the OFA. We could not prove the release from adaptation to unexpected changes of identity.

I will now relate the above-mentioned main findings of our studies to previous work focusing on implemented design and stimulus material used. For a better comparison, all experimental designs of our studies are depicted in a simplified version in Figure 28. It should be emphasized that none of the studies used block-based probability manipulations (compare Figure 12A). All presented conditions of the respective studies were shown in a randomized or pseudo-randomized manner in fMRI runs but never grouped in blocks with differences in proportions.

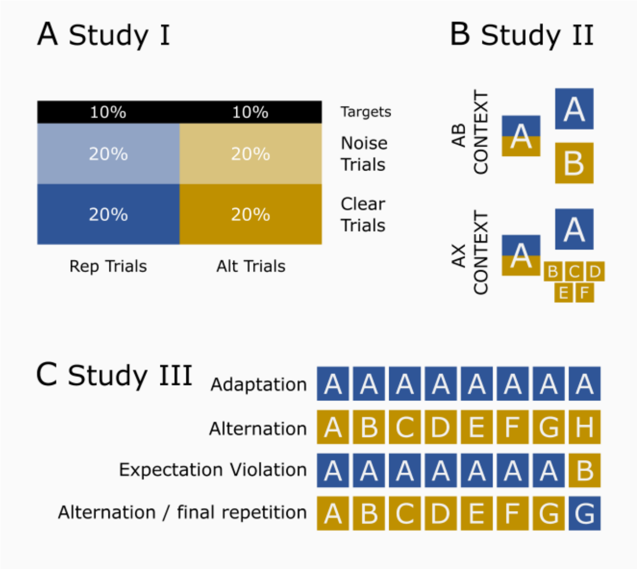


Figure 28. Schematic illustrations of the designs used in the current experiments. Blue = repetition; Gold = alternation. A Study I: Manipulation of sensory precision in an RS design. B Study II: Manipulation of stimulus predictability for repetition and alternation trials in two contexts. C Study III: Manipulation of identity-specific expectation effects in short stimulus sequences. The same letters represent different images from one ID.

Table 4 summarizes the results of our empirical neuroimaging studies (yellow cells) and shows whether the current state of research offers comparable results from other laboratories (blue cells). Only neuroimaging studies investigating the different phenomena in occipito-temporal regions are presented here. Further, only results from univariate analyses are chosen in favor of better comparison to the present studies in Table 4. However, evidence for related effects investigated with other methods or different purposes is still discussed in this section.

Repetition suppression is an often-investigated phenomenon, and its examination for faces is not exempt from this. An entire review article was devoted to ERP results on repetition-related effects for faces alone (Schweinberger & Neumann, 2016) as well as for repetition suppression in the FFA (Henson, 2016). Therefore, evidence for RS to unfamiliar and familiar faces already exists (Table 4 - Repetition suppression). The present results support the finding of RS for unfamiliar faces. However, there is also evidence for the opposite effect for unfamiliar stimuli: repetition enhancement (Henson et al., 2000). Further, we provide evidence for RS for familiarized faces in our second study outside of repetition likelihood and carefully controlling for confounding effects of stimulus novelty. However, as we show that RS magnitude differs according to the predictability of stimulus alternations, we emphasize that the true RS effect might be lower than assumed. RS for famous faces was not tested within our studies but is evident from neuroimaging (Henson et al., 2000) and electrophysiological studies (Neumann & Schweinberger, 2008; Schweinberger et al., 2002) of others.

		Unfamiliar faces	**Visually familiar / familiarized faces**	**Famous familiar faces**
Repetition Suppression		**FFA, OFA (Alink et al., 2018)**	○	**FFA (Henson et al., 2000)**
		FFA, OFA	**FFA, (right) OFA**	/
RS modulation by	Sensory noise/ precision	○	○	○
		No evidence	/	/
	Stimulus predictability	Theoretically impossible	**FFA (Pajani et al., 2017)**	○
		Theoretically impossible	**FFA**	/
Identity-specific effects (image-independent)	Adaptation	**No evidence (Weibert et al., 2016; Davies-Thompson et al., 2013)**	○	**MTL (Weibert et al., 2016); No evidence (Davies-Thompson et al., 2013)**
		/	/	**FFA (+marginal effect in OFA)**
	Expectation Violation	●	○	**FFA, OFA (Amado et al., 2018a)**
		/	/	**No evidence**

Table 4. Results of the current thesis (yellow) and evidence from previous (univariate) neuroimaging studies (blue). FFA: fusiform face area. OFA: occipital face area. MTL: medial temporal lobe. ○ indicates that no other study tested this effect. ● indicates that this effect has been studied with another technique.

Although precision is a significant factor in predictive coding processing (Auksztulewicz & Friston, 2016), and is therefore for prediction-related neuronal effects equally important, the impact of sensory noise on RS has not been tested a lot. To the best of my knowledge, no neuroimaging study investigated measures of RS under different sensory precision conditions to faces (Table 4; RS modulation – sensory noise/precision). Evidence for an effect of sensory uncertainty on RS comes from Turk-Browne et al. (2007). By reducing the contrast in images depicting visual scenes, the authors observed repetition enhancement, whereas RS was observed for high contrast images. As we could not find an effect of visual noise on RS (Study I; 3.1), its contribution to prediction error magnitude might be less critical than prior precision (volatility of environment; Den Ouden et al., 2010; Summerfield et al., 2011).

We found stimulus predictability (Table 4; RS modulation – stimulus predictability) to be a significant factor in repetition designs as it can confound repetition effects (Study II, chapter 3.2). We provide evidence for this by training participants and thereby visually familiarize them with the stimulus material. Although Pajani et al. (2017) also tested the effect of stimulus predictability, their results were confounded with effects of repetition probability and stimulus novelty (see Feuerriegel et al., 2018a for a related EEG study). In order to anticipate upcoming stimuli, a prior neural representation is needed. Therefore, in theory, no such effect can be examined with unfamiliar faces presented for the first time. To the best of my knowledge, no study investigated this effect for famous faces. However, there is evidence for priming effects in studies where the first stimulus was not a face but the name or a different but related face (Amado et al., 2018a; Schweinberger, 1996). Although those paradigms can only be compared to our approach to a limited extent, I include these studies here as they offer different approaches to investigate stimulus predictability.

Although identity-specific adaptation (Table 4; Identity-specific effects – adaptation) has been tested with unfamiliar and familiar face stimuli in the past (Davies-Thompson et al., 2009), the authors found this effect to be image-dependent. In other words, an adaptation effect to identity was only evident when one image was repeated and not when different images of one person were presented. When neuronal responses to different images of the same identity were compared with responses to images from different identities, no effect was present in univariate analyses. A similar study by Weibert et al. (2016) implementing the same stimuli with a much larger sample size found image-independent adaptation effects in the medial temporal lobe (MTL). We provide evidence for an identity-specific and image-independent adaptation effect for famous faces in the FFA and a tendency in the OFA in our third study. This is in accordance with results from MVPA studies providing support for identity discrimination for familiar faces in the FFA (Axelrod & Yovel, 2015) and OFA (Tsantani et al., 2021), as well as unfamiliar faces in ATL, FFA, and OFA (Anzellotti et al., 2014; Nestor et al., 2011). Recent studies show that identity can be decoded from EEG data as well for familiar (Ambrus et al., 2018; Dobs et al., 2019) and unfamiliar faces (Nemrodov et al., 2018). Testing effects for familiarized faces is always more challenging as training, or prolonged experimental settings are needed. However, fast periodic visual stimulation paradigms are promising candidates for addressing this effect for familiarized faces (Verosky et al., 2020).

Expectation violation effects to an unpredictable identity change (Table 4; Identity-specific effects – expectation violation) did not reveal a larger neuronal response attributed to predictive processes in Study III. However, evidence for such a response modulation comes from electrophysiological studies on familiar faces (Johnston et al., 2016) and unfamiliar faces (Robinson et al., 2020). Robinson and colleagues (2020) manipulated identity prediction signals in an MEG study. Theoretically, expectation violation signals to an unknown ID are difficult to investigate as unfamiliar identities are no neural representation has developed. Yet, expectation effects based on the repeated presentation of one ID can be examined. In addition to blocks of five images from different IDs and the same ID, Robinson et al. (2020) included a condition in which ID changed after four images of the same ID. They found prediction error signals in the occipito-temporal gyrus to trials with violation of identity expectation compared to when no expectation was induced. However, although stimuli varied in viewpoint, their stimulus material does not meet the requirements for testing image-independent effects. Therefore, this study does only partially relate to the present one. Amado et al. (2018a) investigated the expectation of a specific famous identity presented within a priming design and found reduced BOLD signal in the FFA and OFA for congruent name-face trials. As they used the name as a cue for the visual input, their finding can be interpreted as an image-independent identity expectation effect. Furthermore, Kouider et al. (2009) found view-independent response reductions in the FFA in a subliminal priming paradigm for famous but not unfamiliar faces. To the best of my knowledge, no study investigated identity expectation effects for familiarized faces.

All in all, the influence of sensory precision, in particular, has not yet been sufficiently investigated. One explanation for why we could not find the effect could be that prior beliefs for repetition are much more precise, hence robust than assumed. Faces as stimuli could enhance this as stimuli since identity is a very strong prior itself. Also, since we have shown the importance of stimulus predictability, I will relate the two in the next chapter.

4.2 Predictive precision in response suppression studies

The importance of weighting prediction errors according to their precision in predictive processing is has been explained in chapter 1.2.4. Here, I will evaluate the current results from the perspective of predictive precision and predictability.

In the first study, we manipulated the sensory precision in a repetition design study, which did not show any impact on RS magnitude. In Study II, we realized a context in which alternations were equally predictable as repetition trials, which nearly abolished RS effects. Although the current results do not directly relate to each other, it stands to reason that precision and predictability are linked (Friston, 2005), which is why these two factors will be put in relation to each other.

4.2.1 The relationship of predictive precision and predictability

According to Bayesian inference, the probability distribution representing a status or event is referred to as a belief (Adams et al., 2014). Before this event is apparent or sensory data are available, beliefs before observation are termed prior beliefs. Those prior beliefs can be precise, represented by a narrow PDF, or imprecise, represented by a broad PDF (compare Figure 13). Possible values in this distribution are concentrated around the mean, representing the expectation since this is the event or state of something most likely to occur. Therefore, precision is always an estimate of the variance of the distribution. In contrast, predictability is a part of the stimulus nature or a stimulus attribute and contributes to estimates of prior precision. In experimental settings, both depend on the implemented design (compare Figure 12 and 28), hence the influence of perceptual expectations and stimulus material. For example, using more than just one exemplar image of identities in Study II would significantly act on precision and stimulus-specific expectations.

Detailed predictability of an upcoming stimulus is the outcome of either an inital presentation of the same stimulus (repetition) or the training on statistical regularities because priors (S1 in a trial) are learned to be reliable. The interleaved nature of our design in Study II ensured implicit learning processes, which is slightly different from previous studies investigating the predictability of face exemplars (Feuerriegel et al., 2018a; Pajani et al., 2017). After seeing face A (sensory input), the prior is updated in terms of precision weighting. The posterior belief is high in precision for AB trials, as repetition and alternations can both be predicted. In AX trials, the former posterior belief, which is now the prior, would be precise for a stimulus repetition but much lower for a stimulus alternation, as one of five faces can be presented. The respective PDFs are illustrated in Figure 29. As probabilities for repetition and alternation are equal, the prior belief for the upcoming stimulus (S2) results from a precise prior for repetition in both contexts but a lower precision for alternation trials in the AX context. Therefore, the posterior belief is determined by the sensory input, and an unpredictable alternation elicits a larger neuronal response.

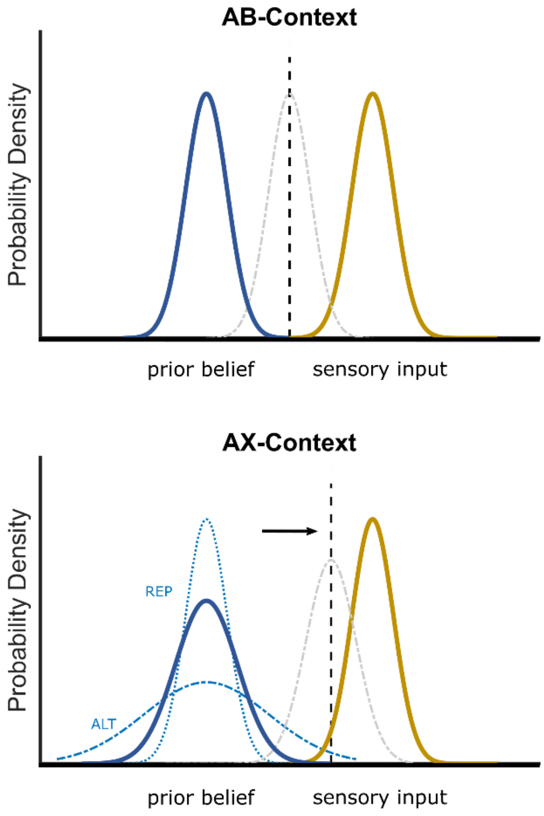


Figure 29. Schematic illustration of predictive precision in Study II after seeing the first stimulus in a trial for AB context (upper panel) and AX context (bottom panel). Probability density (y-axis) as functions for prior beliefs (dark blue), sensory input (gold) and posterior beliefs (grey). Light blue PDFs illustrate theoretical precision for Rep and Alt trials, which was equalized in the AB-Context.

In fact, the probability functions weighted by precision for the AX-context in Figure 29 illustrate prediction error weighting in a standard repetition-alternation design. Assuming the information that a stimulus repetition is more likely than a change is added, as in a repetition block (Summerfield et al., 2008), the priors change according to that. The prior belief would be much more precise because (1) encountering a repetition is the default prior in predictive coding theories, and (2) occurs with a higher probability. Hence, posterior beliefs are mainly determined by prior beliefs. If the prediction of repetition fails to come true, the resulting prediction error is increased. In other words, alternating stimuli elicit a bigger response, and RS is increased (P_{REP} effect).

This can be transferred to Study III, in which we show a reduced response to trials, showing the same ID. The repeated presentation of the same identity, together with identity being a strong prior itself, increases the precision of the prior belief. However, since we could not find an effect for violating this expectation, we cannot prove this assumption.

Logically, stimulus predictability is also strongly connected to surprising stimuli. The term surprise in studies investigating response suppression phenomena as well as in the predictive coding framework is used to describe unpredictable or unexpected sensory input. The role of surprise in experiments examining effects of expectation suppression has been explained in detail in the past (Amado et al., 2016). The authors provide evidence that surprise contributes more to ES measures than the response reduction due to correctly predicted input which is supported by a recent study on MMRs (Feuerriegel et al., 2021b). Surprise effects might have contributed to the results of Study II. However, we tried to control for surprise effects by stimulus novelty, as the full dataset consisted of only twelve stimuli and became highly familiar during the training sessions.

I have already touched on the different forms of predictability throughout this work. Stimuli can also be predictable outside of repeated presentation (Study I) or learned stimulus associations (Study II). For example, studies that examined priming effects used a name or a different but related face as an adapter stimulus for an upcoming face (Schweinberger, 1996). This relates to Study III, as we used famous faces and assumed a stable identity representation to be activated when encountering them. The image-specific adaptation effect could be partly due to the fact that we can predict familiar faces because we have a robust prior activated when we see the first image. Although stimuli can not be predicted accurately, as in Study II, one could imagine some sort of template being activated and used for prediction.

Therefore, I will briefly discuss the various facets of stimulus predictability in experimental design in the next chapter.

4.2.2 Different levels of stimulus predictability

In Study II, we showed the importance of stimulus predictability in response suppression experiments. I have also demonstrated how predictability is related to precision estimations, which play a crucial role in prediction error weighting processes in predictive processing. I will now briefly discuss different forms of predictability and how they are implemented in prediction-related investigations focusing on face stimuli.

The first type of stimulus predictability I will address, refers mainly to using repetition probability to study the effects of expectation (for an overview, see Grotheer & Kovács, 2016). Only the outcome of a trial (Rep versus Alt) can be predicted or expected in those cases – and of course, repeatedly presented stimuli. Given the nature of Rep-Alt paradigms where repetitions are predictable, but alternations are not, in a design that uses blocks with different probabilities of repetition and alternation, this effect is likely to be reinforced and could account partly for the P_{REP} effect (Summerfield et al., 2008). As repetitions occur with a high probability in specific contexts (Rep blocks), observers can predict a repeated image and prior beliefs for encountering a recurrence are more precise as well (compare Figure 29). This naturally carries over to many studies that manipulate the probability of repetition, including those that are cue-based (e.g., Amado et al., 2016).

Moreover, a specific class or category of stimuli can be expected. This is related to the manipulation of repetition probability but refers to the stimulus content. Previous studies showed a modulatory effect of the likelihood of encountering a face versus a house on neuronal responses in face-responsive regions (Egner et al., 2010; Trapp et al., 2016). Although those stimuli were not predictable in a detailed manner, the general structure of faces itself serves as a primitive but still present basis for expectation formations. However, in Table 4, I state that investigating the effect of stimulus predictability on RS for unfamiliar faces is theoretically impossible. I argue that there must be a neuronal representation, no matter how vague, to study these effects. Otherwise, the only input that is predictable is, in fact, the one that has already been shown: a face repetition.

In Study II, we report evidence for the impact of predictability on an image level. It can be argued that our approach of familiarization is not very profound, as we use only one exemplar image per identity, and no further information, such as contextual or semantical, is added. Still, this method of familiarization was sufficient to diminish measures of RS. Therefore, even though the identity of the adapter face stimulus served as a predictor for the following stimulus (in alternating trials), this effect is image-based or exemplar-specific. In other words, a specific face image cued the successor face image in alternation trials. Expectation of a specific image and the capacity to predict that image in detail requires extensive training and a representation of the respective image.

Identity-based cueing would request a more abstract neuronal representation. Evidence for such effects comes mainly from priming studies. Although the relationship between repetition priming effects and RS as measured with fMR-a is complex and not fully understood (Kaiser et al., 2013; Wig et al., 2005), they share certain characteristics (for a review, see Henson & Rugg, 2003). Moreover, priming effects are not restricted to the repetition of a stimulus. Electrophysiological (Jemel et al., 2005; Schweinberger, 1996) and neuroimaging studies (Amado et al., 2018a) could show that presenting the name of a famous identity resulted in neuronal response suppression when the following image matched the primed name. This is known as cross-modal priming and can be related to predictive coding principles. For example, Amado et al. (2018a) observed behavioral priming effects and a reduced BOLD signal in the FFA and OFA for congruent name-face trials. They conclude that these effects result from generated predictions transferred to those areas via feedback or lateral connections. To put it simply, the prime enhanced stimulus predictability via pre-activation of a neuronal representation and thereby facilitates identity recognition via top-down information flow (Ganis & Schendan, 2008). However, the degree of stimulus predictability in such cases is still to be discussed. As already mentioned, the structure of a face serves as a simple basis for predictions alone. Naturally, one would assume that the more a neural representation of a given identity establishes, the more predictability increases. In experiments, this would apply to personally and famous familiar faces. However, Carbon (2008) showed that face identification is strongly impaired for modified and rather unusual images of famous persons. This differentiates from recognition performance for personally familiar faces, which was not affected by different image manipulations. It also suggests that the prediction of a famous face in experimental settings is likely to be more accurate provided typical known example images are used. Such images reflect the ‘iconic’ representation of a famous ID and would therefore match our prediction. However, the presentation of an unusual photograph could abolish or even reverse an effect of facilitated processing by pre-activating the associated neural representation. This should be taken into account when famous familiar faces are used, as in Study III. In contrast, if the representation of a personally known identity is more abstract, the prediction may not be as precise and detailed while being more flexible. Therefore, the potentially observed effects are not as susceptible to substantial changes in sensory input.

I have briefly addressed different levels of stimulus predictability that need to be considered when interpreting prediction-related response alterations. As some of the proposed levels of predictability refer specifically to faces, the next chapter will discuss the use of faces in prediction studies.

4.3 Faces in studies on prediction-related response alterations

The current studies focus on neuronal response alterations due to contextual manipulation. Therefore, a lot of questions regarding the relationship between face processing and predictive coding remain unanswered. Still, I will briefly discuss the current studies in this context.

The more familiar we are with a person, the more information is available and can be transferred to lower-level areas when encountering a familiar face or expecting it (Kovács, 2020). At the same time, the higher we are in the 'hierarchy' of face familiarity, the more complex yet abstract the representations are (Hole, 2011). This needs to be taken into account when exploiting faces as stimulus material for investigating effects that are more or less dependent on the interaction of bottom-up and top-down information flow. In the current thesis, the stimulus material was chosen from unfamiliar faces in Study I and II, whereby faces in the latter became visually familiar to our participants. In Study III, however, we used a set of different photographs of famous identities.

In Study I, we examined RS to unfamiliar clear and noisy faces. Face identification has been shown to be more difficult for degraded images (Hermann et al., 2015), but the implemented task did not demand detailed processing of facial information. Another behavioral task that directs attention to the face stimuli and requires a more differential analysis of them could provide us with a clearer picture of face identification under conditions that, according to PC theories, should modulate prediction errors (Auksztulewicz & Friston, 2016).

In Study II, we investigated the effects of statistical learning of face associations, which led to diminished RS effects in fusiform areas. However, no such effect was observed in the OFA. Even though the participants have undergone intensive training to learn stimulus pairs implicitly and were therefore visually familiarized with the stimulus material, we cannot assume that a stable, robust representation about the depicted identities was formed. Much more information, be it visual or semantical, is needed to constitute neuronal representation for a specific identity (Jenkins & Burton, 2011). Therefore, the different levels of familiarity are related to neural representations that vary in robustness and complexity. However, independent from the degree of familiarity, identity in itself is stable. I have referred to the assumption of the world as being stable on a short-time scale before (Dong & Atick, 1995). This is one of the reasons why RS is related to predictive coding processes. Identity, in fact, is a much more stable construct because the identity of a person does not change, even when contextual aspects change. Therefore, we can define identity (as an aspect of faces used in experiments) as precise prior. Nevertheless, this prior can be shaped by expectations, as we showed by associating two faces in Study II.

In Study III, we showed a strong tendency for image-invariant identity encoding in the OFA, which could result from feedback information from the FFA or even higher-level areas, such as the ATL (Pyles et al., 2013). A stable identity representation is crucial for such an effect, which seems to be evident in the FFA. However, also for the FFA, we cannot exclude feedback information transfer, which should be tested to reveal the underlying mechanisms of image-independent adaptation effects to faces.

The specific representation of famous faces as icons (Carbon, 2008) further supports the idea of predictive coding as the underlying process of face familiarization and face identification. Indeed, some famous persons' faces are commonly shown in a specific way. Therefore, their representation is built on these experiences, and anomalous photographs impede the identification of this person (Carbon, 2008). Similarly, as with unusual contextual information (e.g., when you meet your dentist in a bar), an image from a different period in their life might not be part of our representation. This can potentially affect identity-specific adaptation measures, as in Study III. It might be less relevant for actors and actresses, such as who we know from very different roles in movies. Also, although we used ambient images, no deviating images were part of our dataset. It is still possible that such exemplar images could reduce or even eliminate the effects of adaptation. However, this also offers new possibilities to examine further the predictive processing mechanisms underlying person identification. For example, common and uncommon photographs of famous identities could be implemented as deviants or as violating stimuli in prediction designs.

Although much research on predictive processes and how expectations shape perception has been done, it is still unclear in what format predictions are passed on to lower levels to facilitate perception. By now, we know a lot about how areas are connected and how information is carried (see chapters 1.1.1 and 1.2.1.3). Results from a recent fMRI study show that when the prior expectation of encountering a face is high, activity in face-responsive areas is increased even before the sensory input arrived (Trapp et al., 2016). Could this be transferred to more detailed predictions of expected input? Since we already know a lot about facial representations in different areas, this could be used to study the transfer of predictions (Loffler et al., 2005; Tsantani et al., 2021). For example, if we know that we are about to encounter an old friend, some representation is already active. Stable face representations are robust because they are flexible for variances in visual appearance. Is the prediction precise because we know this person very well, or is it abstract? Addressing this question in an experimental setting could give insight into both face processing and prediction transfer.

4.4 The underlying mechanisms of reduced neuronal responses

Whether RS relies on predictive coding mechanisms or local adaptation mechanisms is still under debate. As pointed out in chapters 1.2.2.1 and 1.2.2.2, several models can explain neuronal response suppression after repeated presentation of stimuli. Whereas studies reporting the P_{REP} effect suggested a top-down component in the form of fulfilled expectations (Kovács et al., 2012; Summerfield et al., 2008), others could not replicate this kind of RS modulation and suggested different underlying mechanisms for RS (Vinken et al., 2018; for a review, see Kovács & Vogels, 2014). The context of higher-order expectations did not modulate RS magnitude for objects (Kovacs et al., 2013) or unfamiliar stimuli (Grotheer & Kovacs, 2014) or when the participants' attention was diverted from the stimuli (Larsson & Smith, 2012). Even when behavioral priming was observed, suggesting an advantage for recognition processes, an effect in related neuroimaging data was absent (Olkkonen et al., 2017). Furthermore, RS and ES were found to be independent in neuroimaging (Grotheer & Kovács, 2015) and electrophysiological measurements (Feuerriegel et al., 2018a). Thus, the question of under which circumstances RS varies in its magnitude is still open.

A two-stage model of response suppression (Figure 30) was proposed for explaining inconsistent findings (see, for example, Kaliukhovich & Vogels, 2011; Summerfield et al., 2011, 2008; Vinken et al., 2018) for such effects by Grotheer and Kovács (2016). In their model, response suppressions resulting from stimulus repetitions are explained by low-level inferential processes within the ventral visual stream. Expectation-related effects showing response reductions result from higher-level predictions that originate in frontal areas. Empirical studies can support this differentiation (e.g., Jiang et al., 2013; Summerfield et al., 2011; Todorovic & de Lange, 2012). A DCM study on neuronal connectivity changes in EEG signals revealed repetitions to systematically change connectivity within and between neuronal regions (Garrido et al., 2009a). Garrido et al. (2009a) proposed that those connectivity changes are related to the predictability of repetitions (changes in extrinsic connectivity) and prior precision (changes in intrinsic connectivity). This again supports the idea that both local within-region and hierarchical between-region modulatory processes characterize RS. As of now, it is very likely that a combination of multiple mechanisms underlies RS effects. Therefore, it is important to investigate under which circumstances feedback from higher-order processing levels shapes responses of local adaptation mechanisms and when not.

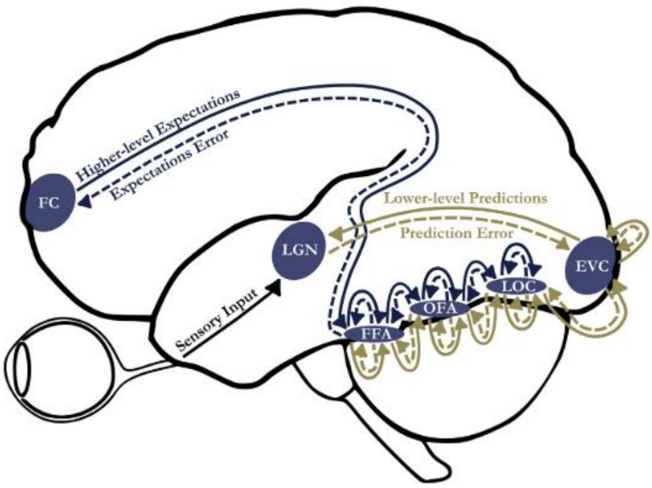


Figure 30. Schematic illustration of the two-stage model of response suppression in the ventral visual stream. Higher-level expectations (blue) originate in frontal regions and are transmitted to the FFA, the OFA and the LOC. Lower-level predictions (gold) originate in the FFA and are transmitted, via the OFA, the LOC and the EVC, to the LGN. FC: frontal cortex, FFA: fusiform face area, OFA: occipital face area, LOC: lateral occipital complex, EVC: early visual cortex, LGN: lateral geniculate nucleus. Adapted from Grotheer & Kovács (2016) with permission.

Changing the precision of prior beliefs has been shown to modulate RS magnitude, as proposed by PC models (Summerfield et al., 2011). However, differences in the precision of sensory input in a simple RS design in Study I (chapter 3.1) did not resolve in a change of RS magnitude. This suggests that RS can derive from local neuronal mechanisms and computations that can be modulated by top-down mechanisms when put into a context in which higher-order information is present (e.g., probability of repetition). In such a design, sensory input precision might play a significant role in determining RS magnitude, too. Although we cannot exclude inferential processes underlying RS with the present design, manipulating the sensory precision rather than priors might be a better way to test for this assumption.

In Study II, we showed that controlling for stimulus predictability can lead to similar response reduction in alternation trials as in repetition trials. RS could therefore be partly explained by stimulus predictability as well. Although unpredictable stimulus alternations were also part of the training session, they still elicited a higher neuronal response than any other condition in Study II. A generally enhanced response to unpredictable stimuli is in accordance with the assumption that feedback connections are mostly inhibitory (although inhibitory effects can be modulated, Bastos et al., 2012) because in those cases, higher cortical areas are unable to suppress activity (Meyer & Olson, 2011; Todorovic et al., 2011; Wacongne et al., 2011). Furthermore, the counterpart – smaller responses to predictable stimuli - cannot be easily explained by mechanisms such as local adaptation processes because those stimulus features are not represented in earlier cortical regions (Bastos et al., 2012). The two-stage model (Grotheer & Kovács, 2016) proposes repetition-related effects depend on low-level inference processes within the ventral visual stream. This is in accordance with our results, including activations in fusiform and middle temporal regions that were found in an exploratory whole-brain analysis when contrasting alternation trials and repetition trials in Study II (see chapter 3.2, supplementary material). Contrasting the two contexts (AX versus AB), however, revealed activation clusters in frontal regions. At first glance, this seems to fit with the proposed model as expectation-related predictions are assumed to originate in frontal areas. Critically, as the AB context revealed the stimulus-specific expectation effects (i.e., a reduced response to predictable alternation stimuli after extensive training), these additional activations of higher-level areas are not easy to explain. Frontal activations for the AX context might reflect more difficult computations for these trials. It is important to further elucidate the processes within and between cortical regions to understand underlying mechanisms of RS (Ewbank et al., 2013; Garrido et al., 2009a; Kohn & Movshon, 2003). We have focused on the FFA, which is, according to Grotheer and Kovács (2016), a good candidate for examining prediction-related effects for faces. Indeed, effects of stimulus-specific expectations were absent in OFA data.

As I pointed out in chapter 1.3, face identification requires inferential processing in the hierarchy of face-processing regions. Therefore, disentangling bottom-up and top-down processes in face perception is challenging. Top-down processes can be demonstrated as early as in the course of detecting faces (Brodski-Guerniero et al., 2017; Ganis & Schendan, 2008). Surely, processing familiar faces go along with a large proportion of higher-level influences. As we presented face images of one identity that differed in visuospatial properties in Study III, a complex network of areas likely contributes to perceiving them as belonging to the same person. Study III was not designed to elucidate neuronal mechanisms underlying RS but rather addressed expectation-related response alterations to identity. In fact, fMR-a studies the consequences of adaptation on a neuronal population level, and therefore, the exact underlying neuronal mechanisms are inaccessible (Larsson et al., 2016). An image-independent identity-specific adaptation effect was evident in the FFA, suggesting that this area does encode identity. However, the hypothesized PE response to an unexpected change of identity was absent. We assume that this effect can be found by employing methods and analyses with higher spatial and temporal resolution.

In summary, different methods capture different stimulus processing mechanisms, and those mechanisms, even if their measured output is the same (a reduced neural response), are likely to vary a lot in their contribution to what is finally measured when different designs, stimulus material, or tasks are applied. All here presented experiments rely on neuroimaging data, and therefore their results reflect the activity of a large neuronal population. The current studies do not provide direct evidence for a top-down mechanism producing (Study I) RS. However, as stimulus predictability is capable of mimicking RS effects (Study II), it can argue that measures labeled as RS can indeed be the outcome of several different neuronal computations - among them are local neuronal scaling mechanisms (Alink et al., 2018), inherited adaptation effects from upstream regions (Kohn & Movshon, 2003), microcircuit computations (Bastos et al., 2012; Westerberg et al., 2019), and additional hierarchical inference processes (Ewbank et al., 2011; Garrido et al., 2009a).

4.5 Limitations of the presented studies

In this section I will discuss some limitations of our investigations.

4.5.1 Implemented tasks and attention

A critical aspect of experiments testing prediction-related effects, not only in behavioral but neuroimaging and electrophysiological investigations, is the task chosen for the participants to perform. One of the reasons is the direct effect the selected task has on the focus of attention. Whenever two or more conditions are compared, differences in attention to these can potentially confound results because significant effects of conditions can result from different attentional loads instead of experimental manipulations. In general, stimuli evoke higher neuronal responses when attended compared to unattended (Maunsell & Cook, 2002; Moran & Desimone, 1985). Therefore, it is crucial to choose a task that either diverts attention away from the stimuli features that make the experimental manipulation or draws attention equally to everything, depending on the research question. In the first of the present studies, participants had to perform a target-detection task, and logically those target trials were excluded from further analyses of the neuroimaging data. However, it is theoretically possible that clear trials draw more attention than noisy trials, although we would assume to observe an effect of noise level on RS in this case. As already mentioned in chapter 1.2.4, attention is related to predictive precision and can be a confound in prediction error experiments. Kok and colleagues (2012) investigated prediction error signals to gratings in human V1, V2, and V3 using fMRI. They manipulated prediction and attention and found their interaction. Their results revealed that attention could reverse common prediction-related response suppression. Therefore, if attention is allocated to clear and noisy trials differently, this could influence response suppression.

Another example of how important the choice of the behavioral task is, comes from an fMRI study by Summerfield and Koechlin (2008). They used simple perceptual decision-making tasks, one of which biased decisions towards one choice and the other did not. Specifically, participants were asked to decide whether a target matches or not matches cue in orientation in one task. In other parts of the experiment, participants were asked to make an A versus B decision. During blocks in which participants had to decide whether a target matched the cue occipital and fusiform regions showed higher activity than during an A-versus-B task (Summerfield & Koechlin, 2008). The authors reported that behavioral results also revealed different strategies for the two tasks. An advantage in reaction times for a ‘same’ in comparison with a ‘different’ answer was termed as the ‘fast-same effect’ (for a review on this phenomenon, see Farell, 1985). Recently this effect has been explained in the predictive coding framework (Friston, 2005). In Summerfield and Koechlin (2008), the bias towards the ‘same’ answer results from top-down mechanisms. The advantage in behavioral responses relies on the faster accumulation of information that passes the threshold for a decision. We have implemented this kind of task in the second of the reported studies. Therefore, we cannot exclude attention as contributing to our results. However, implementing another task would likely have impeded learning of the statistical regularities, which is crucial in this design.

4.5.2 Stimulus sets

In Study I, we created noisy stimuli by reducing phase coherence. The absence of an effect of sensory precision on RS cannot be clearly explained. However, one possibility is that our stimulus processing procedures were not adequate for this purpose. However, since we found an effect of noise in the behavioral results and neuroimaging results, this is unlikely.

In Study III, ambient images of different celebrities were chosen. Since the familiarity of individuals is constantly changing depending on which sample one chooses, consideration should be given to periodically testing the stimuli for familiarity in behavioral experiments.

In this step, one could also test the stimulus set directly for highly deviant stimuli. Although a stable representation should include such exemplar images as well, it has been shown that this is not true for particular cases (Carbon, 2008).

4.5.3 Methodological limitations

As we measure the BOLD signal with fMRI, one disadvantage is the low temporal resolution – especially when compared to electrophysiological methods. This is important to keep in mind when thinking about the characteristics of specific phenomena that are going to be manipulated. Therefore, neuronal timing differences for the first and second presentation of stimuli proposed, for example, by the fatigue model for RS, are difficult to test. This applies as well to differences in temporal dynamics of RS and ES. Although it is not impossible to investigate those with fMRI (Grotheer & Kovács, 2015), electrophysiological methods are much more appropriate.

However, there is good development regarding the timing parameters when using fMRI, meaning the TR can be reduced. To be more precise, the time between two MR pulses can be shortened to approximately 500 ms. This is mostly achieved by multi-slice acquisition at one timepoint (Xu et al., 2013) and provides a significant advantage for specific paradigms, especially for a study with a paradigm used in Study III (chapter 3.3). The possibility to separate responses to the single stimuli within a trial, be it a trial with two or more stimuli, offers the opportunity to classify the observed phenomena more thoroughly (Pajani et al., 2017). In the case of Study III (3.3), we probably would have been able to improve the signal-to-noise ratio. Although there is evidence that differences between trials containing only standards and trials with a deviant stimulus even in a block design (Amado et al., 2018b) can be detected, we could not find this effect when testing for identity-related violation in a similar design. The separation of the signal to the single images within blocks would have allowed to examine only the outcome of the trials and investigate the difference between fulfilled and violated identity expectations independently from responses to the preceding stimuli.

The high spatial resolution has been considered to be the advantage of fMRI over EEG. In light of the presented phenomena of response suppression findings and keeping in mind the many possible underlying processes, the spatial resolution is not high enough to make specific statements about how the individual mechanisms work together or distinguish them from each other. Although most models of RS predict changes in amplitude, which can be measured with fMRI, it has been shown that univariate analyses fail to differentiate between models (Weiner et al., 2010). In the last years, multivariate pattern analyses have been offered more complex investigative approaches in neuroscience. An example that shows that these analyses can be superior to conventional ones comes from research on differences in face identity representations in different levels of familiarity (for a review, see Kovács, 2020). Nevertheless, we were able to prove identity-specific encoding that is image-invariant in the FFA with univariate analyses.

As already mentioned, one disadvantage of fMRI methods applied in the current experiment is that only averaged activity can be measured. In other words, the acquired neuroimaging data collapse the activity of multiple neuronal populations. Therefore, it is unable to differentiate between activity driven by bottom-up and top-down processing streams. Laminar fMRI (Lawrence et al., 2019) is a promising technique to tackle specific characteristics of processing stimuli in an experimental setting.

A more general limitation is the chosen sample size. Obviously, the best approach to estimate the needed sample size reliably in a power analysis is conducting a pilot experiment. However, due to limited resources, this is often not applicable for fMRI experiments. We therefore followed other studies that have investigated related phenomena to determine the sample size. We acknowledge that samples in Study I and II were rather small. However, in Study I, we conducted a Bayes factor ANOVA additionally to traditional ANOVA, to demonstrate that the absence of an RS modulation by sensory precision is not due to the small sample size. In Study II, our data collection was severely limited by external conditions due to the strictly defined procedure of the training with subsequent fMRI measurement.

4.6 Open questions and future directions

In the course of the discussion, I have already mentioned some considerations that could improve our studies or initiate new research. I would like to briefly review some of the already mentioned methodological approaches and techniques with their specific advantages for investigating predictive processes in the ventral visual stream. Furthermore, I will propose some ideas for follow-up studies based on the presented investigations.

I have already mentioned the possibility that attention is allocated differently to noisy and clear stimuli. Different accuracies for noisy and clear stimuli suggest that clear trials draw more attention. Nevertheless, our sensory precision manipulation did not affect RS magnitude, although attention and precision are closely related as proposed by PC models.

One possible explanation is that the task was indeed more difficult for noisy trials (Banko et al., 2011). Implementing a different task could provide better insights into the influence of sensory precision and attention. Furthermore, based on the present results, we cannot conclude that RS should be explained in terms of feedforward models. Another manipulation of sensory precision or a task demanding processing the stimulus content could shed light on whether RS reflects the prediction error shaped by precision weighting. However, to not introduce more predictive processes just by demanding face identification, sensory noise could be manipulated differently. For example, reducing the stimulus presentation time can also change uncertainty in the sensory signal.

In Study II (chapter 3.2), we showed that measures of repetition suppression might be confounded with effects of different degrees of stimulus predictability in repetition and alternation trials. Through training, the participants were able to predict stimulus changes as well as stimulus repetitions. This led to reduced response differences between repetition and alternation trials in this context (i.e., no repetition suppression effect). Different processes likely led to response suppression for repetition and alternation trials in the trained AB context. To further investigate top-down and bottom-up influences, it would be insightful to examine the timing parameters of these effects with electrophysiological methods (Feuerriegel et al., 2018a).

We have also raised the question of whether the differences in the relative likelihood of specific face images in the two contexts led to weighting expectations towards repetitions in the AX context. To further test the impact of predictability, a graded manipulation of predictability could be implemented. In Study II, alternating faces in context AB were always predictable. In contrast, in context AX, predictability was downgraded as face X was always one of five faces that occurred with equal probability. If, however, probabilities in the AX context were changed and biased towards one face, one could investigate the impact of the relative likelihood of encountering a specific face and shed light on the role of surprise.

In chapter 4.2.2, I emphasize that different levels of stimulus predictability exist, especially when utilizing faces for studies on response suppression. One can predict the outcome of a trial (repeating – alternating), the class of stimuli (faces – houses), the identity depicted in an image (as in priming studies), or a specific image is in Study II (chapter 3.2). We visually familiarized our participants with the face stimuli. However, we used only one exemplar image per identity to ensure that the exact image can be predicted similar to repeatedly presented stimuli. The use of ambient images (as in Study III) could strengthen a neural representation. This could be one way of investigating stimulus predictability effects on a more abstract level. Of course, it is possible to use famous familiar faces right away, as is associative priming studies (Schweinberger et al., 1995). However, one advantage of previously unknown faces is that familiarity levels are equal at the start, and no associations between faces are present. Whereas familiarity with stimulus material can be assessed easily by a questionnaire in experiments, the evaluation of associations participants might have for specific celebrities is difficult.

In Study III, we provided evidence for identity-specific adaptation effects in the FFA and a tendency for the same effect in the OFA. Although previous univariate studies failed to find these effects with similar designs (e.g., Davies-Thompson et al., 2009), multivariate approaches were successfully proved identity representations in those areas (Tsantani et al., 2021). This suggests that more powerful analyses could provide a more robust effect of image-independent adaptation to identity for the FFA and OFA and maybe even show the identity-specific release from adaptation when identity changes. A related effect in a slightly different design was shown by Johnston et al. (2016) in event-related potentials and in an MEG study, even coined to the occipito-temporal gyrus (Robinson et al., 2020). Therefore, a combination of univariate and multivariate analyses for fMRI is recommended for future studies. This could also elucidate how stimulus predictability affects different measures of response suppression as it was shown that predictability reduces activity as measured with fMRI but enhanced classification performance in multivariate analyses (Kok et al., 2012a). However, paradigms have to be planned carefully as not every dataset is suitable for multivariate analyses.

The potential of faces as stimulus material is not yet fully exhausted. Faces lend themselves a good basis for investigating basic principles of visual perception. However, they need to be chosen with special care. Depending on their familiarity characterization ranging from unfamiliar to personally familiar, processing faces engages different cortical (and subcortical) areas and is likely influenced by higher-order information on different scales (Duchaine & Yovel, 2015; Kovács, 2020). As faces are of special interest to humans, their processing is, in general, very complex. That comes with the disadvantage that slightly different stimulus sets may cause significantly different results. Future studies should aim for better integration of knowledge about face processing and perception when using them to examine neuronal responses related to predictive processes.

Undoubtedly, the FFA is a good candidate for investigating response alterations to face stimuli under different manipulations. The FFA is engaged in face processing irrespective of familiarity, plays a crucial role in identity processing (chapter 3.2), and is highly interconnected with other areas (Duchaine & Yovel, 2015). However, in addition to lower-level areas, such as the OFA, future studies should aim to localize higher-level areas such as the ATL.

5 CONCLUSION

The empirical studies in this thesis investigate response alterations in occipito-temporal face areas measured with fMRI. We specifically focus on measures of neuronal behavior resulting from stimulus repetition, fulfillment of image-specific expectations, and higher-level identity expectations. Furthermore, we highlight the importance of stimulus predictability in studies investigating the mentioned effects focusing on facial stimulus material.

Whereas many studies show the modulation of RS by expectations, and it therefore has recently been explained by theories of PC, other findings support intrinsic neuronal adaptation mechanisms to cause this effect. Under which conditions RS reflects either of them is therefore not finally clarified. As predicted by hierarchical inference models, measures of RS vary according to the precision of prior beliefs. We were the first to show that RS to faces is not modulated by sensory precision as it was shown for the precision of priors (Summerfield et al., 2011) in Study I. FFA responses to unfamiliar faces revealed robust RS under low and high sensory precision conditions. This supports feedforward models of RS or within-region computational processes based on neuronal circuits. However, as the measured BOLD signal reflects activity on a population level, we cannot differentiate signals of neuronal units representing prediction and prediction errors. Therefore, we cannot exclude RS to be the consequence of inferential processes and perceptual expectations. Nevertheless, if RS reflects the prediction error, the precision of sensory input seems to be less determining its magnitude than the precision of prior beliefs. One possible explanation is that repetition as prior is much more robust than assumed.

We also show the importance of stimulus predictability in measures of RS in Study II. As many studies use RS under different conditions to measure feedback influences of higher levels in the processing hierarchy, the imbalance of stimulus predictability as characteristic of immediate RS designs needs to be considered for correct conclusions. A relatively short amount of training on statistical regularities (compared to generating internal models of the outside world on a long-term scale), significantly altered neuronal responses to faces. Stimulus-specific expectations, evolved after training on specific stimulus associations, eliminated RS. Although this issue was addressed earlier (Pajani et al., 2017), we were the first to investigate this effect independent from higher-level expectations and, in addition, carefully accounted for effects of stimulus novelty. Our results emphasize that stimulus predictability needs to be considered when investigating RS. Especially in paradigms where higher-order expectations are examined using measures of RS, resulting RS magnitudes can be confounded with effects of stimulus-specific expectations.

In Study III, we provide further evidence for image-independent identity encoding of famous faces in the FFA. This supports the idea of stable identity representations in this area. In contrast, earlier studies could find response differences only in higher-order face-processing areas (Weibert et al., 2016) or with multivariate analyses (Axelrod & Yovel, 2015). We even found a tendency for the same effect in the OFA, which needs further investigation due to the explained limitations. Setting about those could also shed light on image-invariant expectation violation effects, signaling prediction errors, which could not be observed within this paradigm.

Additionally, on the basis of our work, we suggest that faces can provide a great basis for investigating neuronal mechanisms as we can draw on a lot of knowledge about the individual areas and their connections. However, they need to be chosen with care, especially with regard to familiarity and the accompanying neuronal representations and resulting differences in predictive precision.

Although our studies addressed hypotheses based on predictive coding principles, such as precision weighting of prediction errors, stimulus-specific expectations, hence stimulus predictability and identity-specific expectations, we cannot provide direct evidence for or against RS being a consequence of predictive processing in ventral visual areas. However, as it has been shown that several proposed models can explain neuronal behavior to stimulus repetitions, our studies are a part of the bigger puzzle. In the end, our brain is a dynamic system, and there is still a long way to go before we understand how it functions.

6 REFERENCES

Adams, R.A., Brown, H.R., Friston, K.J., 2014. Bayesian inference, predictive coding and delusions. AVANT. J. Philos. Vanguard V, 51–88. https://doi.org/10.26913/50302014.0112.0004

Adams, R.A., Stephan, K.E., Brown, H.R., Frith, C.D., Friston, K.J., 2013. The Computational Anatomy of Psychosis. Front. Psychiatry 4, 1–26. https://doi.org/10.3389/fpsyt.2013.00047

Ahlheim, C., Schiffer, A.-M.M., Schubotz, R.I., 2016. Prefrontal cortex activation reflects efficient exploitation of higher-order statistical structure. J. Cogn. Neurosci. 28, 4–5. https://doi.org/10.1162/jocn_a_01005

Aitchison, L., Lengyel, M., 2017. With or without you: predictive coding and Bayesian inference in the brain. Curr. Opin. Neurobiol. https://doi.org/10.1016/j.conb.2017.08.010

Alink, A., Abdulrahman, H., Henson, R.N., 2018. Forward models demonstrate that repetition suppression is best modelled by local neural scaling. Nat. Commun. 9. https://doi.org/10.1038/s41467-018-05957-0

Alink, A., Blank, H., 2021. Can expectation suppression be explained by reduced attention to predictable stimuli? Neuroimage. https://doi.org/10.1016/j.neuroimage.2021.117824

Alink, A., Schwiedrzik, C.M., Kohler, A., Singer, W., Muckli, L., 2010. Stimulus Predictability Reduces Responses in Primary Visual Cortex. J. Neurosci. 30, 2960–2966. https://doi.org/10.1523/JNEUROSCI.3730-10.2010

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. https://doi.org/10.1016/j.neuropsychologia.2016.02.003

Amado, C., Kovács, P., Mayer, R., Ambrus, G.G., Trapp, S., Kovács, G., 2018a. Neuroimaging results suggest the role of prediction in cross-domain priming. Sci. Rep. 8. https://doi.org/10.1038/s41598-018-28696-0

Amado, C., Stoyanova, P., Kovács, G., 2018b. Visual mismatch response and fMRI signal adaptation correlate in the occipital-temporal cortex. Behav. Brain Res. 347, 77–87.

https://doi.org/10.1016/j.bbr.2018.03.001

Ambrus, G.G., Dotzer, M., Schweinberger, S.R., Kovács, G., 2017a. The occipital face area is causally involved in the formation of identity-specific face representations. Brain Struct. Funct. 222, 4271–4282. https://doi.org/10.1007/s00429-017-1467-2

Ambrus, G.G., Kaiser, D., Cichy, R.M., Kovács, G., 2018. The neural dynamics of familiar face recognition. Words count Abstr. 201. https://doi.org/10.1101/393652

Ambrus, G.G., Windel, F., Burton, A.M., Kovács, G., 2017b. Causal evidence of the involvement of the right occipital face area in face-identity acquisition. Neuroimage 148, 212–218. https://doi.org/10.1016/j.neuroimage.2017.01.043

Anzellotti, S., Fairhall, S.L., Caramazza, A., 2014. Decoding representations of face identity that are tolerant to rotation. Cereb. Cortex 24, 1988–1995. https://doi.org/10.1093/cercor/bht046

Apps, M.A.J., Tsakiris, M., 2013. Predictive codes of familiarity and context during the perceptual learning of facial identities. Nat. Commun. 4, 1–10. https://doi.org/10.1038/ncomms3698

Arnal, L.H., Giraud, A.L., 2012. Cortical oscillations and sensory predictions. Trends Cogn. Sci. 16, 390–398. https://doi.org/10.1016/j.tics.2012.05.003

Ashburner, J., Chen, C., Moran, R., Henson, R., Phillips, C., n.d. SPM12 Manual The FIL Methods Group (and honorary members) 1–508.

Attneave, F., 1954. Some informational aspects of visual perception. Psychol. Rev. 61, 183–193. https://doi.org/10.1037/h0054663

Auksztulewicz, R., Friston, K., 2016. Repetition suppression and its contextual determinants in predictive coding. Cortex 80, 125–140. https://doi.org/10.1016/j.cortex.2015.11.024

Axelrod, V., Yovel, G., 2015. Successful decoding of famous faces in the fusiform face area. PLoS One 10, 19–25. https://doi.org/10.1371/journal.pone.0117126

Axelrod, V., Yovel, G., 2013. The challenge of localizing the anterior temporal face area: A possible solution. Neuroimage 81, 371–380. https://doi.org/10.1016/j.neuroimage.2013.05.015

Baker, K.A., Mondloch, C.J., 2019. Two Sides of Face Learning: Improving Between-Identity Discrimination While Tolerating More Within-Person Variability in Appearance.

Perception 48, 1124–1145. https://doi.org/10.1177/0301006619867862

Baldeweg, T., 2007. ERP repetition effects and mismatch negativity generation: A predictive coding perspective. J. Psychophysiol. https://doi.org/10.1027/0269-8803.21.34.204

Banko, E.M., Gal, V., Kortvelyes, J., Kovacs, G., Vidnyanszky, Z., 2011. Dissociating the Effect of Noise on Sensory Processing and Overall Decision Difficulty. J. Neurosci. 31, 2663–2674. https://doi.org/10.1523/JNEUROSCI.2725-10.2011

Bar, M., 2004. Visual objects in context. Nat. Rev. Neurosci. https://doi.org/10.1038/nrn1476

Barbosa, L.S., Kouider, S., 2018. Prior Expectation Modulates Repetition Suppression without Perceptual Awareness. Sci. Rep. 8. https://doi.org/10.1038/s41598-018-23467-3

Barragan-Jason, G., Cauchoix, M., Barbeau, E.J., 2015. The neural speed of familiar face recognition. Neuropsychologia 75, 390–401. https://doi.org/10.1016/j.neuropsychologia.2015.06.017

Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P., Friston, K.J., 2012. Canonical Microcircuits for Predictive Coding. Neuron. https://doi.org/10.1016/j.neuron.2012.10.038

Bayes T, Price R (1763). "An Essay Towards Solving a Problem in the Doctrine of Chances. By the late Rev. Mr. Bayes, communicated by Mr. Price, in a letter to John Canton, MA. and F.R.S." Philosophical Transactions of the Royal Society of London, 53, 370-418

Bendixen, A., Schröger, E., Winkler, I., 2009. I heard that coming: Event-related potential evidence for stimulus-driven prediction in the auditory system. J. Neurosci. 29, 8447–8451. https://doi.org/10.1523/JNEUROSCI.1493-09.2009

Brodski-Guerniero, A., Paasch, G., Wollstadt, P., Özdemir, I., Wibral, M., Lizier, J., 2017. Information-Theoretic Evidence for Predictive Coding in the Face-Processing System. J. Neurosci. 37. https://doi.org/10.1523/JNEUROSCI.0614-17.2017

Brodski, A., Paasch, G.F., Helbling, S., Wibral, M., 2015. The faces of predictive coding. J. Neurosci. 35, 8997–9006. https://doi.org/10.1523/JNEUROSCI.1529-14.2015

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P.J.B., Burton, A.M., Miller, P., 1999. Verification of face identities from images captured on video. J. Exp. Psychol. Appl. 5, 339–360. https://doi.org/10.1037/1076-898X.5.4.339

Bruce, V., Henderson, Z., Newman, C., Burton, A.M., 2001. Matching identities of familiar and unfamiliar faces caught on CCTV images. J. Exp. Psychol. Appl. 7, 207–218. https://doi.org/10.1037/1076-898X.7.3.207

Bruce, V., Young, A., 1986. Understanding face recognition Copyright. Br. J. Psychol. 77, 305–327.

Buchbinder, B.R., 2016. Functional magnetic resonance imaging, in: Handbook of Clinical Neurology. Elsevier B.V., pp. 61–92. https://doi.org/10.1016/B978-0-444-53485-9.00004-0

Bunzeck, N., Thiel, C., 2016. Neurochemical modulation of repetition suppression and novelty signals in the human brain. Cortex 80, 161–173. https://doi.org/10.1016/j.cortex.2015.10.013

Burton, A.M., Jenkins, R., Hancock, P.J.B., White, D., Burton, M., 2005. Robust representations for face recognition: the power of averages, Cognitive Psychology.

Burton, A.M., Wilson, S., Cowan, M., Bruce, V., 1999. FACE RECOGNITION IN POOR-QUALITY VIDEO: Evidence From Security Surveillance. Psychol. Sci. 10, 243–248.

Carandini, M., Ferster, D., 1997. A Tonic Hyperpolarization Underlying Contrast Adaptation in Cat Visual Cortex. Science (80-.). 276, 949–952. https://doi.org/10.1126/science.276.5314.949

Carbon, C., 2008. Famous faces as icons . The illusion of being an expert in the recognition of famous faces. Perception 37, 801–807. https://doi.org/10.1068/p5789

Carbon, C.C., Strobach, T., Langton, S.R.H., Harsányi, G., Leder, H., Kovács, G., 2007. Adaptation effects of highly familiar faces: Immediate and long lasting. Mem. Cogn. 35, 1966–1976. https://doi.org/10.3758/BF03192929

Casagrande, V.A., Royal, D.W., Sáry, G., 2006. Extraretinal Inputs and Feedback Mechanisms to the Lateral Geniculate Nucleus (LGN), in: The Primate Visual System. John Wiley & Sons, Ltd, Chichester, UK, pp. 191–211. https://doi.org/10.1002/0470868112.ch7

Chen, C.C., Henson, R.N., Stephan, K.E., Kilner, J.M., Friston, K.J., 2008. Forward and backward connections in the brain: A DCM study of functional asymmetries. https://doi.org/10.1016/j.neuroimage.2008.12.041

Chiou, R., Lambon Ralph, M.A., 2018. The anterior-ventrolateral temporal lobe contributes to boosting visual working memory capacity for items carrying semantic information. Neuroimage 169, 453–461. https://doi.org/10.1016/j.neuroimage.2017.12.085

Churchill, N.W., Oder, A., Abdi, H., Tam, F., Lee, W., Thomas, C., Ween, J.E., Graham, S.J., Strother, S.C., 2012. Optimizing preprocessing and analysis pipelines for single-subject fMRI. I. Standard temporal motion and physiological noise correction methods. Hum. Brain Mapp. 33, 609–627. https://doi.org/10.1002/hbm.21238

Clark, A., 2018. A nice surprise? Predictive processing and the active pursuit of novelty. Phenomenol. Cogn. Sci. 17, 521–534. https://doi.org/10.1007/s11097-017-9525-z

Clark, A., 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav. Brain Sci. 36, 181–204. https://doi.org/10.1017/S0140525X12000477

Czigler, I., 2007. Visual mismatch negativity: Violation of nonattended environmental regularities. J. Psychophysiol. 21, 224–230. https://doi.org/10.1027/0269-8803.21.34.224

Davies-Thompson, J., Gouws, A., Andrews, T.J., 2009. An image-dependent representation of familiar and unfamiliar faces in the human ventral stream. Neuropsychologia 47, 1627–1635. https://doi.org/10.1016/j.neuropsychologia.2009.01.017

Davies-Thompson, J., Newling, K., Andrews, T.J., 2013. Image-invariant responses in face-selective regions do not explain the perceptual advantage for familiar face recognition. Cereb. Cortex 23, 370–377. https://doi.org/10.1093/cercor/bhs024

De Baene, W., Vogels, R., 2010. Effects of adaptation on the stimulus selectivity of macaque inferior temporal spiking activity and local field potentials. Cereb. Cortex 20, 2145–2165. https://doi.org/10.1093/cercor/bhp277

de Lange, F.P., Heilbron, M., Kok, P., 2018. How Do Expectations Shape Perception? Trends Cogn. Sci. 22, 764–779. https://doi.org/10.1016/j.tics.2018.06.002

Den Ouden, H.E.M., Daunizeau, J., Roiser, J., Friston, K.J., Stephan, K.E., 2010. Striatal prediction error modulates cortical coupling. J. Neurosci. 30, 3210–3219. https://doi.org/10.1523/JNEUROSCI.4458-09.2010

Den Ouden, H.E.M., Kok, P., De Lange, F.P., 2012. How prediction errors shape perception, attention, and motivation. Front. Psychol. 3. https://doi.org/10.3389/fpsyg.2012.00548

Desimone, R., 1996. Neural mechanisms for visual memory and their role in attention. Proc. Natl. Acad. Sci. U. S. A. 93, 13494–13499. https://doi.org/10.1073/pnas.93.24.13494

Diamond, R., Carey, S., 1986. Why Faces Are and Are Not Special. An Effect of Expertise. J. Exp. Psychol. Gen. 115, 107–117. https://doi.org/10.1037/0096-3445.115.2.107

Dobs, K., Isik, L., Pantazis, D., Kanwisher, N., 2019. How face perception unfolds over time. Nat. Commun. 10, 1258. https://doi.org/10.1038/s41467-019-09239-1

Dong, D.W., Atick, J.J., 1995. Statistics of natural time-varying images. Netw. Comput. Neural Syst. 6, 345–358. https://doi.org/10.1088/0954-898X

Duchaine, B., Yovel, G., 2015. A Revised Neural Framework for Face Processing. Annu. Rev. Vis. Sci. 1, 393–416. https://doi.org/10.1146/annurev-vision-082114-035518

Eger, E., Schweinberger, S.R., Dolan, R.J., Henson, R.N., 2005. Familiarity enhances invariance of face representations in human ventral visual cortex: fMRI evidence. Neuroimage 26, 1128–1139. https://doi.org/10.1016/j.neuroimage.2005.03.010

Egner, T., Monti, J.M., Summerfield, C., 2010. Expectation and Surprise Determine Neural Population Responses in the Ventral Visual Stream. J. Neurosci. 30, 16601–16608. https://doi.org/10.1523/JNEUROSCI.2770-10.2010

Eick, C.M., Kovács, G., Rostalski, S.M., Röhrig, L., Ambrus, G.G., 2020. The occipital face area is causally involved in identity-related visual-semantic associations. Brain Struct. Funct. 225, 1483–1493. https://doi.org/10.1007/s00429-020-02068-9

Epstein, R.A., Parker, W.E., Feiler, A.M., 2008. Two Kinds of fMRI Repetition Suppression? Evidence for Dissociable Neural Mechanisms. J. Neurophysiol. 99, 2877–2886. https://doi.org/10.1152/jn.90376.2008

Ewbank, M.P., Henson, R.N., Rowe, J.B., Stoyanova, R.S., Calder, A.J., 2013. Different neural mechanisms within occipitotemporal cortex underlie repetition suppression across same and different-size faces. Cereb. Cortex 23, 1073–1084. https://doi.org/10.1093/cercor/bhs070

Ewbank, M.P., Lawson, R.P., Henson, R.N., Rowe, J.B., Passamonti, L., Calder, A.J., 2011. Changes in “top-down” connectivity underlie repetition suppression in the ventral visual pathway. J. Neurosci. 31, 5635–5642. https://doi.org/10.1523/JNEUROSCI.5013-10.2011

Fabbrini, F., Van den Haute, C., De Vitis, M., Baekelandt, V., Vanduffel, W., Vogels, R., 2019. Probing the Mechanisms of Repetition Suppression in Inferior Temporal Cortex with Optogenetics. Curr. Biol. 29, 1988-1998.e4. https://doi.org/10.1016/j.cub.2019.05.014

Fan, X., Wang, F., Shao, H., Zhang, P., He, S., 2020. The bottom-up and top-down processing of faces in the human occipitotemporal cortex. Elife 9. https://doi.org/10.7554/eLife.48764

Farell, B., 1985. “Same”-"Different" Judgments. A Review of Current Controversies in Perceptual Comparisons. Psychol. Bull. https://doi.org/10.1037/0033-2909.98.3.419

Feldman, H., Friston, K.J., 2010. Attention, Uncertainty, and Free-Energy. Front. Hum. Neurosci. 4, 1–23. https://doi.org/10.3389/fnhum.2010.00215

Felleman, D.J., Van Essen, D.C., 1991. Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex 1, 1–47.

Feuerriegel, D., 2016. Selecting appropriate designs and comparison conditions in repetition paradigms. Cortex 80, 196–205. https://doi.org/10.1016/j.cortex.2015.10.022

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. https://doi.org/10.1016/j.neuroimage.2017.12.029

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. Biol. Psychol. 138, 110–125. https://doi.org/10.1016/j.biopsycho.2018.09.002

Feuerriegel, D., Vogels, R., Kovács, G., 2021a. Evaluating the evidence for expectation suppression in the visual system. Neurosci. Biobehav. Rev. 126, 368–381. https://doi.org/10.1016/j.neubiorev.2021.04.002

Feuerriegel, D., Yook, J., Quek, G.L., Hogendoorn, H., Bode, S., 2021b. Visual mismatch responses index surprise signalling but not expectation suppression. Cortex 134, 16–29. https://doi.org/10.1016/j.cortex.2020.10.006

Freeman, J., Simoncelli, E.P., 2011. Metamers of the ventral stream. Nat. Publ. Gr. 14. https://doi.org/10.1038/nn.2889

Friston, K., 2010. The free-energy principle: A unified brain theory? Nat. Rev. Neurosci. https://doi.org/10.1038/nrn2787

Friston, K., 2009. The free-energy principle: a rough guide to the brain? Trends Cogn. Sci. 13, 293–301. https://doi.org/10.1016/j.tics.2009.04.005

Friston, K., 2005. A theory of cortical responses. Philos. Trans. R. Soc. B Biol. Sci. 360, 815–836. https://doi.org/10.1098/rstb.2005.1622

Friston, K., 2003. Learning and inference in the brain. Neural Networks 16, 1325–1352. https://doi.org/10.1016/j.neunet.2003.06.005

Friston, K., 2002. Functional integration and inference in the brain. Prog. Neurobiol. https://doi.org/10.1016/S0301-0082(02)00076-X

Friston, K., Kiebel, S., 2009. Predictive coding under the free-energy principle. Philos. Trans. R. Soc. B Biol. Sci. 364, 1211–1221. https://doi.org/10.1098/rstb.2008.0300

Friston, K., Kilner, J., Harrison, L., 2006. A free energy principle for the brain. J. Physiol. Paris 100, 70–87. https://doi.org/10.1016/j.jphysparis.2006.10.001

Ganis, G., Schendan, H.E., 2008. Visual mental imagery and perception produce opposite adaptation effects on early brain potentials. Neuroimage 42, 1714–1727. https://doi.org/10.1016/j.neuroimage.2008.07.004

Garrido, M.I., Kilner, J.M., Kiebel, S.J., Stephan, K.E., Baldeweg, T., Friston, K.J., 2009a. Repetition suppression and plasticity in the human brain. Neuroimage 48, 269–279. https://doi.org/10.1016/j.neuroimage.2009.06.034

Garrido, M.I., Kilner, J.M., Stephan, K.E., Friston, K.J., 2009b. The mismatch negativity: A review of underlying mechanisms. Clin. Neurophysiol. 120, 453–463. https://doi.org/10.1016/j.clinph.2008.11.029

Gauthier, I., Tarr, M.J., Anderson, A.W., Skudlarski, P., Gore, J.C., 1999. Activation of the middle fusiform “face area” increases with expertise in recognizing novel objects. Nat. Neurosci. 2, 568–573. https://doi.org/10.1038/9224

Gauthier, I., Tarr, M.J., Moylan, J., Skudlarski, P., Gore, J.C., Anderson, a W., 2000. The fusiform “face area” is part of a network that processes faces at the individual level. J. Cogn. Neurosci. 12, 495–504. https://doi.org/10.1162/089892900562165

Gentile, F., Ales, J., Rossion, B., 2017. Being BOLD: The neural dynamics of face perception. Hum. Brain Mapp. 38, 120–139. https://doi.org/10.1002/hbm.23348

Gobbini, M.I., Haxby, J. V., 2007. Neural systems for recognition of familiar faces. Neuropsychologia 45, 32–41. https://doi.org/10.1016/j.neuropsychologia.2006.04.015

Gobbini, M.I., Haxby, J. V., 2006. Neural response to the visual familiarity of faces. Brain Res. Bull. 71, 76–82. https://doi.org/10.1016/j.brainresbull.2006.08.003

Goesaert, E., Op de Beeck, H.P., 2013. Representations of facial identity information in the ventral visual stream investigated with multivoxel pattern analyses. J. Neurosci. 33, 8549–8558. https://doi.org/10.1523/JNEUROSCI.1829-12.2013

Gorno-Tempini, M.L., Price, C.J., 2001. Identification of famous faces and buildings. A functional neuroimaging study of semantically unique items. Brain 124, 2087–2097. https://doi.org/10.1093/brain/124.10.2087

Gotts, S.J., Chow, C.C., Martin, A., 2012. Repetition priming and repetition suppression: A case for enhanced efficiency through neural synchronization. Cogn. Neurosci. 3, 227–237. https://doi.org/10.1080/17588928.2012.670617

Grill-Spector, K., Henson, R., Martin, A., 2006. Repetition and the brain: Neural models of stimulus-specific effects. Trends Cogn. Sci. 10, 14–23. https://doi.org/10.1016/j.tics.2005.11.006

Grill-Spector, K., Knouf, N., Kanwisher, N., 2004. The fusiform face area subserves face perception, not generic within-category identification. Nat. Neurosci. 7, 555–562. https://doi.org/10.1038/nn1224

Grill-Spector, K., Kourtzi, Z., Kanwisher, N., 2001. The lateral occipital complex and its role in object recognition. Vision Res. 41, 1409–1422. https://doi.org/10.1016/S0042-6989(01)00073-6

Grill-Spector, K., Malach, R., 2001. fMR-adaptation: A tool for studying the functional properties of human cortical neurons. Acta Psychol. (Amst). 107, 293–321. https://doi.org/10.1016/S0001-6918(01)00019-1

Grill-Spector, K., Weiner, K.S., Kay, K., Gomez, J., 2017. The Functional Neuroanatomy of Human Face Perception. Annu. Rev. Vis. Sci. 176, 167–196. https://doi.org/10.1146/annurev-vision-102016-061214

Gross, C. G. (1999). Brain, vision, memory : tales in the history of neuroscience 1. MIT Press paperback ed.). Cambridge, Mass.

Grotheer, M., Kovacs, G., 2014. Repetition Probability Effects Depend on Prior Experiences. J. Neurosci. 34, 6640–6646. https://doi.org/10.1523/JNEUROSCI.5326-13.2014

Grotheer, M., Kovács, G., 2016. Can predictive coding explain repetition suppression? Cortex 80, 113–124. https://doi.org/10.1016/j.cortex.2015.11.027

Grotheer, M., Kovács, G., 2015. The relationship between stimulus repetitions and fulfilled expectations. Neuropsychologia 67, 175–182. https://doi.org/10.1016/j.neuropsychologia.2014.12.017

Groves, P.M., Thompson, R.F., 1970. Habituation: A dual-process theory. Psychol. Rev. 77, 419–450. https://doi.org/10.1037/h0029810

Gschwind, M., Pourtois, G., Schwartz, S., Van De Ville, D., Vuilleumier, P., 2012. White-matter connectivity between face-responsive regions in the human brain. Cereb. Cortex 22, 1564–1576. https://doi.org/10.1093/cercor/bhr226

Harry, B.B., Umla-Runge, K., Lawrence, A.D., Graham, K.S., Downing, P.E., 2016. Evidence for integrated visual face and body representations in the anterior temporal lobes. J. Cogn. Neurosci. 28, 1178–1193. https://doi.org/10.1162/jocn_a_00966

Haxby, J. V., Hoffman, E.A., Gobbini, M.I., 2000. The distributed human neural system for face perception. Trends Cogn. Sci. 4, 223–233. https://doi.org/10.1016/S1364-6613(00)01482-0

Heilbron, M., Chait, M., 2018. Great Expectations: Is there Evidence for Predictive Coding in Auditory Cortex? Neuroscience. https://doi.org/10.1016/j.neuroscience.2017.07.061

Helmholtz, H. (1867). Handbuch der physiologischen Optik. Leipzig: Voss.

Henderson, J.M., Choi, W., Lowder, M.W., Ferreira, F., 2016. Language structure in the brain: A fixation-related fMRI study of syntactic surprisal in reading. Neuroimage 132, 293–300. https://doi.org/10.1016/j.neuroimage.2016.02.050

Henry, C.A., Joshi, S., Xing, D., Shapley, R.M., Hawken, M.J., 2013. Functional characterization of the extraclassical receptive field in macaque V1: Contrast, orientation, and temporal dynamics. J. Neurosci. 33, 6230–6242. https://doi.org/10.1523/JNEUROSCI.4155-12.2013

Henson, R., Shallice, T., Dolan, R., 2000. Neuroimaging evidence for dissociable forms of repetition priming. Science (80-.). 287, 1269–1272. https://doi.org/10.1126/science.287.5456.1269

Henson, R.N., 2016. Repetition suppression to faces in the fusiform face area: A personal and dynamic journey. Cortex 80, 174–184. https://doi.org/10.1016/j.cortex.2015.09.012

Henson, R.N.A., 2003. Neuroimaging studies of priming. Prog. Neurobiol. 70, 53–81. https://doi.org/10.1016/S0301-0082(03)00086-8

Henson, R.N.A., Price, C.J., Rugg, M.D., Turner, R., Friston, K.J., 2002. Detecting latency differences in event-related BOLD responses: Application to words versus nonwords and initial versus repeated face presentations. Neuroimage 15, 83–97. https://doi.org/10.1006/nimg.2001.0940

Henson, R.N.A., Rugg, M.D., 2003. Neural response suppression, haemodynamic repetition effects, and behavioural priming. Neuropsychologia 41, 263–270. https://doi.org/10.1016/S0028-3932(02)00159-8

Hermann, P., Banko, E.M., Gal, V., Vidnyanszky, Z., 2015. Neural Basis of Identity Information Extraction from Noisy Face Images. J. Neurosci. 35, 7165–7173. https://doi.org/10.1523/JNEUROSCI.3572-14.2015

Hermann, P., Grotheer, M., Kovács, G., Vidnyánszky, Z.Z., 2017. The relationship between repetition suppression and face perception. Brain Imaging Behav. 11, 1018–1028. https://doi.org/10.1007/s11682-016-9575-9

Hoffman, E.A., Haxby, J. V., 2000. Distinct representations of eye gaze and identity in the distributed human neural system for face perception. Nat. Neurosci. 3, 80–84. https://doi.org/10.1038/71152

Hohwy, J., Roepstorff, A., Friston, K., 2008. Predictive coding explains binocular rivalry: An epistemological review. Cognition 108, 687–701. https://doi.org/10.1016/j.cognition.2008.05.010

Hole, G., 2011. Identity-specific face adaptation effects: Evidence for abstractive face representations. Cognition 119, 216–228. https://doi.org/10.1016/j.cognition.2011.01.011

Hosoya, T., Baccus, S.A., Meister, M., 2005. Dynamic predictive coding by the retina. Nature 436, 71–77. https://doi.org/10.1038/nature03689

Ishai, A., Schmidt, C.F., Boesiger, P., 2005. Face perception is mediated by a distributed cortical network. Brain Res. Bull. 67, 87–93. https://doi.org/10.1016/j.brainresbull.2005.05.027

James, T.W., Gauthier, I., 2006. Repetition-induced changes in BOLD response reflect accumulation of neural activity. Hum. Brain Mapp. 27, 37–46. https://doi.org/10.1002/hbm.20165

Jemel, B., Pisani, M., Rousselle, L., Crommelinck, M., Bruyer, R., 2005. Exploring the functional architecture of person recognition system with event-related potentials in a within- and cross-domain self-priming of faces. Neuropsychologia 43, 2024–2040. https://doi.org/10.1016/j.neuropsychologia.2005.03.016

Jenkins, R., Burton, M., 2011. Stable face representations. Philos. Trans. R. Soc. B Biol. Sci. 366, 1671–1683. https://doi.org/10.1098/rstb.2010.0379

Jenkins, R., White, D., Van Montfort, X., Mike Burton, A., 2011. Variability in photos of the same face. Cognition 121, 313–323. https://doi.org/10.1016/j.cognition.2011.08.001

Jiang, J., Summerfield, C., Egner, T., 2013. Attention sharpens the distinction between expected and unexpected percepts in the visual brain. J. Neurosci. 33, 18438–18447. https://doi.org/10.1523/JNEUROSCI.3308-13.2013

Johnston, P., Overell, A., Kaufman, J., Robinson, J., Young, A.W., 2016. Expectations about person identity modulate the face-sensitive N170. Cortex 85, 54–64. https://doi.org/10.1016/j.cortex.2016.10.002

Kaiser, D., Walther, C., Schweinberger, S.R., Kovacs, G., 2013. Dissociating the neural bases of repetition-priming and adaptation in the human brain for faces. J. Neurophysiol. 110. https://doi.org/10.1152/jn.00277.2013

Kaliukhovich, D.A., de Baene, W., Vogels, R., 2013. Effect of adaptation on object representation accuracy in macaque inferior temporal cortex. J. Cogn. Neurosci. 25, 777–789. https://doi.org/10.1162/jocn_a_00355

Kaliukhovich, D.A., Vogels, R., 2014. Neurons in Macaque Inferior Temporal Cortex Show No Surprise Response to Deviants in Visual Oddball Sequences. J. Neurosci. 34, 12801–12815. https://doi.org/10.1523/jneurosci.2154-14.2014

Kaliukhovich, D.A., Vogels, R., 2011. Stimulus repetition probability does not affect repetition suppression in macaque inferior temporal cortex. Cereb. Cortex 21, 1547–1558. https://doi.org/10.1093/cercor/bhq207

Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. 17, 4302–11. https://doi.org/10.1098/Rstb.2006.1934

Kanwisher, N., Yovel, G., 2006. The fusiform face area: a cortical region specialized for the perception of faces. Philos. Trans. R. Soc. B Biol. Sci. 361, 2109–2128. https://doi.org/10.1098/rstb.2006.1934

Kar, K., Krekelberg, B., 2016. Testing the assumptions underlying fMRI adaptation using intracortical recordings in area MT. Cortex 80, 21–34. https://doi.org/10.1016/j.cortex.2015.12.011

Kersten, D., Mamassian, P., Yuille, A., 2004. Object perception as Bayesian inference. Annu. Rev. Psychol. 55, 271–304. https://doi.org/10.1146/annurev.psych.55.090902.142005

Koelsch, S., Vuust, P., Friston, K., 2018. Predictive Processes and the Peculiar Case of Music. Trends Cogn. Sci. xx, 63–77. https://doi.org/10.1016/j.tics.2018.10.006

Kohn, A., 2007. Visual adaptation: Physiology, mechanisms, and functional benefits. J. Neurophysiol. 97, 3155–3164. https://doi.org/10.1152/jn.00086.2007

Kohn, A., Movshon, J.A., 2003. Neuronal adaptation to visual motion in area MT of the macaque. Neuron 39, 681–691. https://doi.org/10.1016/S0896-6273(03)00438-0

Kok, P., Bains, L.J., Van Mourik, T., Norris, D.G., De Lange, F.P., 2016. Selective activation of the deep layers of the human primary visual cortex by top-down feedback. Curr. Biol. 26, 371–376. https://doi.org/10.1016/j.cub.2015.12.038

Kok, P., Jehee, J.F.M., de Lange, F.P., 2012a. Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex. Neuron 75, 265–270. https://doi.org/10.1016/j.neuron.2012.04.034

Kok, P., Rahnev, D., Jehee, J.F.M.M., Lau, H.C., de Lange, F.P., 2012b. Attention reverses the effect of prediction in silencing sensory signals. Cereb. Cortex 22, 2197–2206. https://doi.org/10.1093/cercor/bhr310

Kok, R., Taubert, J., Van der Burg, E., Rhodes, G., Alais, D., 2017. Face familiarity promotes stable identity recognition: Exploring face perception using serial dependence. R. Soc. Open Sci. 4. https://doi.org/10.1098/rsos.160685

Köster, M., Kayhan, E., Langeloh, M., Hoehl, S., 2020. Making Sense of the World: Infant Learning From a Predictive Processing Perspective. Perspect. Psychol. Sci. 15, 562–571. https://doi.org/10.1177/1745691619895071

Kouider, S., Eger, E., Dolan, R., Henson, R.N., 2009. Activity in face-responsive brain regions is modulated by invisible, attended faces: Evidence from masked priming. Cereb. Cortex 19, 13–23. https://doi.org/10.1093/cercor/bhn048

Kourtzi, Z., Kanwisher, N., 2001. Representation of Perceived Object Shape by the Human Lateral Occipital Complex. Science (80-.). 293.

Kovács, G., 2020. Getting to know someone: Familiarity, person recognition, and identification in the human brain. J. Cogn. Neurosci. 32, 2205–2225. https://doi.org/10.1162/jocn_a_01627

Kovács, G., Iffland, L., Vidnyánszky, Z., Greenlee, M.W., 2012. Stimulus repetition probability effects on repetition suppression are position invariant for faces. Neuroimage 60, 2128–2135. https://doi.org/10.1016/j.neuroimage.2012.02.038

Kovacs, G., Kaiser, D., Kaliukhovich, D.A., Vidnyanszky, Z., Vogels, R., 2013. Repetition Probability Does Not Affect fMRI Repetition Suppression for Objects. J. Neurosci. 33, 9805–9812. https://doi.org/10.1523/JNEUROSCI.3423-12.2013

Kovács, G., Vogels, R., 2014. When does repetition suppression depend on repetition probability ? Front. Hum. Neurosci. 8, 685. https://doi.org/10.3389/fnhum.2014.00685

Kravitz, D.J., Saleem, K.S., Baker, C.I., Ungerleider, L.G., Mishkin, M., 2013. The ventral visual pathway: An expanded neural framework for the processing of object quality. Trends Cogn. Sci. https://doi.org/10.1016/j.tics.2012.10.011

Kriegeskorte, N., Formisano, E., Sorger, B., Goebel, R., 2007. Individual faces elicit distinct response patterns in human anterior temporal cortex. Proc. Natl. Acad. Sci. 104, 20600–20605. https://doi.org/10.1073/pnas.0705654104

Lamme, V.A.F., Supèr, H., Spekreijse, H., 1998. Feedforward, horizontal, and feedback processing in the visual cortex. Curr. Opin. Neurobiol. 8, 529–535. https://doi.org/10.1016/S0959-4388(98)80042-1

Larsson, J., Smith, A.T., 2012. fMRI Repetition Suppression : Neuronal Adaptation or Stimulus Expectation ? Cereb. Cortex 22, 567–576. https://doi.org/10.1093/cercor/bhr119

Larsson, J., Solomon, S.G., Kohn, A., 2016. fMRI adaptation revisited. Cortex 80, 154–160. https://doi.org/10.1016/j.cortex.2015.10.026

Lawrence, S.J.D., Formisano, E., Muckli, L., de Lange, F.P., 2019. Laminar fMRI: Applications for cognitive neuroscience. Neuroimage. https://doi.org/10.1016/j.neuroimage.2017.07.004

Lee, T.S., Mumford, D., 2003. Hierarchical Bayesian inference in the visual cortex. J. Opt. Soc. Am. A 20, 1434. https://doi.org/10.1364/josaa.20.001434

Lehar, S. (2003). *The world in your head: A Gestalt view of the mechanism of conscious experience.* Lawrence Erlbaum Associates Publishers.

Li, L., Miller, E.K., Desimone, R., 1993. The representation of stimulus familiarity in anterior inferior temporal cortex. J. Neurophysiol. 69, 1918–1929. https://doi.org/10.1152/jn.1993.69.6.1918

Loffler, G., Yourganov, G., Wilkinson, F., Wilson, H.R., 2005. fMRI evidence for the neural representation of faces. Nat. Neurosci. 8, 1386–1390. https://doi.org/10.1038/nn1538

Maccotta, L., Buckner, R.L., 2004. Evidence for neural effects of repetition that directly correlate with behavioral priming. J. Cogn. Neurosci. 16, 1625–1632. https://doi.org/10.1162/0898929042568451

Majid, A., Roberts, S.G., Cilissen, L., Emmorey, K., Nicodemus, B., O'Grady, L., Woll, B., LeLan, B., De Sousa, H., Cansler, B.L., Shayan, S., De Vos, C., Senft, G., Enfield, N.J., Razak, R.A., Fedden, S., Tufvesson, S., Dingemanse, M., Ozturk, O., Brown, P., Hill, C., Le Guen, O., Hirtzel, V., Van Gijn, R., Sicoli, M.A., Levinson, S.C., 2018. Differential coding of perception in the world's languages. Proc. Natl. Acad. Sci. U. S. A. 115, 11369–11376. https://doi.org/10.1073/pnas.1720419115

Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R., Tootell, R.B., 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc. Natl. Acad. Sci. 92, 8135–8139. https://doi.org/10.1073/pnas.92.18.8135

Man, K., Kaplan, J., Damasio, H., Damasio, A., 2013. Neural convergence and divergence in the mammalian cerebral cortex: From experimental neuroanatomy to functional neuroimaging. J. Comp. Neurol. https://doi.org/10.1002/cne.23408

Maunsell, J.H.R., Cook, E.P., 2002. The role of attention in visual processing, in: Philosophical Transactions of the Royal Society B: Biological Sciences. pp. 1063–1072. https://doi.org/10.1098/rstb.2002.1107

Maunsell, J.H.R., Van Essen, D.C., 1983. The connections of the middle temporal visual area

(MT) and their relationship to a cortical hierarchy in the macaque monkey. J. Neurosci. 3, 2563–2586. https://doi.org/10.1523/jneurosci.03-12-02563.1983

Maurer, D., Le Grand, R., Mondloch, C.J., 2002. The many faces of configural processing. Trends Cogn. Sci. https://doi.org/10.1016/S1364-6613(02)01903-4

May, P.J.C., Tiitinen, H., 2010. Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. Psychophysiology 47, 66–122. https://doi.org/10.1111/j.1469-8986.2009.00856.x

McGugin, R.W., Newton, A.T., Tamber-Rosenau, B., Tomarken, A., Gauthier, I., 2019. Thickness of deep layers in the fusiform face area predicts face recognition. J. Cogn. Neurosci. 32, 1316–1329. https://doi.org/10.1162/jocn_a_01551

Meyer, T., Olson, C.R., 2011. Statistical learning of visual transitions in monkey inferotemporal cortex. Proc. Natl. Acad. Sci. 108, 19401–19406. https://doi.org/10.1073/pnas.1112895108

Miller, E.K., Desimone, R., 1994. Parallel neuronal mechanisms for short-term memory. Science (80-.). 263, 520–522. https://doi.org/10.1126/science.8290960

Miller, E.K., Li, L., Desimone, R., 1991. A neural mechanism for working and recognition memory in inferior temporal cortex. Science (80-.). 254, 1377–1379. https://doi.org/10.1126/science.1962197

Mooney, C.M., 1957. Age in the development of closure ability in children. Can. J. Psychol. 11, 219–226. https://doi.org/10.1037/h0083717

Moran, J., Desimone, R., 1985. Selective attention gates visual processing in the extrastriate cortex. Science (80-.). 229, 782–784. https://doi.org/10.1126/science.4023713

Muckli, L., De Martino, F., Vizioli, L., Petro, L.S., Smith, F.W., Ugurbil, K., Goebel, R., Yacoub, E., 2015. Contextual Feedback to Superficial Layers of V1. Curr. Biol. 25, 2690–2695. https://doi.org/10.1016/j.cub.2015.08.057

Mumford, D., 1992. On the computational architecture of the neocortex - II The role of cortico-cortical loops. Biol. Cybern. 66, 241–251. https://doi.org/10.1007/BF00198477

Musmann, H. (1979). Predictive image coding. Image transmission techniques, *73*, 112.

Näätänen, R. (1992). Attention and Brain Function (1st ed.). Routledge. https://doi.org/10.4324/9780429487354

Nagy, K., Greenlee, M.W., Kovács, G., 2012. The lateral occipital cortex in the face perception network: An effective connectivity study. Front. Psychol. 3. https://doi.org/10.3389/fpsyg.2012.00141

Natu, V., O'Toole, A.J., 2011. The neural processing of familiar and unfamiliar faces: A review and synopsis. Br. J. Psychol. 102, 726–747. https://doi.org/10.1111/j.2044-8295.2011.02053.x

Nemrodov, D., Niemeier, M., Patel, A., Nestor, A., 2018. The neural dynamics of facial identity processing: Insights from EEG-based pattern analysis and image reconstruction. eNeuro 5. https://doi.org/10.1523/ENEURO.0358-17.2018

Nestor, A., Plaut, D.C., Behrmann, M., 2011. Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. Proc. Natl. Acad. Sci. 108, 9998–10003. https://doi.org/10.1073/pnas.1102433108

Neumann, M.F., Schweinberger, S.R., 2008. N250r and N400 ERP correlates of immediate famous face repetition are independent of perceptual load. Brain Res. 1239, 181–190. https://doi.org/10.1016/j.brainres.2008.08.039

O'Reilly, J.X., Jbabdi, S., Behrens, T.E.J., 2012. How can a Bayesian approach inform neuroscience? Eur. J. Neurosci. 35, 1169–1179. https://doi.org/10.1111/j.1460-9568.2012.08010.x

Olkkonen, M., Aguirre, G.K., Epstein, R.A., 2017. Expectation modulates repetition priming under high stimulus variability 17, 1–16. https://doi.org/10.1167/17.6.10.doi

Pajani, A., Kouider, S., Roux, P., de Gardelle, V., 2017. Unsuppressible Repetition Suppression and exemplar-specific Expectation Suppression in the Fusiform Face Area. Sci. Rep. 7, 160. https://doi.org/10.1038/s41598-017-00243-3

Pazo-Alvarez, P., Cadaveira, F., Amenedo, E., 2003. MMN in the visual modality: a review. Biol. Psychol. 63, 199–236.

Phillips, C.G., Zeki, S., Barlow, H.B., 1984. Localization of function in the cerebral cortex: Past, present and future. Brain 107, 328–361. https://doi.org/10.1093/brain/107.1.328

Pitcher, D., Charles, L., Devlin, J.T., Walsh, V., Duchaine, B., 2009. Triple Dissociation of Faces, Bodies, and Objects in Extrastriate Cortex. Curr. Biol. 19, 319–324. https://doi.org/10.1016/j.cub.2009.01.007

Pitcher, D., Garrido, L., Walsh, V., Duchaine, B.C., 2008. Transcranial magnetic stimulation disrupts the perception and embodiment of facial expressions. J. Neurosci. 28, 8929–8933. https://doi.org/10.1523/JNEUROSCI.1450-08.2008

Pitcher, D., Ungerleider, L.G., 2021. Evidence for a Third Visual Pathway Specialized for Social Perception. Trends Cogn. Sci. 25, 100–110. https://doi.org/10.1016/j.tics.2020.11.006

Pitcher, D., Walsh, V., Yovel, G., Duchaine, B., 2007. TMS Evidence for the Involvement of the Right Occipital Face Area in Early Face Processing. Curr. Biol. 17, 1568–1573. https://doi.org/10.1016/j.cub.2007.07.063

Pobric, G., Jefferies, E., Ralph, M.A.L., 2007. Anterior temporal lobes mediate semantic representation: Mimicking semantic dementia by using rTMS in normal participants, PNAS.

Posner, M.I., Snyder, C.R., Davidson, B.J., 1980. Attention and the detection of signals. J. Exp. Psychol. Gen. 109, 160–174. https://doi.org/10.1037/0096-3445.109.2.160

Pyles, J.A., Verstynen, T.D., Schneider, W., Tarr, M.J., 2013. Explicating the Face Perception Network with White Matter Connectivity. PLoS One 8, 61611. https://doi.org/10.1371/journal.pone.0061611

Ramon, M., Gobbini, M.I., 2018. Familiarity matters: A review on prioritized processing of personally familiar faces. Vis. cogn. 26, 179–195. https://doi.org/10.1080/13506285.2017.1405134

Rangarajan, V., Jacques, C., Knight, R.T., Weiner, K.S., Grill-Spector, K., 2020. Diverse Temporal Dynamics of Repetition Suppression Revealed by Intracranial Recordings in the Human Ventral Temporal Cortex. https://doi.org/10.1093/cercor/bhaa173

Rao, R.P.N., 2005. Bayesian inference and attentional modulation in the visual cortex. Neuroreport 16, 1843–1848. https://doi.org/10.1097/01.wnr.0000183900.92901.fc

Rao, R.P.N., Ballard, D.H., 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. 2, 79–87. https://doi.org/10.1038/4580

Rice, G.E., Caswell, H., Moore, P., Hoffman, P., Lambon Ralph, M.A., 2018. The Roles of Left Versus Right Anterior Temporal Lobes in Semantic Memory: A Neuropsychological Comparison of Postsurgical Temporal Lobe Epilepsy Patients. Cereb. Cortex 1–15. https://doi.org/10.1093/cercor/bhx362

Robinson, J.E., Breakspear, M., Young, A.W., Johnston, P.J., 2018. Dose-dependent modulation of the visually evoked N1/N170 by perceptual surprise: A clear demonstration of prediction-error signalling. Eur. J. Neurosci. 1–11. https://doi.org/10.1111/ejn.13920

Robinson, J.E., Woods, W., Leung, S., Kaufman, J., Breakspear, M., Young, A.W., Johnston, P.J., 2020. Prediction-error signals to violated expectations about person identity and head orientation are doubly-dissociated across dorsal and ventral visual stream regions. Neuroimage 206, 116325. https://doi.org/10.1016/j.neuroimage.2019.116325

Rossion, B., Caldara, R., Seghier, M., Schuller, A.M., Lazeyras, F., Mayer, E., 2003. A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. Brain 126, 2381–2395. https://doi.org/10.1093/brain/awg241

Rotshtein, P., Henson, R.N.A., Treves, A., Driver, J., Dolan, R.J., 2005. Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. Nat. Neurosci. 8, 107–113. https://doi.org/10.1038/nn1370

Ryu, J.J., Chaudhuri, A., 2006. Representations of familiar and unfamiliar faces as revealed by viewpoint-aftereffects. Vision Res. 46, 4059–4063. https://doi.org/10.1016/j.visres.2006.07.018

Sadeh, B., Podlipsky, I., Zhdanov, A., Yovel, G., 2010. Event-related potential and functional MRI measures of face-selectivity are highly correlated: A simultaneous ERP-fMRI investigation. Hum. Brain Mapp. 31, 1490–1501. https://doi.org/10.1002/hbm.20952

Sayres, R., Grill-Spector, K., 2006. Object-Selective Cortex Exhibits Performance-Independent Repetition Suppression. J. Neurophysiol. 95, 995–1007. https://doi.org/10.1152/jn.00500.2005

Schacter, D.L., Buckner, R.L., 1998. Priming and the Brain Review. Neuron 20, 185–195.

Schwartenbeck, P., FitzGerald, T., Dolan, R.J., Friston, K., 2013. Exploration, novelty, surprise, and free energy minimization. Front. Psychol. 4. https://doi.org/10.3389/fpsyg.2013.00710

Schwarz, L., Kreifelts, B., Wildgruber, D., Erb, M., Scheffler, K., Ethofer, T., 2019. Properties of face localizer activations and their application in functional magnetic resonance imaging (fMRI) fingerprinting. PLoS One 14, 1–21. https://doi.org/10.1371/journal.pone.0214997

Schweinberger, S.R., 1996. How gorbachev primed yeltsin: Analyses of associative priming in person recognition by means of reaction times and event-related brain potentials. J. Exp. Psychol. Learn. Mem. Cogn. 22, 1383–1407. https://doi.org/10.1037/0278-7393.22.6.1383

Schweinberger, S.R., Burton, A.M., 2003. Covert recognition and the neural system for face processing. Cortex. https://doi.org/10.1016/S0010-9452(08)70071-6

Schweinberger, S.R., Neumann, M.F., 2016. Repetition effects in human ERPs to faces. Cortex 80, 141–153. https://doi.org/10.1016/j.cortex.2015.11.001

Schweinberger, S.R., Pfütze, E.M., Sommer, W., 1995. Repetition Priming and Associative Priming of Face Recognition: Evidence From Event-Related Potentials. J. Exp. Psychol. Learn. Mem. Cogn. 21, 722–736. https://doi.org/10.1037/0278-7393.21.3.722

Schweinberger, S.R., Pickering, E.C., Jentzsch, I., Burton, A.M., Kaufmann, J.M., 2002. Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions, Cognitive Brain Research.

Schwiedrzik, C.M., Freiwald, W.A., 2017. High-Level Prediction Signals in a Low-Level Area of the Macaque Face-Processing Hierarchy. Neuron 96, 89-97.e4. https://doi.org/10.1016/j.neuron.2017.09.007

Siman-Tov, T., Granot, R.Y., Shany, O., Singer, N., Hendler, T., Gordon, C.R., 2019. Is there a prediction network? Meta-analytic evidence for a cortical-subcortical network likely subserving prediction, Neuroscience & Biobehavioral Reviews. Pergamon. https://doi.org/10.1016/J.NEUBIOREV.2019.08.012

Sobotka, S., Ringo, J.L., 1994. Stimulus specific adaptation in excited but not in inhibited cells in inferotemporal cortex of Macaque. Brain Res. 646, 95–99. https://doi.org/10.1016/0006-8993(94)90061-2

Spratling, M.W., 2017. A review of predictive coding algorithms. Brain Cogn. 112, 92–97. https://doi.org/10.1016/J.BANDC.2015.11.003

Sterzer, P., Adams, R.A., Fletcher, P., Frith, C., Lawrie, S.M., Muckli, L., Petrovic, P., Uhlhaas, P., Voss, M., Corlett, P.R., 2018. The Predictive Coding Account of Psychosis. Biol. Psychiatry 84, 634–643. https://doi.org/10.1016/j.biopsych.2018.05.015

Summerfield, C., Koechlin, E., 2008. A Neural Representation of Prior Information during Perceptual Inference. Neuron 59, 336–347. https://doi.org/10.1016/j.neuron.2008.05.021

Summerfield, C., Trittschuh, E.H., Monti, J.M., Mesulam, M.-M., Egner, T., 2008. Neural repetition suppression reflects fulfilled perceptual expectations. Nat. Neurosci. 11, 1004–1006. https://doi.org/10.1038/nn.2163

Summerfield, C., Wyart, V., Johnen, V.M., de Gardelle, V., 2011. Human scalp electroencephalography reveals that repetiti suppression varies with expectation. Front. Hum. Neurosci. 5, 1–13. https://doi.org/10.3389/fnhum.2011.00067

Todorovic, A., de Lange, F.P., 2012. Repetition Suppression and Expectation Suppression Are Dissociable in Time in Early Auditory Evoked Fields. J. Neurosci. 32, 13389–13395. https://doi.org/10.1523/JNEUROSCI.2227-12.2012

Todorovic, A., van Ede, F., Maris, E., de Lange, F.P., 2011. Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: An MEG study. J. Neurosci. 31, 9118–9123. https://doi.org/10.1523/JNEUROSCI.1425-11.2011

Trainor, L.J., 2012. Predictive information processing is a fundamental learning mechanism present in early development: Evidence from infants. Int. J. Psychophysiol. https://doi.org/10.1016/j.ijpsycho.2011.12.008

Trapp, S., Lepsien, J., Kotz, S.A., Bar, M., 2016. Prior probability modulates anticipatory activity in category-specific areas. Cogn. Affect. Behav. Neurosci. 16, 135–144. https://doi.org/10.3758/s13415-015-0373-4

Trapp, S., Schweinberger, S.R., Hayward, W.G., Kovács, G., 2018. Integrating predictive frameworks and cognitive models of face perception. Psychon. Bull. Rev. 25. https://doi.org/10.3758/s13423-018-1433-x

Tsantani, M., Kriegeskorte, N., McGettigan, C., Garrido, L., 2019. Faces and voices in the brain: A modality-general person-identity representation in superior temporal sulcus. Neuroimage 201, 116004. https://doi.org/10.1016/J.NEUROIMAGE.2019.07.017

Tsantani, M., Kriegeskorte, N., Storrs, K., Williams, A.L., McGettigan, C., Garrido, L., 2021. FFA and OFA encode distinct types of face identity information. J. Neurosci. 41, JN-RM-1449-20. https://doi.org/10.1523/jneurosci.1449-20.2020

Turk-Browne, N.B., Yi, D.J., Leber, A.B., Chun, M.M., 2007. Visual quality determines the direction of neural repetition effects. Cereb. Cortex 17, 425–433. https://doi.org/10.1093/cercor/bhj159

Ungerleider, L.G. and Mishkin, M. (1982) Two cortical visual systems. In Analysis of Visual Behavior (Ingle, D.J. et al., eds), pp. 549–586, MIT Press

Verosky, S.C., Zoner, K.A., Marble, C.W., Sammon, M.M., Babarinsa, C.O., 2020. Familiarization increases face individuation measured with fast periodic visual stimulation. Biol. Psychol. 153, 107883. https://doi.org/10.1016/j.biopsycho.2020.107883

Vinken, K., Op de Beeck, H.P., Vogels, R., 2018. Face repetition probability does not affect repetition suppression in Macaque inferotemporal cortex. J. Neurosci. 38, 7492–7504. https://doi.org/10.1523/JNEUROSCI.0462-18.2018

Vogels, R., 2016. Sources of adaptation of inferior temporal cortical responses. Cortex 80, 185–195. https://doi.org/10.1016/j.cortex.2015.08.024

Von Der Heide, R.J., Skipper, L.M., Olson, I.R., 2013. Anterior temporal face patches: a meta-analysis and empirical study. Front. Hum. Neurosci. 7, 1–18. https://doi.org/10.3389/fnhum.2013.00017

Wacongne, C., Labyt, E., Van Wassenhove, V., Bekinschtein, T., Naccache, L., Dehaene, S., 2011. Evidence for a hierarchy of predictions and prediction errors in human cortex. Proc. Natl. Acad. Sci. U. S. A. 108, 20754–20759. https://doi.org/10.1073/pnas.1117807108

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. Ann. N. Y. Acad. Sci. https://doi.org/10.1111/nyas.14321

Walther, C., Schweinberger, S.R., Kovács, G., 2013. Adaptor Identity Modulates Adaptation Effects in Familiar Face Identification and Their Neural Correlates. PLoS One 8, 70525. https://doi.org/10.1371/journal.pone.0070525

Weibert, K., Harris, R.J., Mitchell, A., Byrne, H., Young, A.W., Andrews, T.J., 2016. An image-invariant neural response to familiar faces in the human medial temporal lobe. Cortex 84, 34–42. https://doi.org/10.1016/j.cortex.2016.08.014

Weiner, K.S., Grill-Spector, K., 2012. The improbable simplicity of the fusiform face area. Trends Cogn. Sci. https://doi.org/10.1016/j.tics.2012.03.003

Weiner, K.S., Sayres, R., Vinberg, J., Grill-Spector, K., 2010. fMRI-Adaptation and Category Selectivity in Human Ventral Temporal Cortex: Regional Differences Across Time Scales. J. Neurophysiol. 103, 3349–3365. https://doi.org/10.1152/jn.01108.2009

Westerberg, J.A., Cox, M.A., Dougherty, K., Maier, A., 2019. V1 microcircuit dynamics: Altered signal propagation suggests intracortical origins for adaptation in response to visual repetition. J. Neurophysiol. 121, 1938–1952. https://doi.org/10.1152/jn.00113.2019

Whitmire, C.J., Stanley, G.B., 2016. Rapid Sensory Adaptation Redux: A Circuit Perspective. Neuron. https://doi.org/10.1016/j.neuron.2016.09.046

Wig, G.S., Grafton, S.T., Demos, K.E., Kelley, W.M., 2005. Reductions in neural activity underlie behavioral components of repetition priming. https://doi.org/10.1038/nn1515

Wiggs, C.L., Martin, A., 1998. Properties and mechanisms of perceptual priming. Curr. Opin. Neurobiol. 8, 227–233. https://doi.org/10.1016/S0959-4388(98)80144-X

Winkler, I., Czigler, I., 1998. Mismatch negativity: Deviance detection or the maintenance of the “standard.” Neuroreport 9, 3809–3813. https://doi.org/10.1097/00001756-199812010-00008

Xu, J., Moeller, S., Auerbach, E.J., Strupp, J., Smith, S.M., Feinberg, D.A., Yacoub, E., Uğurbil, K., 2013. Evaluation of slice accelerations using multiband echo planar imaging at 3T. Neuroimage 83, 991–1001. https://doi.org/10.1016/j.neuroimage.2013.07.055

Xu, X., Yue, X., Lescroart, M.D., Biederman, I., Kim, J.G., 2009. Adaptation in the fusiform face area (FFA): Image or person? Vision Res. 49, 2800–2807. https://doi.org/10.1016/j.visres.2009.08.021

Zeki, S., Shipp, S., 1988. The functional logic of cortical connections. Nature. https://doi.org/10.1038/335311a0

Zhang, J., Li, X., Song, Y., Liu, J., 2012. The fusiform face area is engaged in holistic, not parts-based, representation of faces. PLoS One 7. https://doi.org/10.1371/journal.pone.0040390

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

A

AB	*context type (StudyII)*
ADA	*adaptation condition (Study III)*
aFFA	*anterior fusiform face area*
AI	*anterior insula*
Alt	*alternation*
ALT	*alternation condition (Study III)*
ALT-II	*alternation with final repetition condition (Study III)*
ANOVA	*analysis of variance*
aSTS-FA	*anterior superior temporal sulcus face area*
ATL(-FA)	*anterior temporal lobe (face area)*
AX	*context type (Study II)*

B

BA	*Brodmann area*
BOLD	*blood oxygen level dependent*

C

CBF	*cerebral blood flow*

D

DCM	*dynamic causal modeling*
DICOM	*digital imaging and communications in medicine format*

E

ECoG	*electrocorticography*
EEG	*electroencephalography*
ERP	*event-related potential*
ES	*expectation suppression*
EV	*expectation violation condition (Study III)*
EVC	*early visual cortex*

F

FC	*frontal cortex*
FFA (l/r)	*(left/right) fusiform face area*
fMR-a	*functional magnetic resonance adaptation*
fMRI	*functional magnetic resonance imaging*
FWE	*family-wise error*
FWHM	*full-width at half-maximum*

G

GLM	*general linear model*

H

HRF	*hemodynamic response function*

I

ID	*identity*
iFFA	*inferior frontal face area*
IFG(-FA)	*inferior frontal gyrus (face area)*
ISI	*inter-stimulus interval*
ITI	*inter-trial interval*

L

LGN	*lateral geniculate nucleus*
LO (l/r)	*(left/right) lateral occipital cortex*
LOC	*lateral occipital complex*

M

M	*mean*
MEG	*magnetoencephalography*
mFFA	
MFG	*middle frontal gyrus*
MMN	*mismatch negativity*
MMR	*mismatch response*
MNI	*Montreal Neurological Institute*
MP-RAGE	*magnetization prepared rapid gradient echo*
MR(I)	*magnetic resonance imaging*
MST	*medial superior temporal area*
MT	*middle temporal visual area*
MTL	*middle temporal lobe*
MVPA	*multivariate pattern analysis*

O

OFA (l/r)	*(left/right) occipital face area*
OSF	*open science framework*

P

PC	*predictive coding*
PDF	*probability density function*
pFFA	*posterior fusiform face area*
PMC	*premotor cortex*
PPA	*parahippocampal place area*
pre-SMA	*pre-supplementary motor area*
P_{REP}	*repetition probability*
pSTS-FA	*posterior superior temporal sulcus face area*

R

Rep	*repetition*
ROI	*region of interest*
RS	*repetition suppression*
RSA	*representational similarity analysis*
RT	*reaction time*

S

S1	*first stimulus*
S2	*second stimulus*
SD	*standard deviation*
SE	*standard error*
STS	*superior temporal sulcus*

T

TE	*time of echo*
TMS	*transcranial magnetic stimulation*
TPJ	*temporoparietal junction*
TR	*repetition time*

V

V1	*primary visual cortex*
V2	*secondary visual cortex*
V3	*visual area V3 in the third visual cortex*
vMMN	*visual mismatch negativity*

AUTHOR CONTRIBUTIONS

Study I. Repetition suppression for noisy and intact faces in the occipito-temporal cortex.

This work was supported by a grant from the Deutsche Forschungsgemeinschaft (KO 3918/5-1).

The idea of the 1st study was developed by Dr. Catarina Amado and Prof. Dr. Gyula Kovács. Dr. Catarina Amado programmed the software and implemented it with the fMRI sequence. With the help of Sandrine Hinrichs and under the supervision of Dr. Catarina Amado, I collected the data presented in this study. Sandrine Hinrichs wrote a bachelor's thesis about a subsample of the participants and was therefore involved in data acquisition and supervised for data analysis. However, the results presented in this article were analyzed by me under the supervision of Dr. Catarina Amado and Prof. Dr. Gyula Kovács. The manuscript was written by me and revised by Prof. Dr. Gyula Kovács and Dr. Catarina Amado. I submitted the manuscript and made all the revisions associated with it.

Study II. Measures of repetition suppression in the fusiform face area are inflated by co-occurring effects of statistically learned visual associations.

This work was supported by a grant from the Deutsche Forschungsgemeinschaft (KO3918/5-1).

The design of the 2nd study relies partly on an EEG experiment performed and published by Dr. Daniel Feuerriegel. Together with Prof. Dr. Gyula Kovács and Dr. Catarina Amado, the idea for an fMRI investigation was developed, and the design was conceptualized. Dr. Daniel Feuerriegel and Dr. Catarina Amado prepared the software. Stimulus material was provided by Dr. Carolin Altmann. I collected both the behavioral data in training sessions and fMRI data, with initial guidance from Dr. Catarina Amado. Part of this dataset was analyzed for my master's thesis at FSU, Jena entitled 'The effects of learning on repetition-related phenomena and stimulus predictions.' However, additional data were collected afterward, and the manuscript was written from scratch. Dr. Daniel Feuerriegel and I performed all analyses. Prof. Dr. Gyula Kovács, Dr. Daniel Feuerriegel, and I wrote the article. Dr. Daniel Feuerriegel and I prepared the publication included the publication of data and analysis codes. I submitted the manuscript and made all the revisions associated with it under the supervision of Prof. Dr. Gyula Kovács, Dr. Daniel Feuerriegel.

Study III. Person identity-specific adaptation effects in the ventral occipito-temporal cortex.

The work was partly funded by the German academic exchange Service DAAD (57446347) and by the Deutsche Forschungsgemeinschaft (KO3918/5-1).

The 3rd study is the result of German-Australian cooperation supported by DAAD. Assoc. Prof. Patrick Johnston and Prof. Dr. Gyula Kovács devised and supervised the project. The experiment was designed by Assoc. Prof. Patrick Johnston, Dr. Géza Ambrus, Dr. Jonathan Robinson, and me. Dr. Géza Ambrus and I prepared the software and implemented it with the fMRI. I selected the stimulus material in consultation with Prof. Dr. Gyula Kovács and Dr. Géza Ambrus, who also preprocessed the chosen stimuli. I carried out the experiment with the help of our Master student Philipp Böhm, and performed all analyses. Prof. Dr. Gyula Kovács and I wrote the manuscript with support from Assoc. Prof. Patrick Johnston, Dr. Géza Ambrus, and Dr. Jonathan Robinson. I submitted the manuscript and prepared the publication of data and analyses.

ACKNOWLEDGEMENTS

Over the past few weeks, I have often thought of all the people around me who have helped me in one way or another to finish this work. And I am very grateful for each of you and any help, big or small, that you offered. I started writing this section before I actually finished this work because I had to put down a few words to thank many people who got me to this point, even if it wasn't the finish line, but 'just' the home stretch.

I was lucky to get the opportunity to do this work and take this difficult step in a great environment. So, first of all, I would like to thank Prof. Dr. Gyula Kovács, who welcomed me to his research team and new department. I have learned a lot from you, and I thank you for all your guidance, advice, patience - especially with my writing, - ideas, all your comments - not only those included in this paper - and research insights. I have always felt very comfortable and gladly accepted any criticism. It was a great atmosphere to work and learn and gain new experiences. Also, thank you for your efforts in arranging some enjoyable collaborations.

I would also like to thank the rest of the team, Géza Ambrus, Charlotta Eick, and Chenglin Li, especially for sharing work, ideas, sometimes just a coffee or some cake, and sometimes just the same room for a few hours. Not only for this reason, but especially after the last somewhat lonely year, I am very grateful for shared office hours with you. I also want to thank Catarina Amado, who was still a member of the team when I just started and taught and supervised me in the early stages of this journey.

Of course, all the other permanent or temporary members of the department deserve to be mentioned here as well. I want to thank Marie Wächter, Lukas Korn, Johannes Lehnen, Rico Stecher, and Alexia Dalski, but also Kathrin Wiese, Bettina Kamchen, and Kathrin Rauscher for their help and support with teaching, participant recruitment, data collection, data analysis, or organizational tasks during my time at FSU. A special thanks go to Ines, who spent hundreds of hours with me in the MRI center.

I want to thank many people who contributed to one project or another. Some of them I never met in person, but they all contributed to my journey, nonetheless. However, a very special thank you goes to Patrick Johnston and Jonathan Robinson! I had the great pleasure of spending a few weeks working with you two in Jena and Brisbane, respectively, and it was a lot of fun. I learned a lot from both of you, and I'm looking forward to finishing our next project. I very much hope that we will meet again soon. I also want to thank Daniel Feuerriegel, without whom the second study would not have been possible at all.

I would also like to thank Prof. Dr. Stefan Schweinberger for being the second examiner for the last part of my PhD thesis. I am very grateful for the time you will invest in reading and evaluating my work.

The PhD club that formed at the end of my first year is a fantastic step towards a shared PhD experience. Sometimes it's just important to know that others are struggling with similar issues, and you can support each other when needed. Thanks so much to all of you guys. Keep up the good work!

Thanks to the best Master study group, Celina, Julia, Lisa, and Tabitha. Thank you for your help, support and patience, especially in the last year. You guys always had my back. You have guided me through every crisis, big and small, with soothing words, distractions, cake, wine, and many other things, and I am incredibly grateful for all of it.

Franzi, Philipp, Karo, Brenda, and Micha - I couldn't have taken the first step without you, and I love that you are still by my side. A special thank you goes to Franzi, who endured my worries and complaints in hours and hours of online co-working sessions during the last weeks.

Of course, many other people have helped me through one crisis or another. Among them are people who have been with me for a very long time: Johanni and Marie! Thank you so much that you still have an open ear for my worries after so many years. Also, many people have accompanied me since the beginning of my studies - have stood by me with advice or sometimes just a coffee or an after-work beer. Many thanks to Alex, Dirk, Franzi, Lina, Mike, Philipp, Thomas & Thomas.

A thank you also goes to my flatmates Sarah and Traci. So many months of lock-in and home office are not easy for anyone, and they weren't for me either. I am happy that I was not alone during this time. And I would like to add that I am very grateful that you have all put up with and forgiven my whims over the past few months.

Last but not least, a very, very special thank you goes to my family, especially to my mother Anne-Katrin, my father Heiko, and my sister Susann. You are always there for me. You have brought me to where I am now, and I know that I owe it to you that I have come this far at all. You always supported me in everything, always believed in me, and gave me confidence when I needed it. You can't imagine how grateful I am to be part of our family. I also want to thank my grandparents for their love and support and for teaching me about life. Thank you for being such a good role model for me, and thank you for just always believing in me.

STATUTORY DECLARATION

Ehrenwörtliche Erklärung

Hiermit erkläre ich, dass mir die geltende Promotionsordnung der Fakultät für Sozial- und Verhaltenswissenschaften bekannt ist.

Des Weiteren bestätige ich, dass ich die vorliegende Dissertation selbst angefertigt habe ohne Textabschnitte eines Dritten oder einer eigenen Prüfungsarbeit ohne Kennzeichnung übernommen zu haben. Alle von mir benutzten Hilfsmittel, persönliche Mitteilungen und Quellen in meiner Arbeit habe ich angegeben.

Bei der Auswahl und Auswertung des Materials sowie bei der Herstellung der Dissertation wurde ich von den Co-Autoren der vorgelegten Manuskripte unterstützt. Die jeweilige Beteiligung ist unter „Beiträge der Autoren“ (Author Contributions) detailliert beschrieben.

Die Hilfe eines kommerziellen Promotionsvermittlers habe ich nicht in Anspruch genommen und Dritte haben weder unmittelbar noch mittelbar geldwerte Leistungen von mir für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Ich bestätige außerdem die Dissertation noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht zu haben, sowie die gleiche, eine in wesentlichen Teilen ähnliche oder eine andere Abhandlung bei einer anderen Hochschule als Dissertation eingereicht zu haben.

Jena, den ______________________________

____

Sophie-Marie Rostalski