

ORIGINAL ARTICLE

Statistical Learning Signals in Macaque Inferior Temporal Cortex

Peter Kaposvari^{1,2,†}, Susheel Kumar^{1,†} and Rufin Vogels¹

¹Laboratorium voor Neuro-en Psychofysiologie, Departement of Neurosciences, Campus Gasthuisberg, B 3000 Leuven, Belgium and ²Department of Medical Physiology, University of Szeged, Dóm square 10, 6720 Szeged, Hungary

Address correspondence to Rufin Vogels, Laboratorium voor Neuro-en Psychofysiologie, Departement of Neurosciences, Campus Gasthuisberg, B 3000 Leuven, Belgium. Email: rufin.vogels@med.kuleuven.be

[†]These authors contributed equally to this work.

Abstract

Humans are sensitive to statistical regularities in their visual environment, but the nature of the underlying neural statistical learning signals still remains to be clarified. As in human behavioral and neuroimaging studies of statistical learning, we exposed rhesus monkeys to a continuous stream of images, presented without interstimulus interval or reward association. The stimulus set consisted of 3 groups of 5 images each (quintets). The stimulus order within each quintet was fixed, but the quintets were presented repeatedly in a random order without interruption. Thus, only transitional probabilities defined quintets of images. Postexposure recordings in inferior temporal (IT) cortex showed an enhanced response to stimuli that violated the exposed sequence. This enhancement was found only for stimuli that were not predicted by the just preceding stimulus, reflecting a temporally adjacent stimulus relationship, and was sensitive to stimulus order. By comparing IT responses with sequences with and without statistical regularities, we observed a short latency, transient response suppression for stimuli of the sequence with regularities, in addition to a later sustained response enhancement to stimuli that violated the sequence with regularities. These findings constrain models of mechanisms underlying neural responses in predictable temporal sequences, such as predictive coding.

Key words: inferior temporal cortex, macaques, predictions, statistical learning, temporal regularities

Introduction

Animals are sensitive to spatial and temporal regularities in their sensory environment. These regularities correspond to properties of the environment that repeat, for example, for primates the configuration of features that constitute a face or the sequence of buildings encountered when driving home along a familiar road. Behavioral studies have shown that some of these regularities can be learned (for reviews see Krogh et al. 2012; Turk-Browne 2012; Dehaene et al. 2015). For instance in the visual domain, mere exposure to a set of scenes or sequences of visual stimuli is sufficient to learn their embedded regularities (Fiser and Aslin 2001, 2002; Turk-Browne 2012). Such extraction of statistical regularities is

often referred to as “statistical learning” (Saffran et al 1996; Turk-Browne 2012).

In the present study, we examined the learning of temporal statistical regularities in macaque inferior temporal (IT) cortex. IT represents complex object features (Tanaka 1996) and previous studies have shown that the object selectivity of adult macaque IT neurons can be altered by visual experience or by visual discrimination training (Miyashita et al. 1993; Kobatake and Tanaka 1994; Baker et al. 2002; Freedman et al. 2006; Op de Beeck et al. 2007; Li and DiCarlo 2008, 2010). For instance, extensive passive exposure to a set of images of objects decreases the response of the population of IT neurons whereas the selectivity for the familiar images increases (Freedman et al. 2006; Woloszyn and Sheinberg 2012).

Given this plasticity of the object selectivity of IT neurons, it is natural to ask whether their responses also reflect temporal regularities in object sequences.

In studies of visual statistical learning of temporal regularities, subjects are exposed to multiple short sequences, each consisting of stimuli that are always presented in the same fixed order. The short sequences themselves are presented in a random order without interruption. Turk-Browne et al. (2009) found that functional magnetic resonance imaging (fMRI) blood oxygen level-dependent (BOLD) activations in the human lateral occipital complex (LOC) were greater for learned fixed sequences than for sequences of randomly ordered stimuli with equal familiarity, suggesting a statistical learning signal in human LOC, a possible homolog of macaque IT (Denys et al. 2004). Recent single unit studies in macaque IT argued instead for a decreased response to predictable stimulus members of familiar doublets (Meyer and Olson 2011; Ramachandran et al. 2016) or triplets (Meyer et al. 2014). Unlike in classical statistical learning studies, the short sequences in these macaque studies were each followed by a reward and clearly demarcated in time with an intersequence break during the familiarization phase. This may not only have enhanced sequence learning in the monkeys but also altered the underlying mechanisms (e.g., the animals might have employed the stimuli to predict the upcoming juice reward). Furthermore, the apparent suppressed responses to the learned sequences observed in the Meyer et al. IT studies (Meyer and Olson 2011; Meyer et al. 2014; Ramachandran et al. 2016) contradicted the stronger neural responses to predicted compared with deviant or randomly ordered stimuli reported in mice visual cortex (Gavornik and Bear 2014).

At present, the reason(s) for these discrepancies between the different studies is unclear, in particular because different stimulation protocols were employed in different species. To resolve this, we employed in the present study a classical statistical learning paradigm, similar to the one employed by Turk-Browne et al. (2009), to examine visual statistical learning in macaque IT. Unlike in the previous studies in macaque IT, we exposed monkeys to an uninterrupted, continuous stream of visual stimuli and reward delivery was independent of the stimulus sequence or stimulus identity. The stream consisted of 3 short sequences of 5 different stimuli each. During exposure, the stimulus order was fixed within each quintet: for example, stimuli A, B, C, D, and E of a quintet were shown always in the same order "ABCDE" but never in a different order (e.g., ACDBE). Importantly, the 3 quintets were presented, uninterrupted, in a random order (e.g., ABCDEKLMNOFGHIJKLMNOABCDE...) and the 3 quintets were presented with equal probability.

In the present study, we performed 2 experiments to examine whether and how IT neurons were sensitive to the statistical regularities embedded in the continuous stream of equally familiar stimuli. In the first experiment, we tested the responses of IT neurons for violations of the exposed quintets. To do so, we compared the spiking activity between the same stimuli presented at the exposed order within the quintet (e.g., to "D" in ABCDE) and when presented at a "deviant" location within one of the other two quintets (e.g., to "D" in KLMDO). We asked whether the exposure produced different responses for the deviant stimulus compared with the same stimulus presented at its correct position in the exposed quintet. Such a differential response would indicate a statistical learning signal in IT. This differential response could be an enhanced response to the

deviant stimulus, as seen in previous studies of sequence learning in macaque IT (Meyer and Olson 2011; Meyer et al. 2014; Ramachandran et al. 2016), or instead a suppressed response to the deviant stimulus, as seen in rodent visual cortex (Gavornik and Bear 2014) and suggested by the human fMRI study by Turk-Browne et al. (2009).

In addition, we examined to what aspects of the statistical structure embedded in the continuous sequences IT neurons were sensitive. Thus, we asked whether the learning signal in IT depended only on the immediate preceding stimulus or instead reflected a larger part of the quintet. This was tested by replacing 2 successive stimuli of a quintet by 2 successive stimuli of another quintet (e.g., LM of the quintet KLMNO replaced CD of ABCDE, yielding ABLME). If monkey IT statistical learning signals are restricted to temporally adjacent stimuli then one would expect that there would be no change in the response to the deviant that follows the first deviant of a pair in the correct position (i.e., response to M in ABLME and in KLMNO), although the deviant pair is inserted in the wrong quintet. Furthermore, the response to a stimulus at the correct position that follows a single deviant, for example, response to "O" in KLMDO, should show a change in its response when the learning effect in IT is determined by the preceding stimulus since O never followed D during the exposure, although O is at its correct position in the quintet (KLMNO). No change in response to the stimulus following the deviant, combined with a differential response to the deviant, would indicate learning of nonadjacent dependencies. Another aspect we examined is the possible cumulative effect of the number of preceding stimuli within a quintet: one may expect a greater differential response to the deviant when it occurs at a later position in the quintet (e.g. greater response to deviant H in ABHDE vs. H in ABCDH). Such dependency of the differential response to a deviant on its position within the quintet would be expected when the response to a stimulus depends on more than one preceding stimulus of a quintet.

Next we asked whether IT neurons were sensitive to the temporal order of the stimuli within a quintet or were coding merely joint statistics. To answer this, we manipulated the position of a stimulus within a quintet, by replacing a stimulus of a quintet by another stimulus of the same quintet (e.g. comparing the response to the first E presentation in ABEDE with that in ABCDE). For both the across-quintet and within-quintet stimulus replacements, we controlled for mere repetition effects, since repetition, even with intervening stimuli, is known to influence the response of IT neurons (Miller et al. 1991; Sawamura et al. 2006; McMahon and Olson 2007; Kaliukhovich and Vogels 2014).

In the first experiment, we observed an enhanced activity for the first deviant of a stimulus pair that was inserted at the wrong position in a quintet, compared with the same stimulus presented at the correct position within a quintet. Then we asked in a second experiment whether this enhanced response reflects 1) a response increase to the deviant stimulus relative to the response to a stimulus in a random sequence or 2) a decrease of the response to a predicted stimulus and without an enhanced response to the deviant with respect to a stimulus in a random sequence. In other words, is there a surprise-like, enhanced response to the deviant stimulus and/or expectation suppression (Todorovic and de Lange 2012; Summerfield and de Lange 2014). Note that thus far, no evidence for surprise-related responses has been obtained in macaque IT (Kaliukhovich and Vogels 2014). We answered this question by comparing the spiking activity between stimuli of the exposed quintets and

equally familiar unpredictable stimuli that were presented in a random order during the exposure phase, and by comparing the activity to deviant stimuli and random sequence stimuli.

Materials and Methods

Subjects and Recording Locations

Data were collected from 2 male rhesus monkeys (H and O; *Macaca mulatta*). All animal care and experimental protocols complied with national and European guidelines and were approved by the KU Leuven Ethical Committee for animal experiments. Both monkeys had an MRI compatible plastic headpost and recording chamber fixed to the skull with acrylic cement and ceramic screws. The recording positions were estimated by an MRI scan obtained before the recording sessions. Multi-unit activity (MUA) recordings were made from the ventral bank of the rostral Superior Temporal Sulcus (STS) of the right hemisphere. The estimated recording locations of monkey H ranged between 14 and 18 mm anterior to the ear canal and between 20 and 25 mm lateral to the midline. For monkey O, the anterior/posterior coordinates of the recording locations ranged from 15 to 19 mm and the medial/lateral coordinates ranged from 23 to 26 mm.

Apparatus

MUA recordings were performed with epoxylite-insulated tungsten microelectrodes (impedance ranging between 0.4 and 1.8 M Ω). In each daily session, a single electrode was lowered with a Narishige microdrive into the brain using a stainless steel guide tube. Spikes were thresholded online using a custom amplitude discriminator. No effort was made to isolate single units from the thresholded MUA since in the present study we were interested in the effects that occur at the population level. By pooling the activity of multiple neurons, MUA has the advantage of reducing stimulus specific effects, which is an advantage when comparing responses across different stimuli as was done here in some of the analyses.

The position of one eye was continuously tracked by means of an infrared video-based tracking system (SR Research EyeLink; sampling rate 1 KHz). Gamma corrected stimuli were displayed on a 20-inch CRT display (Panasonic Panasync/ProP110i; 1024 \times 768 screen resolution; 75 Hz vertical refresh rate) at a distance of 57 cm from the monkey's eyes.

Fixation Task

The animals were required to fixate within a 2 degree square fixation window, centered around a small fixation target, in order to obtain a juice reward. Juice rewards were given with decreasing intervals as long as the monkeys maintained fixation, encouraging long fixation. Importantly, the timing of the juice delivery and the presentation of the stimulus sequences were uncorrelated. A red square (0.13 degree) served as fixation target and was located at the center of the display, superimposed on the center of mass of the stimuli.

Stimuli and Paradigms

The stimuli consisted of modified Snodgrass and Vanderwart images of animals and objects (Fig. 1B), taken from the [Rossion and Pourtois \(2004\)](#) database. The stimuli were presented on a gray background.

Experiment 1

The 15 stimuli were grayscale images of vertebrate animals, all with their head oriented to the left (Fig. 1B). We resized the images so that their maximal horizontal or vertical extent was 6 degrees and equated their mean luminance. The 15 images were sorted twice into groups of 3 quintets. For each of the 2 groups of 3 quintets, we selected quintets so that the area, maximum vertical extent, maximum horizontal extent, aspect ratio and standard deviation of the pixel graylevels of the images did not differ significantly among the 3 quintets of the group and we maximized the overlap of these values (one-way ANOVA, $P > 0.5$). One of the two groups served as stimuli for one monkey and the other group for the other monkey. To select the order of the stimuli within a quintet, we computed pixel-based pairwise Euclidean distances between the graylevels of the images for all possible permutations of the 5 images of a quintet ("within-quintet distances"). In addition, we computed the Euclidean distance between the last image of a quintet and the first image of the same or another quintet and this for all 9 possible combinations of the 3 quintets ("between-quintet distances"). The selected order of the stimuli within a quintet was chosen to have strongly overlapping within- and between-quintet distances. Also, the within-quintet distances were on average (across the 3 quintets) highly similar for the 4 successive doublets of a quintet. Thus, we equated as much as possible low-level image differences between successive presentations within and across quintets.

During the exposure phase of the experiment, the selected 3 quintets, which we will label "standard quintets," were shown in a random order, but with a fixed stimulus order within a quintet. An individual image was shown for 293 ms and was immediately succeeded by the next image, without any inter-stimulus or interquintet interval (Fig. 1A). Unlike in other experiments, the stimulus presentation continued even when the animal aborted fixation.

The exposure phase lasted for 60 and 28 daily sessions, of approximately 2 h each, for monkeys H and O, respectively. After the exposure phase, we performed 2 sets of manipulations in which the order of the stimuli within the sequences was violated. In the first set of manipulations of experiment 1, one or two successive images of a quintet (the "source quintet") were inserted into another quintet by replacement (Fig. 1C). We will label the inserted images "deviants" and the quintets with deviants "target quintets" (see second and third rows of Fig. 1C). For the first 84 MUA sites, the deviants were any of the 5 images of a source quintet but could replace only the third, fourth, or fifth image of the target quintet. For the next 44 MUA sites, we excluded the first image of a source quintet as possible deviant. The deviants consisted of either a pair of stimuli ("deviant pair") or a single stimulus ("single deviant") from another source quintet. We will label the first stimulus of a deviant pair as "first deviant" and the second stimulus of the pair as "follower deviant" (Fig. 1C). Note that the follower deviant (e.g. D in KLMCD with CD being the deviant pair) is identical to the image that follows the first deviant image in its source quintet (ABCDE). By having deviant pairs, we could assess whether statistical learning signals are restricted to temporally adjacent stimuli. If this was the case, then one would expect that there would be no change in the response to the follower deviant since it follows the first deviant of a pair at the same position as in the exposed sequence. Single deviants are by definition all "first deviants." The stimulus that follows a single deviant will be labeled "follower standard" (Fig. 1C). In order to have a sufficiently large number of presentations for both

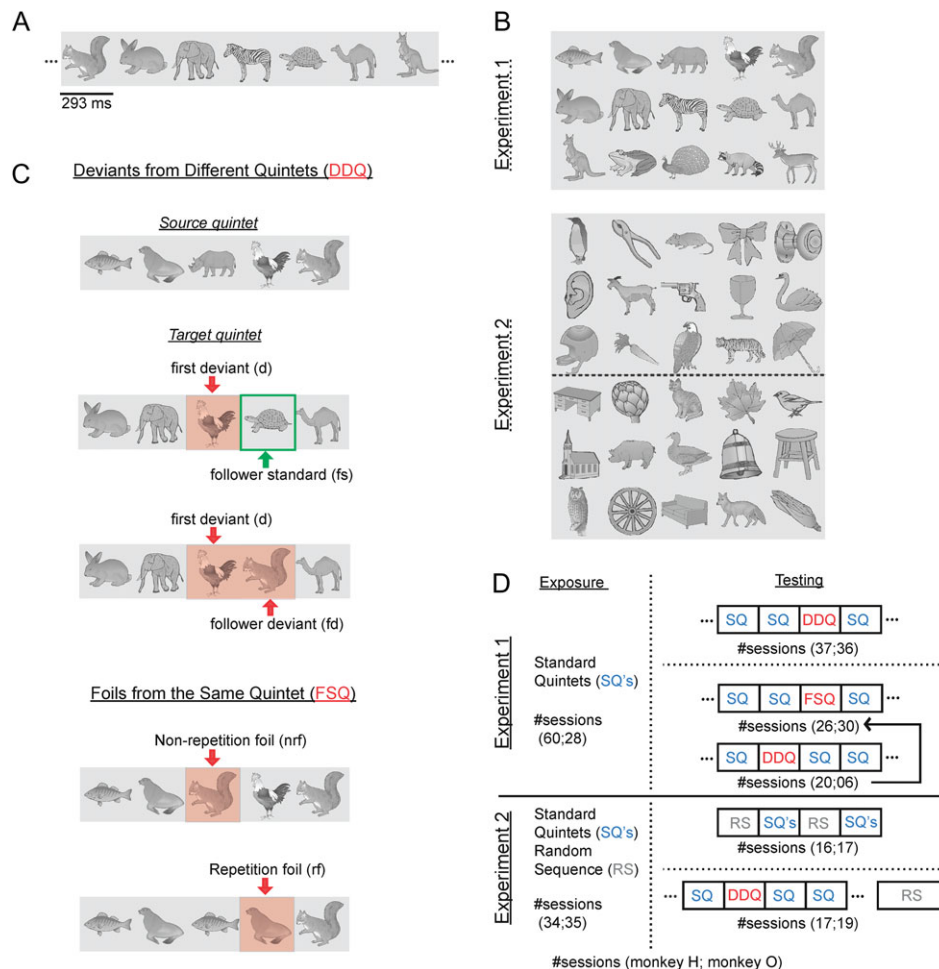


Figure 1. Experimental design. (A) Stimuli were presented in a continuous sequence, trailing each other, without an interstimulus time interval. The exposure duration of a single stimulus was 293 ms. The figure does not show the fixation target. (B) The top panel shows the 15 stimuli (all animals) of experiment 1 ordered in the 3 quintets (rows) presented to monkey H. The assignment of the same stimuli into the 3 quintets differed for the other monkey. During exposure, the 3 quintets were presented in a random order, but the stimulus order within a quintet was fixed. The bottom panel shows the stimuli of experiment 2. The 30 stimuli were sorted in 2 groups of 15 each as indicated by the stippled line. One group of stimuli was presented in quintets (rows) as in experiment 1, while the stimuli of the other group were presented in a random order. The upper 3 rows show the quintets employed for monkey H, while the bottom 3 rows were the quintets for monkey O. Stimuli of the lower 3 and upper 3 rows were presented in the random sequence in monkeys H and O, respectively, thus counterbalancing the 2 groups across the 2 animals. (C) Top: example of deviants from different quintets (DDQ). Illustration of a source quintet (first row), a target quintet containing a single deviant (second row) and a target quintet containing a deviant pair (third row). The first deviant (d; red shaded box) and the follower standard (fs; green box) are indicated in the single deviant target quintet. The follower standard is identical to the stimulus presented at that position in the standard quintet to which the monkey was exposed (see stimulus at position 4 at row 2 in B). The first deviant (d) and the follower deviant (fd) are indicated in the deviant pair target quintet in a shaded box. Bottom: example of foils from the same quintet (FSQ). The fourth and fifth rows show an example of a target quintet with a nonrepetition foil (nrf) and a repetition foil (rf), respectively. The foils are indicated by a red shaded box. (D) Summary, in chronological order, of the different exposure (left) and testing (right) phases. Experiment 1: during the exposure phase, only standard quintets (SQs) were presented. The testing phase consisted of 2 parts. First, target quintets containing deviants (DDQ) were randomly interleaved with standard quintets in continuous sequences (see Materials and Methods). Second, target quintets containing a foil (FSQ) were randomly interleaved with standard quintets (see arrow). In this second part, we performed also daily sessions with sequences having deviants and the latter were interleaved with daily sessions consisting of the sequences containing foils. Experiment 2: the monkeys were exposed to blocks containing standard quintets or random sequences (RS) of stimuli. The testing phase of experiment 2 consisted of 2 parts. First, alternating blocks of standard and random sequences. Second, we presented sequences containing standard quintets and target quintets with only single deviant DDQs, and these were alternated with random sequences. The number of sessions for the exposure phases and testing phase parts is indicated. The first and second number (within brackets) shows the number of sessions for monkeys H and O, respectively.

single and deviant pairs, different MUA sites were tested with either deviant pairs or single deviants in a random manner. During the recordings, target quintets with deviants were presented at random in 25% of all the quintet presentations, with the restriction that at least one standard quintet should be presented in-between 2 target quintets. Tests with deviants were performed during 57 and 36 daily recording sessions in monkeys H and O, respectively.

After manipulation 1 of experiment 1, we exposed again monkeys H and O to standard sequences without deviants in 12 daily

sessions, followed by the second manipulation. After this exposure, in the second set of manipulations, we replaced stimuli from the same quintet ("within-quintet" foils). All stimuli of a particular quintet could replace any of the other images at positions 2–5 of that quintet. To reduce repetition suppression effects, we allowed no immediate repetitions of images (e.g., ABCCE). We distinguished 2 sorts of within-quintet foils: 1) foils that repeated a stimulus of the same quintet (e.g., "B" in quintet "ABCBE"), which we label "repetition foils" and 2) foils that were not a repetition of a stimulus within the quintet (e.g., "D" in "ADCDE"),

labeled as “nonrepetition foils” (Fig. 1C). The reason for separating the 2 sorts of foils is that pooling foils, irrespective of repetition, may underestimate a response enhancement to the repetition foil because of counteracting repetition suppression in the case of the repetition foils. Quintets with foils were presented at random in 25% of all the quintet presentations, with the restriction that at least one standard quintet should be presented in-between 2 target quintets. Tests with foils were performed during 26 and 30 daily recording sessions in monkeys H and O, respectively. In-between within-quintet foil recording sessions, we also performed daily sessions in which we presented across-quintet deviants.

For both across- and within-quintet manipulations, we presented target quintets only during the measurement of the MUA to these sequences, after isolating a responsive MUA site. During the recordings, each deviant or foil stimulus was presented 120 times, yielding an approximate recording duration of 36 min per site. In-between these MUA recordings and while searching for responsive MUA sites, the monkeys were exposed to the standard sequences using the same procedure as during the exposure phase. On purpose, we limited the number of presentations of the deviants or foils, thus reducing unlearning of the sequences. We searched for responsive neurons by presenting the standard sequence. MUA sites that were judged to be responsive to at least one standard sequence stimulus (based on online peristimulus time histograms) were tested further with a sequence that included deviants or foils.

Experiment 2

The stimuli of experiment 2 consisted of 2 groups of 15 stimuli each, all different from those of experiment 1 (Fig. 1B). Each group consisted of a mixture of animals and objects. One group of 15 images consisted of 3 standard quintets with a fixed order of stimuli, designed using the same procedure as in experiment 1. These 3 standard quintets were presented in a random order, using the same protocol as in experiment 1, defining “standard sequences.” The other group of stimuli consisted of 15 other animals and objects and these were presented in a pseudorandom order. These sequences will be labeled as “random.” We required that the number of stimuli in-between presentations of the same stimulus should be at least 4. This ensured that on average the number of intervening stimuli between repetitions of the same stimulus was equal for both the standard and random sequences. The stimulus groups presented in standard or random sequences were counterbalanced across the 2 animals.

We exposed the animals to both types of sequences using a block design. Each block consisted of 4050 stimuli (270 presentations per stimulus), lasting approximately 20 min, and standard and random sequence blocks were alternated in daily sessions of approximately 2 h. The sequence type of the first block of a daily session was randomized across sessions. We kept track of the number of presentations per sequence type and when necessary increased temporally the number of presentations for a particular type in order to equate the number of presentations per sequence type. Thus, we ensured that the stimuli of the standard quintets and random sequences had equal familiarity. The exposure phase in experiment 2 lasted 34 and 35 daily sessions in monkeys H and O, respectively. The exposure phase of experiment 2 started 2 and 48 days after the end of the recording sessions of experiment 1 in monkeys H and O, respectively.

After the exposure phase, we recorded MUA for blocks of standard and random sequences. We searched for responsive

MUA using either the standard quintets or random sequences, alternating between sites. This avoided biasing responses toward one or the other sequence type. We tested each responsive MUA site with 2 blocks of each sequence type, that is, 4 blocks in total, and the 2 types alternated. The sequence type of the first block was randomized across sites. Each block contained 70 presentations per stimulus and lasted approximately 5 min, yielding a total test time of 21 min per MUA site. We kept track of the number of presentations per sequence, in particular during the search periods in which a single sequence type was presented, and when necessary compensated by presenting the less frequent stimuli more in later sessions, e.g. at the end of recording or during the next days. We performed 16 and 17 such daily recording sessions in monkeys H and O, respectively.

After comparing the MUA for standard and random sequences, we introduced quintets with deviants, starting in the next recording session. The deviants in a target quintet originated from the 2 other quintets (Fig. 1C) and replaced stimuli of the third, fourth, and fifth positions of the target quintet. As in experiment 1, only 25% of the quintets contained a deviant but only single deviants were presented. The presentation of the target quintets followed the same protocol as in experiment 1. As before, responsive MUA was searched using either standard quintets (without deviant) or random sequences and this was done with an equal frequency. When testing the MUA sites, we alternated blocks of standard and target quintets and blocks of random sequences. Because of the inclusion of deviants, we increased the duration of the blocks containing both standard and target quintets by presenting each stimulus 480 times per block, of which 25% were target quintets. The number of presentations per stimulus was 140 for the random block. We equated the presentation frequency of stimuli of the random and standard quintets by increasing the duration of the random sequences in sessions that occurred in-between the recording sessions. MUA to both types of sequences was measured in 17 and 19 daily recording sessions in monkeys H and O, respectively.

Figure 1D summarizes in chronological order the different exposure and test phases, including the number of sessions for each phase per monkey.

Data Analysis

All statistical analyses were performed with nonparametric or permutation tests, except noted otherwise. For all analyses, we included only those presentations during which the monkey was fixating the fixation target. Also, the fixation duration was restricted to be at least 2344 ms and data of the last 250 ms of a fixation period were removed to avoid effects of saccade preparation. We included for further analysis only those MUA sites that showed a statistically significant response to the sequences presented without deviants. This testing was performed for the presentations of quintets that were presented in full during a fixation period. We employed a permutation test of the variances of the mean responses computed across the 15 stimuli of a neuron, following a similar procedure as adopted by De Baene et al. (2007). To do so, we binned the responses of each stimulus presentation in 75 ms long bins and then we concatenated the bins for the 15 stimuli of a sequence for a single presentation of the stimuli. This was repeated for each of the minimum number of presentations, computed across the different stimuli of the sequence. Then we computed the variance of the mean responses, averaged across presentations, in the different bins.

This variance metric, which reflects the modulation of the binned response by the stimulus sequence, was then compared with the null distribution of the variances expected when there was no relationship between stimulus and binned responses. The latter variance was computed after permuting the order of the bins separately per presentation number. The null distribution was obtained by computing 1000 such variances on permuted data. If the variance of the observed binned responses was larger than the 95th percentile of the null distribution, the MUA of the site was considered to be significantly modulated by the stimuli. Following the permutation test, 87% of the MUA sites recorded in experiment 1 of monkey H were accepted, while all recorded MUA sites of both experiments in monkey O and all MUA sites recorded in experiment 2 of monkey H were accepted for further analysis.

We computed firing rates per stimulus presentation using an analysis window of 280 ms that started 100 ms after stimulus onset (to account for the response latency we observed in our neuronal sample) and that ended 87 ms after stimulus offset (to reduce contamination of the response by the following stimulus). The firing rates were averaged across presentations of the same stimulus. When comparing responses to deviants or foils with presentations of the same stimuli in their correct position ("standard element") for an MUA site, we equated the number of the 2 sorts of presentations by randomly sampling the same number of standard element presentations as the available number of deviant/foil stimuli of that MUA site. Responses to deviants, foils, and corresponding standard elements were only considered when these were preceded by at least 5 images and succeeded by at least 2 images during a fixation period. Also, stimuli presented as the first stimulus of a quintet were excluded, except when otherwise noted, because these followed another stimulus with a transitional probability of only 1/3. To analyze the time course of the responses, we computed firing rates in nonoverlapping 20 ms long bins. We tested whether there was a significant difference between 2 stimulus conditions (e.g. deviant vs. standard element) with a 2-sided Wilcoxon matched pairs test for each bin between 100 and 380 ms after stimulus onset. The *P* values were corrected for multiple comparisons using the [Benjamini and Hochberg \(1995\)](#) False Discovery Rate method ($q < 0.05$ was considered as statistically significant). In all analyses, we first averaged across stimuli tested for a site and this average provided the data point for a site. In other words, the number of observations in the statistical tests corresponded to the number of MUA sites.

To compare the firing rates between the first deviant and standard element presentations as a function of the effectiveness of that stimulus in driving the MUA, we randomly split the presentations of a stimulus into 2 halves. This was done after equating the number of presentations for the deviant and standard element conditions. Then, using the mean responses to the standard element in one half of the presentations, we selected 3 stimuli for each MUA site: one with the lowest mean firing rate (*w* in Fig. 3), one with the median firing rate (*m* in Fig. 3) and one with the highest firing rate (*b* in Fig. 3). The other half of the presentations of the standard elements were employed to compute the responses to the standard element, which were compared with the responses to the same stimuli presented as first deviants in half of their presentations. This procedure ensured that the responses that were employed to compute the stimulus preference of the MUA site were independent from the responses that entered the comparison of standard element and deviant stimulus presentations as a function of the stimulus preference

(plotted in Fig. 3), avoiding biasing the stimulus selectivity for a particular stimulus condition.

All analyses were performed for both raw firing rates and firing rates that were normalized for each neuron separately. The normalization was performed by subtracting the mean firing rate, averaged across conditions (standard and deviant or foil presentations), from the firing rate of a stimulus and dividing that difference by the mean firing rate. The figures display the unnormalized data. Since all reported effects were also present for normalized spike counts, we are assured that the effects were not driven by outlier sites with high firing rates.

Results

Experiment 1

In the first experiment, we exposed 2 monkeys for several weeks to a continuous stream of visual images (Fig. 1A,B). The 15 visual images were grouped into 3 groups of standard quintets, and the stimulus order was fixed within each quintet. The 3 standard quintets were shown in a random order, without any interruption. After the exposure phase, we tested whether anterior IT neurons carry a signal that reflects the temporal regularities of the stimulus sequence by inserting in 25% of the quintet presentations 1 or 2 successive stimuli ("deviants") that belonged to the other quintets or by inserting a stimulus ("foil") from the same quintet at the "wrong" position within that quintet. The responses to the deviants or foils were compared with the responses to the same stimuli ("standard elements") presented at their correct position in the standard quintets.

Deviants from Different Quintets

The deviant stimuli consisted of either a pair of stimuli ("deviant pair": first deviant and follower deviant), that is, 2 successive stimuli from another source quintet, or a single stimulus ("single deviant": first deviant) from another source quintet (Fig. 1C). In a first analysis, we determined whether there is a difference in the response to a single deviant and its corresponding standard element. To do this, we pooled the MUA sites tested with deviant pairs and single deviants and analyzed for the MUA sites tested with deviant pairs the response to the first deviant. We found that the average MUA to the first deviants was significantly greater than to the physically identical stimuli presented as a standard elements ($P = 1.62 \times 10^{-13}$, $n = 128$ sites; Wilcoxon matched pairs test; Fig. 2A) and this response enhancement was present in each animal (monkey H: $P = 2.92 \times 10^{-7}$, $n = 62$, Supplementary Fig. 1A; monkey O: $P = 1.10 \times 10^{-7}$, $n = 66$, Supplementary Fig. 1B).

To determine whether the response enhancement for the first deviant depended on how well that stimulus was driving the neurons, we selected from the stimuli that were presented as first deviants those that elicited the worst, median and best response when presented as a standard, using an unbiased stimulus selection procedure (see Materials and Methods). We performed this analysis for all the MUA sites with at least 6 presentations of each first deviant stimulus (105 sites). The response enhancement for the deviant was present for each stimulus, irrespective of its effectiveness in driving the MUA (Fig. 3). This additive effect of the first deviant stimulus was confirmed by a repeated measures ANOVA with factors 1) deviant versus standard and 2) stimulus rank that showed a significant main effect of both factors (main effect of deviant/standard: $F_{1,104} = 27.2$; $P = 1.0 \times 10^{-6}$; main effect of stimulus rank: $F_{1.2,128.5} = 55.7$; $P = 2.21 \times 10^{-13}$; Greenhouse-Geisser

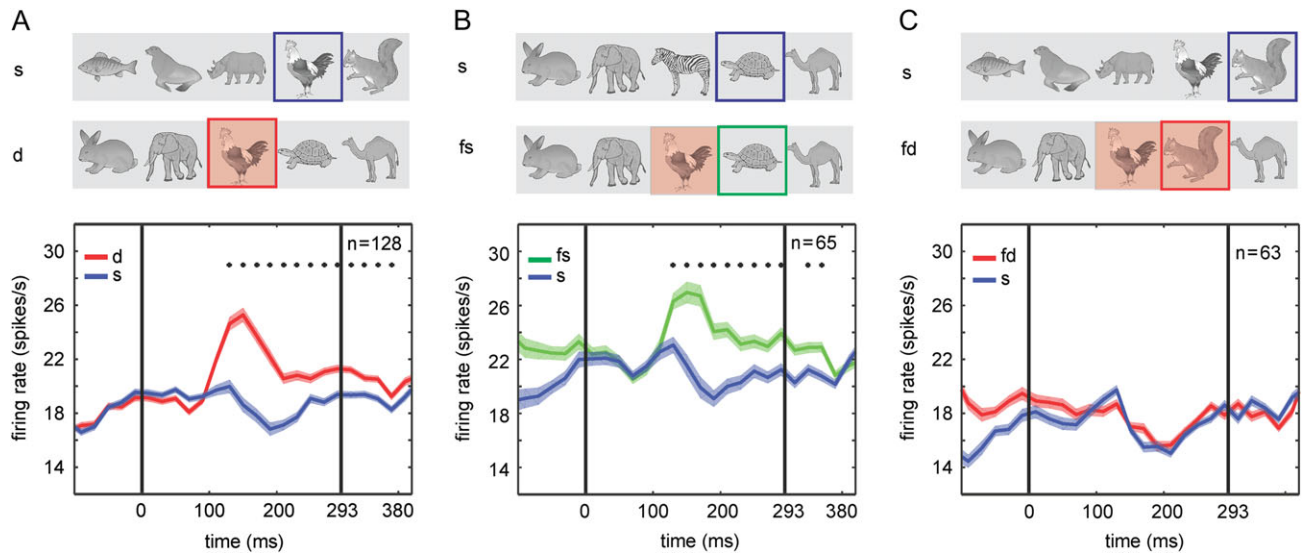


Figure 2. Mean spiking activity for a first deviant, for a follower standard and a follower deviant compared with the activity for corresponding standard elements. The deviants originated from another quintet (cross-quintet deviants). (A) Top: illustration of a source quintet and a target quintet containing a first deviant. The stimuli for which the responses are compared are indicated by a colored square, with the blue and red indicating the standard element (s) and first deviant (d) stimulus, respectively. The shaded square indicates a deviant. Bottom: Firing rate averaged across both monkeys ($n = 128$ MUA sites) for a first deviant (red) and the standard element (blue). (B) Top: illustration of a source quintet and a target quintet containing a first deviant. The stimuli of which the responses are compared are indicated by a colored square, with the blue and green indicating the standard element and a follower standard stimulus, respectively. The shaded square indicates the deviant. Bottom: firing rate averaged across both monkeys ($n = 65$ MUA sites) for a follower standard (fs) (green) and the corresponding standard element (blue). (C) Top: illustration of a source quintet and a target quintet containing 2 successive deviants. The stimuli of which the responses are compared are indicated by a colored square, with the blue and red indicating the standard element (s) and a follower deviant (fd), respectively. The shaded squares indicate the deviants. Bottom: firing rate averaged across both monkeys ($n = 63$ MUA sites) for a follower deviant (fd; red) and the standard element (s; blue). In each data panel, the stars indicate significant differences between the 2 conditions (Wilcoxon matched pairs test; FDR corrected $q < 0.05$), tested per 20 ms bin. Zero on the abscissa indicates stimulus onset and the vertical bars correspond to stimulus onset and offset. Shaded bands indicate standard error of the mean, computed following the procedure by Loftus and Masson (1994) which removes the variance due to the differences in the overall mean response across MUA sites. Data from experiment 1. See Supplementary Figure 1 for data of the individual monkeys.

sphericity violation correction), but no significant interaction ($F_{1.6,171.5} = 0.03$; $P = 0.95$; Greenhouse-Geisser sphericity violation correction). Similar results were obtained when normalizing for each site by dividing the MUA by the maximum response of the 6 conditions.

Next, we examined what aspects of the statistical structure of the sequences IT neurons were sensitive to. First, we examined whether only the preceding stimulus influences the response to a stimulus, that is, whether IT is sensitive only to temporal adjacent stimulus relations. The single deviants were followed by a standard stimulus (“follower standard”) in the target quintet, except when the deviant stimulus was located at the fifth position of the target quintet. The follower standard stimulus occupies the same position within the target quintet as the corresponding standard element would occupy in the standard quintet. However, in the standard quintets, the corresponding standard element to the follower standard stimulus never follows the deviant stimulus. Hence, when the IT responses are sensitive to adjacent stimuli, that is pair-based, and do not take into account preceding nonadjacent stimulus associations in a quintet, one would expect a response enhancement also for the follower standard (e.g., response enhancement to N in KLHNO compared with N in KLMNO). Indeed, we found that the response to a follower standard was enhanced with respect to the response to the same stimulus when presented in a standard quintet without a deviant (“standard element”; $P = 9.42 \times 10^{-9}$; $n = 65$; Wilcoxon matched pairs test; Fig. 2B). This response enhancement for the follower standard was present in each monkey (monkey H: $P = 2.21 \times 10^{-5}$, $n = 35$, Supplementary Fig. 1C; monkey O: $P = 8.19 \times 10^{-5}$,

$n = 30$, Supplementary Fig. 1D) and similar in magnitude and time course to the response enhancement observed for the first deviant (compare Fig. 2A,B). Furthermore, when the learning signal in IT is pair-based, one would expect also no enhancement to the follower deviant of a deviant pair (e.g. comparing response to M in ABLME and KLMNO), since the former stimulus is at the correct position given the first deviant. In agreement with this prediction, we found that the mean response to the follower deviant of a deviant pair did not differ significantly from that to the corresponding standard element ($P = 0.19$; $n = 63$; Wilcoxon matched pairs test; Fig. 2C). This absence of response enhancement for the follower deviant (Fig. 2C) together with the presence of response enhancement for a follower standard (Fig. 2B) strongly suggests that the statistical learning signal in IT represents learned pairs.

If the statistical learning signal represents mainly learned pairs, then one would not expect an effect of the position of the first deviant in the target quintet on the response to that deviant (e.g. response to deviant H in ABHDE vs. H in ABCDH). Only the preceding stimulus would matter, not its global position in the quintet. To examine this, we computed for each MUA site the response to the first deviant and corresponding standard element for each of the 3 positions (3, 4, and 5, e.g., response to deviant H in ABHDE, ABCHE, and ABCDH) of the first deviant in the target quintet. Statistical analysis showed no effect of position in the target quintet on the response to the first deviant ($P = 0.94$; $n = 65$; Friedman ANOVA), nor on the difference between the responses to the first deviant and the standard element ($P = 0.56$; $n = 65$; Friedman ANOVA; see Supplementary Fig. 2B). Note that at each position the difference between the

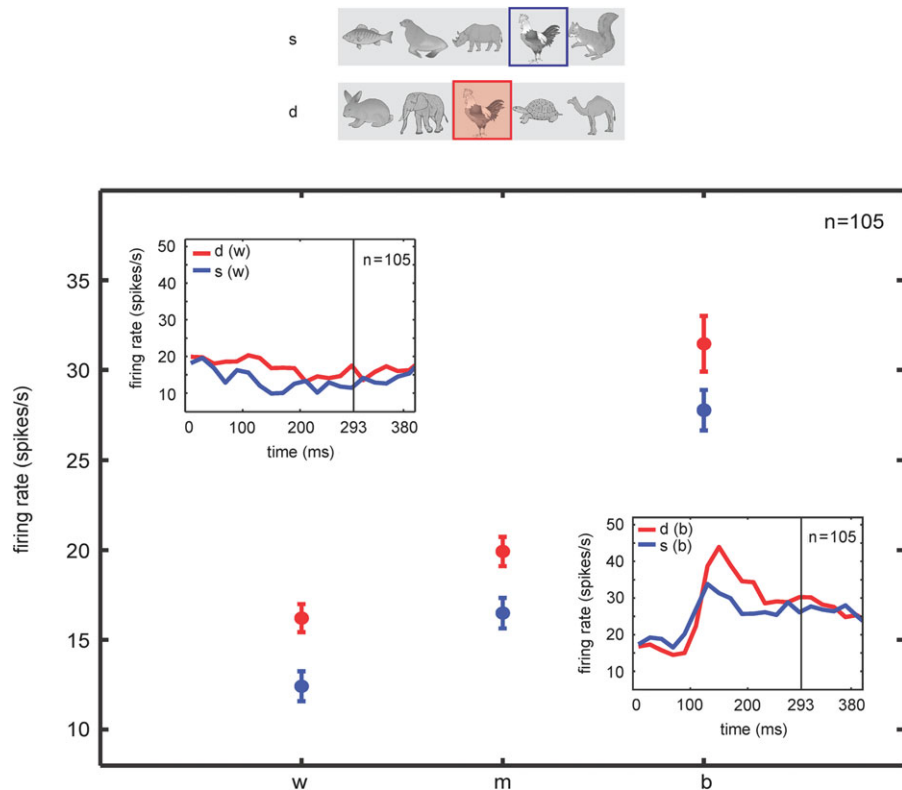


Figure 3. Responses to across-quintet deviants and the corresponding standard elements as a function of the response strength to the standard stimulus. Mean firing rate (105 MUA sites of both monkeys) to the stimuli presented in a standard quintet (blue) and the same stimuli presented as first deviants (red). For each MUA site, the stimuli were sorted according to their response strength for the standard stimuli. w: stimulus with lowest response strength to the standard stimulus; m: stimulus with median response strength to the standard stimulus; b: stimulus with highest response strength to the standard stimulus. The 3 stimuli were selected for each neuron separately using responses for the other half of the presentations than for which we plotted the responses here (see Materials and Methods). Bars indicated standard errors of the mean, computed using the within-subject (= site) design method of Loftus and Masson (1994), that is, after removal of across site mean activity differences. The 2 insets illustrate the population peri-stimulus time histograms for the first deviant (red) and standard element (blue) stimuli for the worst (upper left) and best (bottom right) response conditions. Zero indicates stimulus onset and the vertical line stimulus offset. Data from experiment 1.

response to the first deviant and the standard element was significant (Wilcoxon matched pairs test; FDR corrected q values $< 2.93 \times 10^{-6}$). Also, we examined whether the effect of the first deviant depended on the position of the corresponding standard element in the source quintet. To do this, we analyzed those MUA sites that had first deviants from each of the 5 positions in the source quintet. Although there was a trend toward a smaller response enhancement when the standard element was at the first position of the source quintet, this failed to reach statistical significance ($P = 0.25$; $n = 61$; Friedman ANOVA). Also, there was no significant effect of the position in the source quintet on the difference between the responses for the first deviant and the corresponding standard element ($P = 0.24$; $n = 61$; Friedman ANOVA). Note that at each of the 5 positions the difference between response to the first deviant and the corresponding standard element was significant (Wilcoxon matched pairs test; FDR corrected q values $< 4.1 \times 10^{-4}$; see Supplementary Fig. 2A).

Finally, we examined whether stimulus repetition can explain the enhancement of the responses for the first deviant and the follower standard stimuli. Previous studies suggest that the number of intervening stimuli between repetitions of the same stimulus determines the strength of repetition suppression in IT with stronger repetition suppression with a smaller number of intervening stimuli (McMahon and Olson 2007; Kaliukhovich and Vogels 2014). The insertion of deviants in a sequence of stimuli caused differences between the

distributions of the number of stimuli that are presented between repetitions of the same stimulus of that sequence compared with the standard quintet sequence. We computed for the sequences that were presented to the animals, and contained target quintets, the number of intervening stimuli between the deviant and the preceding presentation of the same stimulus identity and between a corresponding standard element and its preceding presentation. Although the mean number of intervening stimuli was nearly identical for the standard elements (mean = 13.9) and the first deviants (mean = 14.3), the distributions of the number of intervening stimuli differed between the 2 stimulus types (Supplementary Fig. 3A,B). To elaborate, 32% of the standard elements and 16% of the first deviants had stimuli in-between repetitions lesser than 5. However, the distributions of the number of intervening stimuli were highly similar between the standard elements and the follower standard stimuli (Supplementary Fig. 3B,D), although the neural responses differed markedly between these 2 stimulus conditions (Fig. 2B). The same held for the distributions of the number of intervening stimuli and the responses for the first and follower deviant (see for distributions Supplementary Fig. 3A,C; see for neural responses Fig. 2A,C). Hence, differences among the distributions of the number of intervening stimuli for the different stimulus conditions cannot explain the response enhancement present for the first deviant and the follower standard and the lack of enhancement for the follower deviant. In other words, the

response enhancement for the first deviant and follower standard cannot be caused by repetition-related effects, such as repetition suppression. This is not unexpected given the broad and strongly overlapping distributions of the number of intervening stimuli for the different stimulus conditions.

Foils from the Same Quintet

For 82 MUA sites of both monkeys, we compared the response to a standard element with the response to the same stimulus when it replaced another stimulus of the same quintet (foils), thus changing its position within the quintet. We distinguished 2 sorts of foils: “repetition foils” and “nonrepetition foils.” Unlike nonrepetition foils (e.g., “D” in “ADCDE”), repetition foils repeat a stimulus of the same sequence (e.g., “B” in quintet “ABCBE”). Because repetition suppression can occur for repetition foils, which can counteract a response enhancement for the foil, the overall response enhancement may be underestimated for these foils.

We observed a significantly greater response for the nonrepetition foils compared with the same stimuli presented as standard elements ($P = 1.21 \times 10^{-4}$; $n = 82$; Wilcoxon matched pairs test; Fig. 4A). This held true in each of the 2 animals (monkey H: $P = 0.026$, $n = 42$; monkey O: $P = 2.7 \times 10^{-4}$, $n = 40$). A similar response enhancement was present for a follower standard stimulus that trailed a nonrepetition foil ($P = 2.56 \times 10^{-3}$; $n = 82$; Wilcoxon matched pairs test; Fig. 4C) and a repetition foil ($P = 4.02 \times 10^{-4}$; $n = 82$; Wilcoxon matched pairs test; Fig. 4D). Not unexpectedly because of the possibility of repetition suppression due to the small number of intervening stimuli for the repetition foils (mean = 1.4 with 100% smaller than 4; Supplementary Fig. 4B), the response enhancement for the repetition foil failed to reach significance ($P = 0.065$; $n = 82$; Wilcoxon matched pairs test; Fig. 4B). Note that differences among the distributions of the number of intervening stimuli between repetitions (Supplementary Fig. 4) cannot explain the enhancement for the nonrepetition foil and the follower standard stimulus (Fig. 4A,C). In fact, the distribution of the number of intervening stimuli for the nonrepetition foils contained relatively more instances of <4 intervening stimuli than the standard elements, which would predict relatively more repetition suppression (instead of an enhanced response) for the nonrepetition foils. These effects are expected to be weak because of the broad and strongly overlapping distributions of the nonrepetition foils (mean number of intervening stimuli = 12.2) and standard elements (13.8).

These analyses show that a stimulus that is presented at its incorrect position within its quintet produces an enhanced response. In addition, an enhancement is present for a correctly positioned stimulus that follows the incorrectly positioned stimulus. This suggests that the responses in IT may depend on the order of the stimuli, even for pairs of stimuli. To assess whether the order of the stimuli within a pair is critical, we tested whether a response enhancement is present when the order of the members of a manipulated stimulus pair is opposite to that in a standard quintet. For each standard quintet, we had 2 such cases: ABCDE versus ADCDE and ABCDE versus ABEDE, with the relevant pairs indicated in **bold** and the tested stimulus underlined. We found that for these pairs the mean response to the follower standard stimulus that trailed a nonrepetition foil (e.g., C in ADCDE) was significantly greater than to the corresponding standard element (e.g., C in ABCDE; $P = 1.14 \times 10^{-3}$; $n = 82$; Wilcoxon matched pairs test; Fig. 5B), and this response enhancement was present in each animal (monkey H: $P = 0.03$, $n = 42$; monkey O: $P = 6.28 \times 10^{-4}$, $n = 40$).

This demonstrates that the statistical learning signal in IT reflects the order of adjacent stimuli.

Experiment 2

Next we asked whether the greater response for the deviant and foils compared with the standard stimuli corresponds to a suppressed response to the stimuli in standard quintets or an enhanced response to the deviant and foil stimuli, compared with the responses before statistical learning or with stimuli presented in sequences without any regularity. Thus, we exposed the 2 monkeys to 2 groups of 15 stimuli each (Fig. 1B). One group of stimuli consisted of 3 standard quintets, with the stimuli in each quintet having a fixed order. The 3 standard quintets were presented in a random order without any interquintet interval as in experiment 1 (see Materials and Methods). The stimuli of the other group were presented in a random order (“random sequence”). Which group of stimuli was presented as standard quintets or in the random sequence differed between the 2 monkeys. The standard and random sequences were presented in alternating blocks for several weeks and we took care that the familiarity of the standard and random sequence stimuli was equated in each animal (see Materials and Methods).

Comparison of Standard and Random Sequence Responses

After exposing the monkeys to both standard and random sequences, we recorded MUA to both sequences in 119 sites. We observed a significantly larger mean response to the stimuli of the random sequence compared with the stimuli in standard quintets ($P = 0.042$; $n = 119$; Wilcoxon matched pairs test; Fig. 6A). Although this response suppression for the standard quintet stimuli was similar in the population responses of the 2 animals, it was statistically significant only in monkey O when considering the 280-ms analysis window ($P = 1.72 \times 10^{-3}$, $n = 61$; monkey H: $P = 0.68$, $n = 61$; but in monkey H: $P = 0.037$ in bin 100–120 ms after stimulus onset). Because we counterbalanced the 2 groups of stimuli across animals, and since similar trends and time courses were observed in the 2 animals (Fig. 6C,E), it is highly unlikely that the response difference between the 2 sequences was due to differences between the stimuli of the 2 groups. To further address this, we bootstrapped stimuli (with replacement; 1000 bootstraps) and then recomputed the mean response across the resampled stimuli for each bootstrapped sample. We resampled separately the standard and random group stimuli and then we computed the response difference between the 2 stimulus groups per sample pair. For the raw responses, a 0 response difference corresponded to the 1.9th percentile of the distribution of the 1000 bootstrapped response differences while for the normalized responses it corresponded to the 1.5th percentile, both percentiles below 5 (type 1 error = 0.05). This analysis strongly suggests that the smaller response for the standard compared with the random sequence is not due to stimulus differences.

We took care that the standard and random sequences showed on average a nominally identical number of intervening stimuli between repeated stimuli (means of 13.8 and 13.9 for the standard and random sequences, respectively; Supplementary Fig. 5). The smallest number of intervening stimuli between repeated stimuli was 4 by (deliberate) design (see Materials and Methods) in both sequences. The random sequences contained 9% of such repetitions with 4 intervening stimuli compared with 33% in the standard sequence which might have caused stronger repetition

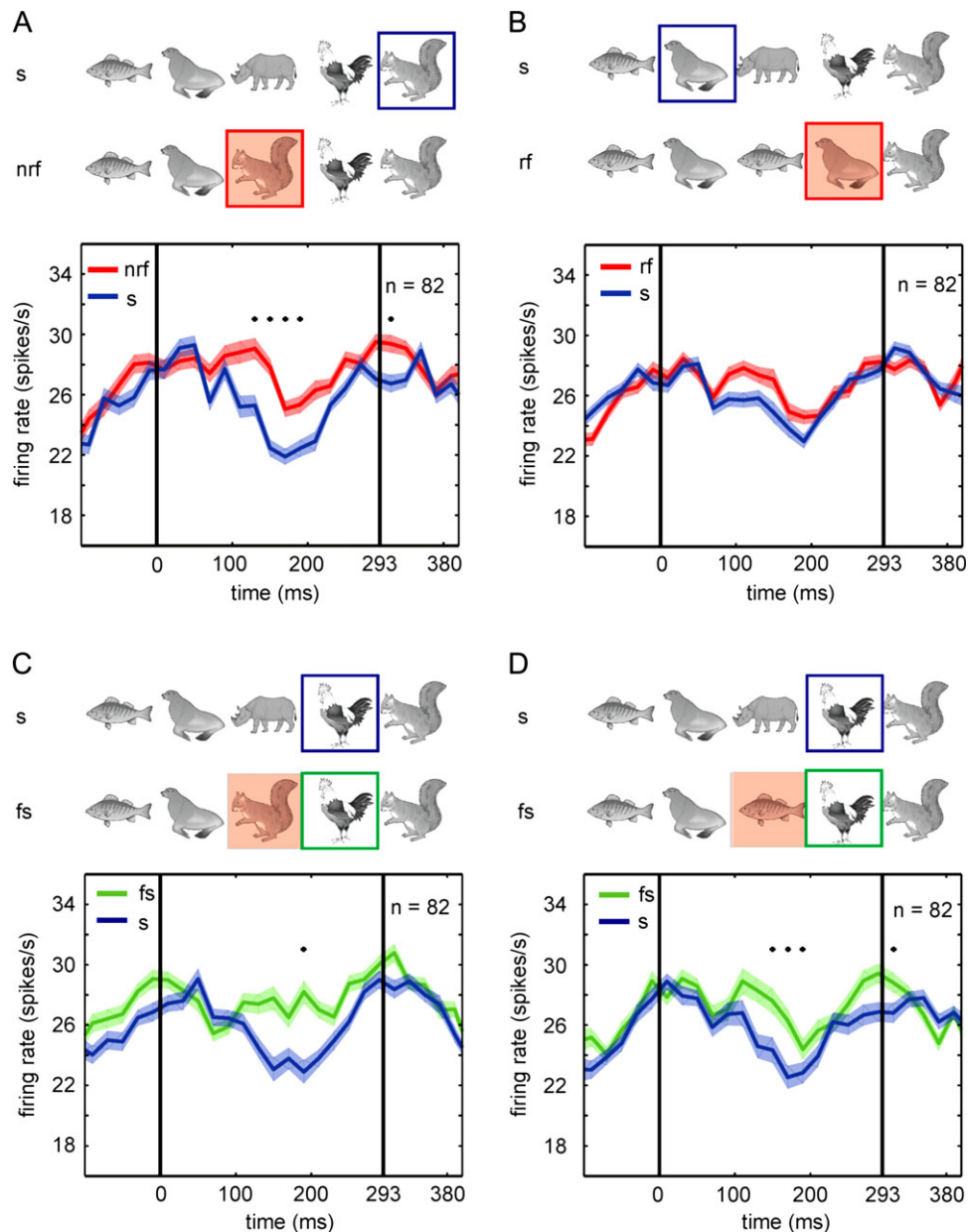


Figure 4. Mean spiking activity for a within-quintet foil and for a follower standard compared with the activity for the corresponding standard elements. (A) Top: illustration of a standard quintet (s) and a target quintet containing a nonrepetition foil (nrf). The stimuli for which the responses are compared are indicated by a colored square, with the blue and red indicating the standard element and the nonrepetition foil, respectively. The shaded square indicates a foil. Bottom: Firing rate averaged across both monkeys ($n = 82$ MUA sites) for a nonrepetition foil (red) and the corresponding standard element (blue). (B) The same as for (A), except that only responses to repetition foils (rf) were considered ($n = 82$ MUA sites). (C) Top: illustration of a standard quintet and a target quintet containing a nonrepetition foil. The stimuli for which the responses are compared are indicated by a colored square, with the blue and green indicating the standard element and the follower standard stimulus, respectively. The shaded square indicates the foil. Bottom: firing rate averaged across both monkeys ($n = 82$ MUA sites) for a follower standard stimulus (green) and the corresponding standard elements (blue). Only responses to a follower standard of nonrepetition foils and responses to their corresponding standard elements are plotted. (D) The same as for (C), except that responses to follower standard stimuli of repetition foils were considered ($n = 82$ MUA sites). Same conventions as in Figure 2. Shaded bands indicate standard error of the mean, computed following the procedure by Loftus and Masson (1994), which removes the variance due to the differences in the overall mean response across MUA sites. Data from experiment 1.

suppression in the standard compared with the random sequence, which may explain the reduced neural response to the stimuli in the former sequences. To address this concern, we verified first whether repetition of stimuli causes repetition suppression in the standard and random sequences. Indeed, the mean response to a stimulus was lower when the fifth preceding stimulus was the same stimulus (e.g., BCDEAB) than when the fifth preceding

stimulus was different (e.g., GHIJAB; Supplementary Fig. 6). This repetition suppression with 4 intervening stimuli was present in the standard ($P = 6.88 \times 10^{-5}$; $n = 119$; Wilcoxon matched pairs test; tested using the 280-ms long analysis window; Supplementary Fig. 6A) and random sequences ($P = 0.002$; $n = 119$; Wilcoxon matched pairs test; Supplementary Fig. 6B). However, the size of the repetition suppression (2 spikes/s at response peak) was

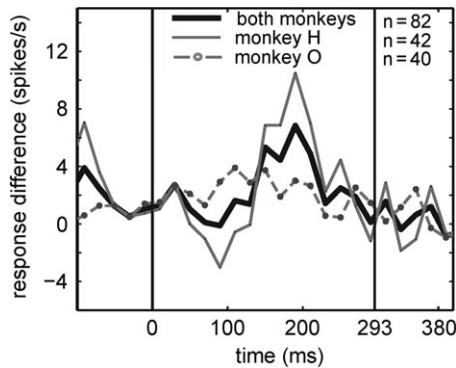


Figure 5. Response differences obtained when the order of neighboring stimuli was reversed with respect to that in the standard quintet. For instance, comparing the response to stimulus M in KNMNO minus response to the stimulus M in KLMNO; see Results. Both the data pooled across animals and for the individual monkeys separately are plotted. Zero on the abscissa indicates stimulus onset and the vertical bars correspond to stimulus onset and offset. Data from experiment 1.

smaller compared with the difference between the random and standard sequences observed in the same neurons (6 spikes/s at response peak). The repetition suppression was not detectable for the standard ($P = 0.27$; $n = 119$; Wilcoxon matched pairs test; Supplementary Fig. 6C) and random sequences ($P = 0.71$; $n = 119$; Wilcoxon matched pairs test; Supplementary Fig. 6D) when 9 intervening stimuli were present between repeated stimuli (note that the number of intervening stimuli for the standard sequence was 4, 9, 14, etc.). To ensure that the difference between the response to the random and standard sequences was not due to repetition suppression, we selected only stimulus presentations for which the number of intervening stimuli between the presentation and the previous presentation of the same stimulus was at least 9, abolishing any possible effect of repetition suppression. The results we obtained with this selected set of presentations were similar to those for all presentations, demonstrating a significantly higher response to the random compared with the standards sequence stimuli in each animal (Fig. 6D,F), and when pooling the 2 animals (Fig. 6B). Hence, the suppressed response to the standard sequence stimuli is not caused by repetition suppression.

Responses to Deviants from Different Quintets

Figure 6 suggests that the suppressed response for the standard quintet stimuli is present for a short duration, till 150 ms at most, whereas the enhanced response to the first deviants in experiment 1 (Fig. 2) was more sustained. We pondered whether this striking difference between the time courses of the suppressed response for the standard quintet stimuli and the enhanced response to the first deviant could be due to uncontrolled differences between experiment 1 and 2. To address this, we presented deviants in the quintet blocks (see Materials and Methods), similar to experiment 1, during MUA recordings in both monkeys (monkey O: 30 sites; monkey H: 32 sites). We replicated the enhanced response to the first deviant compared with its corresponding standard element ($P = 3.09 \times 10^{-4}$; $n = 62$; Wilcoxon matched pairs test; Fig. 7A). As in experiment 1, we also observed a significantly enhanced response for the follower standard ($P = 3.52 \times 10^{-6}$; $n = 62$; Wilcoxon matched pairs test; Fig. 7C). Furthermore, as observed in experiment 1, the

enhanced response to the first deviant was present during the late part of the response, up to 360 ms after stimulus onset (Fig. 7A). Figure 7A also plots the responses to the stimuli of the random sequences for the 62 sites tested with deviants. Although the mean response to the stimuli of the random sequence was higher than to the standard quintet stimuli, this difference failed to reach statistical significance ($P = 0.26$), even when tested binwise. Furthermore, the mean response to the stimuli of the random sequence was significantly smaller than that to the first deviant ($P = 0.017$; $n = 62$; Wilcoxon matched pairs test; Fig. 7A) and the follower standard stimuli ($P = 8.6 \times 10^{-3}$; $n = 62$; Wilcoxon matched pairs test; Fig. 7C). For the standard sequences, the above analyses were performed using the stimuli that were shown as deviants or follower standards for each site. To increase statistical power, we included in a subsequent analysis all 15 stimuli that were shown in the standard quintets, but the difference between the responses for the standard and random sequences still failed to reach significance when integrating the responses in the 280-ms window ($P = 0.92$).

Overall, these data suggest the presence of 2 effects with very different time courses during the response of IT neurons: an early, transient suppression of the response to stimuli shown in standard quintets and a greater sustained enhancement of the response to first deviants. These dissociable effects can be demonstrated clearly when pooling the responses to the first deviants and corresponding standard element presentation of all sites recorded in the quintet sequences of experiment 1 and 2 ($n = 190$ sites; Fig. 7B) and pooling the responses to standard and random sequences of all sites of experiment 2 ($n = 181$ sites; Fig. 7D). Binwise testing (FDR corrected) showed a sustained significant enhancement for the deviant stimulus between 120 and 380 ms after stimulus onset with the maximum deviant response at 150 ms. The response suppression for the standard quintet stimuli, relative to the random sequence stimuli, was present only transiently between 100 and 140 ms, with a maximum response at 110 ms. The MUA of each individual monkey demonstrated this difference in time course between the response suppression for the standard elements and the enhancement for the first deviant (see Supplementary Fig. 7).

The transitional probability from the fifth stimulus of a standard quintet to the first stimulus of the next standard quintet is smaller ($P = 1/3$) than for the last 4 stimuli of a quintet ($P = 1$). We asked whether this difference in transitional probability is sufficiently large to cause a response difference between the first and four other stimuli of a standard quintet. To address this, we pooled the responses of all MUA sites that were recorded in experiments 1 and 2 with standard quintets, thus averaging the responses of a large number of sites across 6 stimuli per position in a quintet per monkey (and 12 across monkeys), reducing spurious differences between positions that can arise because of stimulus (identity) selectivity differences among sites. When pooling the sites ($n = 391$ sites) of the 2 monkeys, we found an effect of position ($P = 1.15 \times 10^{-3}$; $n = 391$; Friedman ANOVA; Fig. 8A) with the greatest response to the first stimulus of a quintet. FDR corrected multiple comparisons with Wilcoxon matched pairs tests showed significant differences between the first and third to fifth positions ($q < 0.021$; corrected for the 10 possible comparisons). The difference between the first and second positions failed to reach significance ($q = 0.067$), and this was because of a difference between the 2 animals in their response to the second position. Monkey O showed a greater response to the first compared with the

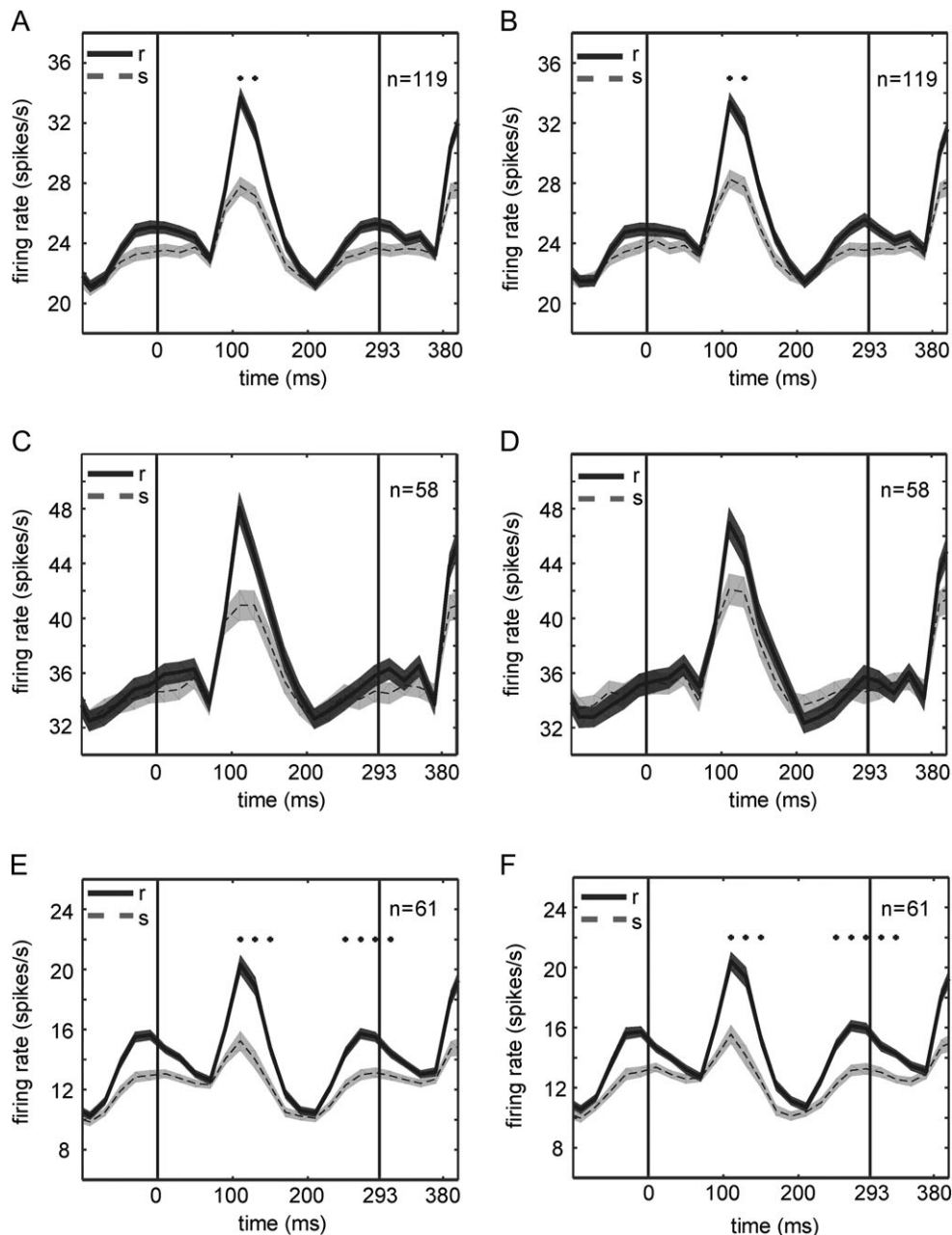


Figure 6. Mean responses to the standard (s) and the random sequence stimuli (r). (A) Data pooled across the 2 animals. (C, E) Data of monkeys H and O, respectively. (B) Data pooled across the 2 animals for the stimulus presentations for which the number of intervening stimuli between the presentation and the previous presentation of the same stimulus was at least 9, controlling for repetition suppression. (D, F) Data with at least 9 intervening stimuli of monkeys H and O, respectively. Note that the standard and random sequence stimuli were counterbalanced across the animals. Same conventions as in Figure 2. Shaded bands indicate standard error of the mean, computed following the procedure by Loftus and Masson (1994). Data from experiment 2 including only standard quintets, obtained before inserting deviants into target sequences.

other stimulus positions ($P = 0.011$; $n = 197$; Friedman ANOVA; Fig. 8C), as predicted by the transitional probabilities. However, we found no statistical evidence for a difference between the first and second positions for monkey H ($q = 0.82$), despite an overall statistically significant effect of position in this animal ($P = 4.61 \times 10^{-3}$; $n = 194$; Friedman ANOVA; Fig. 8B). It is possible that monkey H did not learn the transitional probability of the first stimulus pair of the quintets.

Given this effect of within-quintet stimulus position on response strength, we recomputed the average responses to the standard and random sequences by including only the

third to fifth stimuli of a standard quintet and randomly drawing 9 of the 15 stimuli presented in the random sequences. As expected from the within-quintet position effect (Fig. 8), the average response to the stimuli presented in the standard quintets of experiment 2 (Fig. 7D; stippled line) dropped slightly but this drop was small compared with the size of the response difference between the random and standard sequences. Also, this analysis showed that the response to the standard quintet stimuli was well below the 95% confidence interval obtained by randomly drawing (without replacement) 1000 times 9 of the 15 random sequence stimuli (gray band in

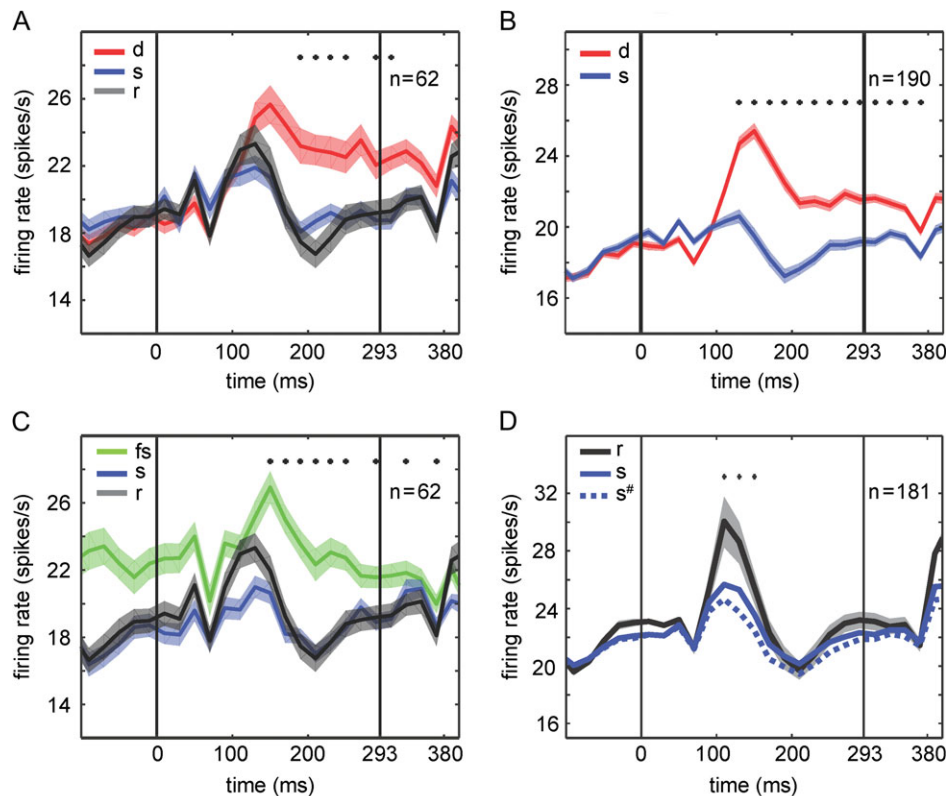


Figure 7. Comparison of responses to first deviants, to follower standards, to their corresponding standard elements and to responses to random sequence stimuli. (A) Comparison of responses to first deviants (red), standard elements (blue) and stimuli in random sequences (gray) for the same multi-units ($n = 62$). Data of experiment 2 of both monkeys were pooled—see Supplementary Figure 7 for data of the individual monkeys. (B) Responses to the first deviants (red) and standard elements (blue) for the combined data of experiments 1 and 2 (190 multi-unit sites, pooled across animals). Shaded bands in (A), (B) and (C) indicate standard error of the mean, computed following the procedure by Loftus and Masson (1994). (C) Comparison of responses to first deviants (green), standard elements (blue) and stimuli in random sequences (gray) for the same multi-units. Data of both monkeys were pooled—see Supplementary Figure 7 for data of the individual monkeys. The stars indicate significant differences between the first deviant or the follower standard and the stimulus in the random sequence (Wilcoxon matched pairs test; FDR corrected $q < 0.05$), tested per 20 ms bin. For other conventions, see Figure 2. Data were collected after those of Figure 6 (see Materials and Methods). (D) Responses to all standard quintet stimuli (15 per monkey; “s,” thick blue line) and random sequence stimuli (15 per monkey; “r,” thick black line) of 181 multi-unit sites of experiment 2 (combining data from Figs 6A and 7A). The stippled blue line (“s#”) shows the mean response to the last 3 stimuli of a standard quintet (9 stimuli in total per monkey). The gray band shows the 95% confidence interval of the mean response to 9 randomly drawn stimuli (out of 15) of the random sequence. The stars indicate significant differences between the mean response for the standard quintet stimuli and random sequence stimuli (Wilcoxon matched pairs test; FDR corrected $q < 0.05$), tested per 20 ms bin. These significance tests were performed using the response for the last 3 stimuli of the 3 standard quintets and the response for 9 randomly drawn stimuli of the random sequence.

Fig. 7D). These differences in responses between the standard and random sequence stimuli remained present when controlling for repetition suppression. We controlled for repetition suppression by taking only the presentations of 9 stimuli in the standard and random sequences for which the number of intervening stimuli between the presentation and the previous presentation of the same stimulus was at least 9 (see above; Supplementary Fig. 8).

Discussion

Passive exposure to an uninterrupted sequence of visual stimuli is known to induce learning of the statistical temporal regularities embedded in the sequence. Here we show that exposing monkeys to continuous sequences with a temporal regularity affects the responses of IT neurons. We found that the early phase of the response was reduced for stimuli that belonged to a sequence with a temporal regularity as compared with responses for stimuli that were presented in a random order during the exposure. Importantly, this response reduction

for stimuli of a sequence with a temporal regularity was not caused by stimulus familiarity since familiarity was equated between the random and standard sequences. Also, it remained present when controlling for repetition effects on the response. A second, sustained statistical learning signature was an enhanced response to stimuli that violated the fixed short sequence or were presented in the wrong order. This enhancement was present for stimuli that were not expected to follow the previous stimulus, irrespective of whether the trailing stimulus was a deviant (first deviant) or was present at the correct position within the standard quintet but followed a deviant (follower standard), suggesting that the statistical learning signal in IT is mainly based on temporally adjacent stimuli. These novel data not only demonstrate statistical learning signals in IT but also advances our understanding of the nature of these signals, that is, suppression of the response to learned sequences versus enhancement of the response to stimuli that violated a learned sequence.

Behavioral and fMRI studies have demonstrated that non-human primates learn statistical regularities in auditory

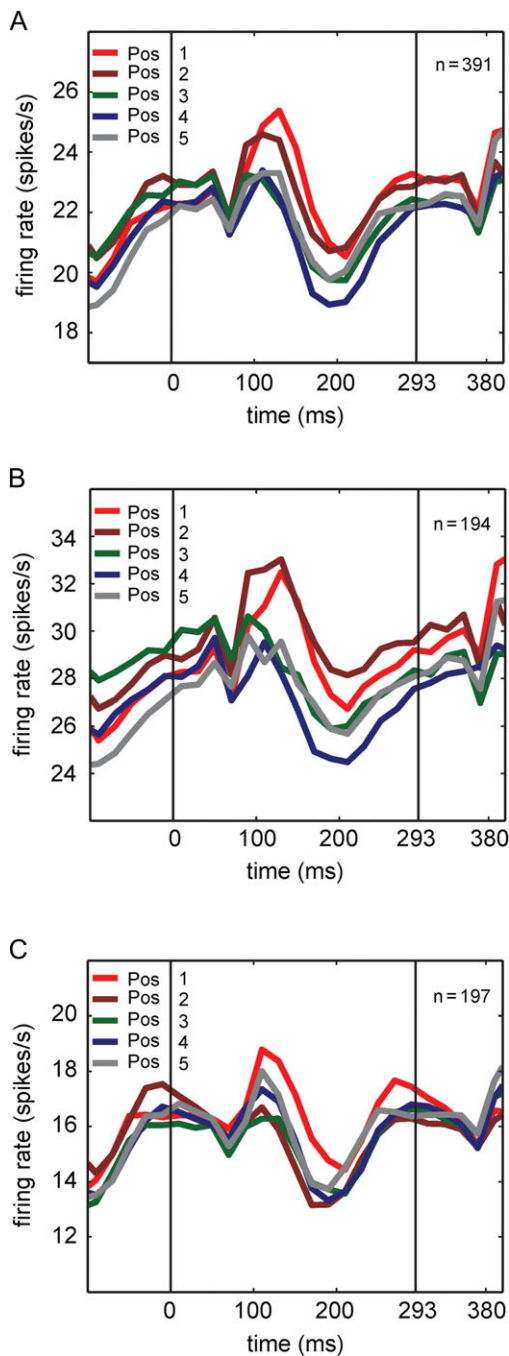


Figure 8. Averaged responses to the stimuli of standard quintets, sorted according to the position of the stimulus within a quintet. Data of experiments 1 and 2 were combined. Pos 1 indicates mean responses to the first stimulus of a quintet, Pos 2 to the second stimulus of a quintet, etc. (A) Data pooled across animals. (B, C) Data of monkeys H and O, respectively. Note that stimuli differed between the 2 animals. Number of multi-unit sites (n) is indicated in each panel.

sequences to which they were passively exposed (marmosets, Wilson et al. 2013; rhesus monkeys, Wilson et al. 2013, 2015; Wang et al. 2015; squirrel monkeys, Ravignani et al. 2013). Interestingly, Wilson et al. (2015) found that the behavioral responses were sensitive to the relationship of temporally adjacent stimuli but not of nonadjacent stimuli. This agrees well with our finding that the response to deviants was driven by temporally adjacent stimuli (also see Meyers et al. 2014). This

also concurs with behavioral studies in humans that show learning of nonadjacent regularities under specific conditions only, for instance when adjacent dependencies are weak (Gómez 2002; Newport and Aslin 2004; Perruchet and Pacton 2006; Bays et al. 2015; but see Otsuka and Saiki 2016). Thus, it is possible that macaques do not learn nonadjacent temporal regularities during passive exposure to sequences with strong adjacent regularities and this was reflected in the IT responses.

To the best of our knowledge, our study is the first to show that macaques learn temporal regularities when being passively exposed to uninterrupted sequences consisting of randomly presented shorter structured sequences of visual stimuli. Previous studies either employed short sequences presented discretely with notable intersequence intervals (auditory stimuli, Wilson et al. 2013, 2015; Wang et al. 2015) and a reward following each short sequence (visual stimuli; Meyers et al. 2014) or trained explicitly baboons to categorize visual “words” from “nonwords” (Grainger et al. 2012). This opens up the possibility to perform comparative studies of statistical learning in monkeys and humans, using the same paradigm that has been employed in human behavioral and imaging studies.

The smaller spiking activity in monkey IT for the standard compared with the random sequences contrasts with the stronger BOLD activation in human LOC for structured compared with sequences without statistical regularity (Turk-Browne et al. 2009). The extent of exposure to the sequences in our monkey study was longer than that employed typically in human studies (e.g., a single session in the Turk-Browne et al. 2009 study) of statistical learning. We did not know the time course of the statistical learning effects in monkeys and we tried to maximize the likelihood of finding a statistical learning effect in our animals by employing a long exposure. These differences in exposure duration may explain differences between the fMRI findings and our monkey data. We would like to note that in real-life people can also have extensive exposure to some stimulus sequences, for example, when driving daily for years along the same road, and perhaps the human statistical learning studies only reflect mechanisms that are related to the initial learning/exposure phase. In this regard, a recent human psychophysical study (Bays et al. 2015) suggested that different behavioral measures can reflect dissociable mechanisms of statistical learning and the same may be true after different exposure durations. Interestingly, a reversal of a perceptual learning effect in BOLD activation has been reported in the primary visual cortex of human subjects during extensive training in a visual texture discrimination task (Yotsumoto et al. 2008).

MUA is a population measure of spiking activity, pooling the activity of several neurons close to the electrode tip. Previous learning studies in IT have also employed MUA (Li and DiCarlo 2010, 2012). MUA correlates with the stimulus-driven component of the BOLD response (Logothetis 2003), in V1 even better than local field potentials according to one report (Cardoso et al. 2012). However, task-related, anticipatory BOLD responses may correlate less with MUA (Sirotnin and Das 2009) and spiking activity can be dissociated from BOLD under particular perceptual conditions (Maier et al. 2008). A dissociation between BOLD and MUA is another possible cause of discrepancy between the results of our study and the fMRI study of Turk-Browne et al. (2009).

Our data diverge from the stronger neural response in the primary visual cortex of mice for a fixed sequence, consisting of 4 oriented gratings, compared with random or novel sequences of the same gratings (Gavornik and Bear 2014). The discrepancy between our monkey IT and the mice visual cortex data can be caused by species differences, a difference between early and

late visual areas, exposure duration (up to 5 days in the mice study, with a learning effect after 1 day [2 min of exposure]), or due to combinatorial effects of these 3 possibilities.

Our findings agree with those obtained by the Meyer and Olson studies (Meyer and Olson 2011; Meyer et al. 2014; Ramachandran et al. 2016) who exposed monkeys for several weeks to doublets or triplets of visual images. They found a greater IT response to visual images when these stimuli were presented in another doublet or triplet than the ones the monkeys were exposed to, similar to our first deviant condition. Similar to our results obtained with an exposure to continuous sequences, the learning signal reflected mainly adjacent stimuli. In their study, the doublets or the triplets were presented during the exposure and test phases as discrete units, each separated by an intersequence interval, and each individual short sequence was followed by a reward. However, an essential requirement of statistical learning is that the stimulus regularities are demarcated only by the statistics of the input stream without any other segmentation cue (Turk-Browne et al. 2009), as was the case in the present study but not in the Meyer and Olson studies. Furthermore, it is uncertain whether the reward contingency in the Meyer and Olson studies contributed to the learning effects observed in those studies. We can safely exclude the reward factor as playing a causative role in the learning in our study since reward delivery was orthogonal to the continuous stream of quintets.

Meyer et al. (2014) and Ramachandran et al. (2016) described their difference in response between the deviants and standard members of a stimulus pair as reflecting “prediction suppression.” The idea was that the monkeys or IT neurons learn to predict the trailing stimulus and that when the prediction equals the presented stimulus the responses of the neurons are reduced. However, it is impossible to know from those studies whether the response to the expected trailing stimulus decreased or whether instead the response to the unexpected deviant stimulus increased, both relative to a condition where no prediction is present. We dissociated these 2 possibilities by introducing the random sequence condition in experiment 2. We found evidence for both types of effects: a relatively brief suppression of the response in the standard compared with the random sequence and a later, more sustained enhanced response to the deviant compared with the random sequence stimuli. As far as the effects obtained in our statistical learning paradigm with continuous presentation translate to the effects obtained when using reward-contingent discrete presentations, “prediction suppression” reflects only part of the difference in response between the deviant and the expected trailing stimulus in the Meyer and Olson studies. In fact, our data suggest that the enhancement of the response to the deviant is a more dominant factor than the suppression of the response to predicted stimuli when comparing deviant and predicted standard stimulus responses.

Temporal cortical neurons are known to associate temporally contiguous stimuli (Miyashita 1988; Sakai and Miyashita 1991; Erickson and Desimone 1999; Li and DiCarlo 2008, 2010) and this association learning is independent of reward (Li and DiCarlo 2012). For instance, in the context of a delayed matching to sample task, Miyashita (1988) observed that the stimuli to which an anterior ventral IT neuron was selective during the delay period depended on the fixed order in which the sample stimuli were presented during the training of the task. He found that the neurons could associate up to 6 consecutive stimuli. These results appear to contrast with the present data in which the learning signal reflected stimulus pairs. However,

there are several important differences between the Miyashita study and ours that can explain this difference. First, the stimulus exposure was passive in our case while the monkeys were performing a memory task in the Miyashita study. Second, the stimuli in our study were presented in a continuous stream while they were presented in subsequent trials in the Miyashita experiment. Third, our recordings were in the STS while Miyashita’s recordings were in ventral temporal cortex (TEav and perirhinal cortex). It is possible that anterior ventral temporal cortical neurons associate a larger set of stimuli than STS neurons, since anterior ventral temporal cortex is heavily connected with medial temporal structures (Kravitz et al. 2013) which have been implicated in statistical learning (Schapiro et al. 2014).

Fourth, it is important to note that Miyashita (1988) examined the stimulus selectivity of single neurons, finding similar responses to stimuli that occurred contiguously during the training. Instead, we compared the responses to the same stimuli when these were presented in the correct versus incorrect position in a sequence, irrespective of the stimulus selectivity of the neuron. It is expected that an associative learning mechanism in which the selectivity of a single neuron reflects temporal contiguous stimuli (Miyashita 1988; Sakai and Miyashita 1991; Erickson and Desimone 1999; Li and DiCarlo 2008, 2010) was present in our experiments too, resulting in neurons that tend to respond similarly to temporally adjacent stimuli of the standard quintets. One could envisage that neurons that respond similarly to adjacent stimuli would show a reduction of the response in our standard sequences, compared with deviants or random sequences, because of adaptation-induced suppression following presentation of the previous stimulus. In such a scheme, neurons that are selective for both, for example, stimulus A and B of a quintet will respond less to stimulus B when it follows A since the neuron’s response became adapted upon the presentation of stimulus A. However, an associative coding mechanism, as described in the Miyashita studies, implies bidirectional associations: neurons coding the association of A and B should respond to A and B, irrespective of the order of presentation. Such a bidirectional mechanism cannot explain our observation that after the exposure the neurons were sensitive to the order of the members of a pair of stimuli (AB vs. BA). This implies that the statistical learning signal we uncovered in IT is unidirectional or order sensitive, that is, it reflects transitional statistics and not just joint statistics.

Our finding of a transitional statistics signal in IT is in line with predictive coding theories of cognition (Rao and Ballard 1999; Friston 2005; Summerfield and de Lange 2014). In such predictive coding framework, the monkey’s brain learned to predict the stimuli of the quintets and compared the prediction with the presented stimulus. When the predicted stimulus equaled the presented one, no prediction error was present that would produce little response in the cortical neurons that reflect the prediction error, thus explaining the smaller response to the fixed (correct prediction) compared with the random sequence (no prediction). Thus, the latter effect can be viewed as a case of “expectation suppression” (Todorovic and de Lange 2012; Summerfield and de Lange 2014). In the case of the first deviant (or the follower standard stimulus trailing a first deviant), the prediction error is large, explaining the enhanced response to the first deviants, which can be viewed as a “surprise response” (Kaliukhovich and Vogels 2014). Our data suggest that expectation suppression and the surprise response are dissociated in their time course and hence may reflect different mechanisms. Given the early onset of the

expectation suppression, it is possible that this signal originates in the visual cortex. Hebbian plasticity, adapting connections in the local network of neurons that each prefer successive stimuli of a standard quintet, may underlie expectation suppression, as suggested by computational models of sequence learning (Wang and Arbib 1990; Abbott and Blum 1996; Veliz-Cuba et al. 2015; Ramachandran et al. 2016). The longer delay of the surprise response to the first deviants suggests that it may result from a top-down input from areas that anticipate a stimulus and detect the violation of the expectation. It is possible that prediction signals from areas such as prefrontal cortex or hippocampus (Davachi and DuBrow 2015) also assist the development of statistical learning-related response changes in IT during the exposure phase (Turk-Browne et al. 2010). Conversely, the violation of the expected stimulus might be computed locally in IT with the longer and more sustained “surprise” response reflecting recurrent activity in the local IT network.

Conclusion

We showed that IT spiking activity signal violations of a learned statistical regularity, defined by transitional probabilities, in a continuous sequence of visual stimuli. The statistical learning was completely unsupervised and stimulus presentation was independent of reward. We uncovered 2 signals related to statistical learning in IT: an early response suppression to stimuli of a sequence with statistical regularities compared with those of a random sequence, and a more dominant sustained but later enhancement of the response to stimuli that violated the learned sequence. These dissociable effects suggest that theories that rely on unitary predictive signals need to be adapted and that likely multiple signals underlie sequence learning and responses to stimuli that violate a learned sequence. The possibility that different mechanisms may underlie statistical learning is supported by a recent behavioral study in humans showing dissociable behavioral read-outs of statistical learning (Bays et al. 2015). Further work is needed to determine the origin of these statistical learning signals we uncovered here in macaque IT.

Supplementary Material

Supplementary material can be found at: <http://www.cercor.oxfordjournals.org/>.

Funding

Fonds voor Wetenschappelijk Onderzoek Vlaanderen (G.0582.12N and G.00007.12-Odysseus), Interuniversitaire Attractiepool and Programma Financiering (PF 10/008), and the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement number PITN-GA-2008-290011 (ABC).

Notes

We thank P. Kayenbergh, G. Meulemans, I. Puttemans, Christophe Ulens, M. De Paep, W. Depuydt, and S. Verstraeten for technical assistance. *Conflict of Interest*: None declared.

References

- Abbott LF, Blum KI. 1996. Functional significance of long-term potentiation for sequence learning and prediction. *Cereb Cortex*. 6:406–416.
- Baker CI, Behrmann M, Olson CR. 2002. Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nat Neurosci*. 5:1210–1216.
- Bays B, Turk-Browne N, Seitz A. 2015. Dissociable behavioural outcomes of visual statistical learning. *Vis Cogn*. 23: 1072–1097.
- Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol*. 57:289–300.
- Cardoso M, Sirotin Y, Lima B, Glushenkova E, Das A. 2012. The neuroimaging signal is a linear sum of neurally distinct stimulus- and task-related components. *Nat Neurosci*. 15: 1298–1306.
- Davachi L, DuBrow S. 2015. How the hippocampus preserves order: the role of prediction and context. *Trends Cogn Sci*. 19:92–99.
- De Baene W, Premereur E, Vogels R. 2007. Properties of shape tuning of macaque inferior temporal neurons examined using rapid serial visual presentation. *J Neurophysiol*. 97: 2900–2916.
- Dehaene S, Meyniel F, Wacogne C, Wang L, Pallier C. 2015. The neural representation of sequences: from transition probabilities to algebraic patterns and linguistic trees. *Neuron*. 88:2–19.
- Denys K, Vanduffel W, Fize D, Nelissen K, Peuskens H, Van Essen D, Orban GA. 2004. The processing of visual shape in the cerebral cortex of human and nonhuman primates: a functional magnetic resonance imaging study. *J Neurosci*. 24:2551–2565.
- Erickson CA, Desimone R. 1999. Responses of macaque perirhinal neurons during and after visual stimulus association learning. *J Neurosci*. 19:10404–10416.
- Fiser J, Aslin RN. 2001. Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychol Sci*. 12:499–504.
- Fiser J, Aslin RN. 2002. Statistical learning of higher-order temporal structure from visual shape sequences. *J Exp Psychol Learn Mem Cogn*. 28:458–467.
- Freedman DJ, Riesenhuber M, Poggio T, Miller EK. 2006. Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. *Cereb Cortex*. 16:1631–1644.
- Friston K. 2005. A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci*. 360:815–836.
- Gavornik JP, Bear MF. 2014. Learned spatiotemporal sequence recognition and prediction in primary visual cortex. *Nat Neurosci*. 17:732–737.
- Gómez RL. 2002. Variability and detection of invariant structure. *Psychol Sci*. 13:431–436.
- Grainger J, Dufau S, Montant M, Ziegler JC, Fagot J. 2012. Orthographic processing in baboons (*Papio papio*). *Science*. 336:245–248.
- Kaliukhovich DA, Vogels R. 2014. Neurons in macaque inferior temporal cortex show no surprise response to deviants in visual oddball sequences. *J Neurosci*. 34: 12801–12815.
- Kobatake E, Tanaka K. 1994. Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *J Neurophysiol*. 71:856–867.

- Kravitz DJ, Saleem KS, Baker CI, Ungerleider LG, Mishkin M. 2013. The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends Cogn Sci.* 17:26–49.
- Krogh L, Vlach HA, Johnson SP. 2012. Statistical learning across development: flexible yet constrained. *Front Psychol.* 3:598.
- Li N, DiCarlo JJ. 2008. Unsupervised natural experience rapidly alters invariant object representation in visual cortex. *Science.* 321:1502–1507.
- Li N, DiCarlo JJ. 2010. Unsupervised natural visual experience rapidly reshapes size-invariant object representation in inferior temporal cortex. *Neuron.* 67:1062–1075.
- Li N, DiCarlo JJ. 2012. Neuronal learning of invariant object representation in the ventral visual stream is not dependent on reward. *J Neurosci.* 32:6611–6620.
- Loftus GR, Masson ME. 1994. Using confidence intervals in within-subject designs. *Psychon Bull Rev.* 1:476–490.
- Logothetis NK. 2003. The underpinnings of the BOLD functional magnetic resonance imaging signal. *J Neurosci.* 23:3963–3971.
- Maier A, Wilke M, Aura C, Zhu C, Ye F, Leopold D. 2008. Divergence of fMRI and neural signals in V1 during perceptual suppression in the awake monkey. *Nat Neurosci.* 11:1193–1200.
- McMahon D, Olson C. 2007. Repetition suppression in monkey inferotemporal cortex: relation to behavioral priming. *J Neurophysiol.* 97:3532–3543.
- Meyer T, Olson CR. 2011. Statistical learning of visual transitions in monkey inferotemporal cortex. *Proc Natl Acad Sci USA.* 108:19401–19406.
- Meyer T, Ramachandran S, Olson CR. 2014. Statistical learning of serial visual transitions by neurons in monkey inferotemporal cortex. *J Neurosci.* 34:9332–9337.
- Miller E, Li L, Desimone R. 1991. A neural mechanism for working and recognition memory in inferior temporal cortex. *Science.* 254:1377–1379.
- Miyashita Y. 1988. Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature.* 335:817–820.
- Miyashita Y, Date A, Okuno H. 1993. Configurational encoding of complex visual forms by single neurons of monkey temporal cortex. *Neuropsychologia.* 31:1119–1131.
- Newport EL, Aslin RN. 2004. Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cogn Psychol.* 48:127–162.
- Op de Beek HP, Wagemans J, Vogels R. 2007. Effects of perceptual learning in visual backward masking on the responses of macaque inferior temporal neurons. *Neuroscience.* 145:775–789.
- Otsuka S, Saiki J. 2016. Gift from statistical learning: visual statistical learning enhances memory for sequence elements and impairs memory for items that disrupt regularities. *Cognition.* 147:113–126.
- Perruchet P, Pacton S. 2006. Implicit learning and statistical learning: one phenomenon, two approaches. *Trends Cogn Sci.* 10:233–238.
- Ramachandran S, Meyer T, Olson CR. 2016. Prediction suppression in monkey inferotemporal cortex depends on the conditional probability between images. *J Neurophysiol.* 115:355–362.
- Rao RP, Ballard DH. 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci.* 2:79–87.
- Ravignani A, Sonnweber R-S, Stobbe N, Fitch WT. 2013. Action at a distance: dependency sensitivity in a new world primate. *Biol Lett.* 9:20130852.
- Rossion B, Pourtois G. 2004. Revisiting Snodgrass and Vanderwart's object pictorial set: the role of surface detail in basic-level object recognition. *Perception.* 33:217–236.
- Saffran JR, Aslin RN, Newport EL. 1996. Statistical learning by 8-month-old infants. *Science.* 274:1926–1928.
- Sakai K, Miyashita Y. 1991. Neural organization for the long-term memory of paired associates. *Nature.* 354:152–155.
- Sawamura H, Orban G, Vogels R. 2006. Selectivity of neuronal adaptation does not match response selectivity: a single-cell study of the fMRI adaptation paradigm. *Neuron.* 49:307–318.
- Schapiro AC, Gregory E, Landau B, McCloskey M, Turk-Browne NB. 2014. The necessity of the medial temporal lobe for statistical learning. *J Cogn Neurosci.* 26:1736–1747.
- Sirotin Y, Das A. 2009. Anticipatory haemodynamic signals in sensory cortex not predicted by local neuronal activity. *Nature.* 457:475–479.
- Summerfield C, de Lange FP. 2014. Expectation in perceptual decision making: neural and computational mechanisms. *Nat Rev Neurosci.* 15:745–756.
- Tanaka K. 1996. Inferotemporal cortex and object vision. *Annu Rev Neurosci.* 19:109–139.
- Todorovic A, de Lange FP. 2012. Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *J Neurosci.* 32:13389–13395.
- Turk-Browne NB, Scholl BJ, Chun MM, Johnson MK. 2009. Neural evidence of statistical learning: efficient detection of visual regularities without awareness. *J Cogn Neurosci.* 21:1934–1945.
- Turk-Browne N, Scholl B, Johnson M, Chun M. 2010. Implicit Perceptual Anticipation Triggered by Statistical Learning. *J Neurosci.* 30:11177–11187.
- Turk-Browne NB. 2012. Statistical learning and its consequences. In: Dodd MD, Flowers JH, editors. *The influence of attention, learning, and motivation on visual search.* New York: Springer. p. 117–146.
- Veliz-Cuba A, Shouval HZ, Josić K, Kilpatrick ZP. 2015. Networks that learn the precise timing of event sequences. *J Comput Neurosci.* 39:235–254.
- Wang D, Arbib MA. 1990. Complex temporal sequence learning based on short-term memory. *Proc IEEE.* 78:1536–1543.
- Wang L, Uhrig L, Jarraya B, Dehaene S. 2015. Representation of numerical and sequential patterns in macaque and human brains. *Curr Biol.* 25:1966–1974.
- Wilson B, Slater H, Kikuchi Y, Milne AE, Marslen-Wilson WD, Smith K, Petkov CI. 2013. Auditory artificial grammar learning in macaque and marmoset monkeys. *J Neurosci.* 33:18825–18835.
- Wilson B, Smith K, Petkov CI. 2015. Mixed-complexity artificial grammar learning in humans and macaque monkeys: evaluating learning strategies. *Eur J Neurosci.* 41:568–578.
- Woloszyn L, Sheinberg DL. 2012. Effects of long-term visual experience on responses of distinct classes of single units in inferior temporal cortex. *Neuron.* 74:193–205.
- Yotsumoto Y, Watanabe T, Sasaki Y. 2008. Different dynamics of performance and brain activation in the time course of perceptual learning. *Neuron.* 57:827–833.