



**Cite this article:** Southwell R, Baumann A, Gal C, Barascud N, Friston K, Chait M. 2017 Is predictability salient? A study of attentional capture by auditory patterns. *Phil. Trans. R. Soc. B* **372**: 20160105. <http://dx.doi.org/10.1098/rstb.2016.0105>

Accepted: 28 July 2016

One contribution of 15 to a theme issue 'Auditory and visual scene analysis'.

#### Subject Areas:

neuroscience, behaviour, cognition

#### Keywords:

predictive coding, electroencephalography, attention, statistical learning, regularity, auditory scene analysis

#### Author for correspondence:

Maria Chait  
e-mail: [m.chait@ucl.ac.uk](mailto:m.chait@ucl.ac.uk)

<sup>†</sup>These authors contributed equally to this study.

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.fig-share.c.3583490>.

# Is predictability salient? A study of attentional capture by auditory patterns

Rosy Southwell<sup>1,†</sup>, Anna Baumann<sup>1,†</sup>, Cécile Gal<sup>1,†</sup>, Nicolas Barascud<sup>2</sup>, Karl Friston<sup>3</sup> and Maria Chait<sup>1</sup>

<sup>1</sup>Ear Institute, University College London, London WC1X 8EE, UK

<sup>2</sup>École Normale Supérieure, Paris 75005, France

<sup>3</sup>Wellcome Trust Centre for Neuroimaging, University College London, London WC1N 3BG, UK

MC, 0000-0002-7808-3593

In this series of behavioural and electroencephalography (EEG) experiments, we investigate the extent to which repeating patterns of sounds capture attention. Work in the visual domain has revealed attentional capture by statistically predictable stimuli, consistent with predictive coding accounts which suggest that attention is drawn to sensory regularities. Here, stimuli comprised rapid sequences of tone pips, arranged in regular (REG) or random (RAND) patterns. EEG data demonstrate that the brain rapidly recognizes predictable patterns manifested as a rapid increase in responses to REG relative to RAND sequences. This increase is reminiscent of the increase in gain on neural responses to attended stimuli often seen in the neuroimaging literature, and thus consistent with the hypothesis that predictable sequences draw attention. To study potential attentional capture by auditory regularities, we used REG and RAND sequences in two different behavioural tasks designed to reveal effects of attentional capture by regularity. Overall, the pattern of results suggests that regularity does not capture attention.

This article is part of the themed issue 'Auditory and visual scene analysis'.

## 1. Introduction

The human brain is highly sensitive to patterns in sensory input [1–5]. A growing body of work in vision [3,6], touch [7], language [1] and audition [8–13] has demonstrated that subjects rapidly and automatically learn complex sensory statistics, and that these are exploited to improve perceptual inference, even when outside conscious awareness. This capacity is often interpreted as a fundamental element of the predictive mechanisms, which are proposed to constitute the principal substrate of perception [14–16].

In hearing, automatic sequence learning has commonly been studied via the mismatch negativity (MMN), an electrophysiological marker for the processing of sounds that break an established rule [11]. MMN to sequence violations has provided (indirect) evidence that the auditory system can learn complex rules governing sequences [10,17]. The repetition positivity, which increases with the number of repeated stimuli, is another neural marker of simple regularity extraction [18]. Recently, Barascud *et al.* [19] provided direct evidence of the process of regularity extraction in more complex tone sequences. They used rapid sequences of tones with frequencies changing in a regular, cyclical pattern, and matched sequences of tones arranged in a random order. Behavioural reaction times (RT) and neural response dynamics indicated rapid recognition of regularity, on par with the latency predicted from an ideal observer model.

Learned knowledge about regularities, whether from low-level statistical learning or conceptual understanding of the phenomena causing sounds, enables predictions to be formed about future sensory input [20]. Such expectations improve behavioural performance in predictable contexts; for example, by orienting resources to a point in time when a stimulus is expected [21], or by facilitating

selective attention and segregation of concurrent sound streams [10,22–24]. In addition, recognition of regularities can aid detection of changes in the environment, which causes sensory input that is in disagreement with these predictions [9,25,26]. It has been proposed that the same predictive mechanisms underlie both the detection of regularity violations and auditory scene analysis [20,25].

Attention allows the prioritization of useful streams of information for further processing. Within this context, the relationship between predictability and attention is attracting increasing research interest [27–31]. However, there is disagreement as to whether it is unpredictable, surprising events [32–34] or predictable ones [35] that are the most informative in scene analysis, and should therefore (in the parlance of Itti & Koch [36,37]) be flagged as more ‘salient’ and attract selective attention (see also [38–40]). In visual studies, it has been shown that learning of regularities helps guide attention to expected locations [41,42] and features [41,43]. Zhao *et al.* [44] recently proposed a framework for attentional guidance whereby automatically learned regularities in the sensorium bias attention, even if not relevant for performing a task, and demonstrated this to operate in guiding visual search. They presented sequences of abstract shapes; the order of which was statistically structured at a particular location in the search array and random at others. This was followed by a static visual search array. RT were faster to targets presented at the statistically structured array location, despite the regularity carrying no predictive information as to the target location.

Zhao *et al.* [44] suggest that the prioritization of regular features is a means to focus resources on stable aspects of the world, which can then be learnt. The notion that the brain is ‘hardwired’ to prioritize regularities is at the heart of popular models of the brain as a statistical organ of prediction. The expected precision of bottom-up information streams plays a vital role in such predictive processing accounts [14,31]. Reliable prediction errors are up-weighted in proportion to their expected precision, thereby refining the brain’s generative model based on the most informative streams [31].

These ideas may help explain an intriguing recent finding concerning the passive brain response to acoustic patterns. Barascud *et al.* [19] found a substantial increase in the neural response to regularly repeating sound sequences over similar random sequences. This finding seems contrary to a large body of work showing reduced responses to predictable stimuli [45–50]. The proposed explanation for the discrepancy is that, unlike many of the signals used in previous work which often consist of oddball or roving sequences [45,51], the stimuli used in [19] were complex auditory patterns where the predictability of sound sequences was not confounded with neural adaptation resulting from repetition of identical sounds. There are several possible explanations for the increased response to regularity. One is that it reflects the engagement of neural circuits for sequence learning, whose activity in addition to the basic response to the stimulus in auditory cortex results in a net increase in magnetic field strength. Another is that the same neural population is simply more active, with the effect resulting from an increased gain on the activity of auditory neurons responding to the stimuli, potentially signalling greater expected precision. At the cognitive level, the result could potentially indicate that subjects were having their attention spontaneously biased towards the regular sounds, even though they were engaging in an unrelated visual task. Indeed, it has been shown that

neural response magnitude is enhanced to attended, predictable stimuli in audition [52,53], and in vision [29].

In the experiments presented in this paper, we investigate whether regularity captures (exogenous) attention. We use the same stimuli as [19], consisting of tone-pip sequences whose frequency pattern is either regularly repeating (REG) or random (RAND; figure 1). In Experiment 1, we demonstrate that the increased brain response to REG relative to RAND also occurs in electroencephalography (EEG). In a series of behavioural experiments, we then investigate the capacity of REG and RAND to exogenously capture attention when they act as auditory distractors (Experiment 2) and test whether auditory regularity biases attention in scenarios where multiple sound streams are attended and task-relevant (Experiment 3). In both of these paradigms, we find no evidence for attentional capture by acoustic regularity.

## 2. Experiment 1 (electroencephalography)

Experiment 1 investigated EEG brain responses to regularly repeating (REG) and random (RAND) tone-pip sequences of varying complexity (figure 1) while participants, naive to the auditory stimuli, were engaged in an unrelated visual task. The stimuli were identical to those previously used by Barascud *et al.* [19] in a magnetoencephalography (MEG) study. The main aim was to replicate these results using EEG. In addition to the differences in sensitivity between the two techniques [54], a replication in EEG is key to assimilating those findings with the existing literature, where the majority of electrophysiology studies on regularity detection use EEG.

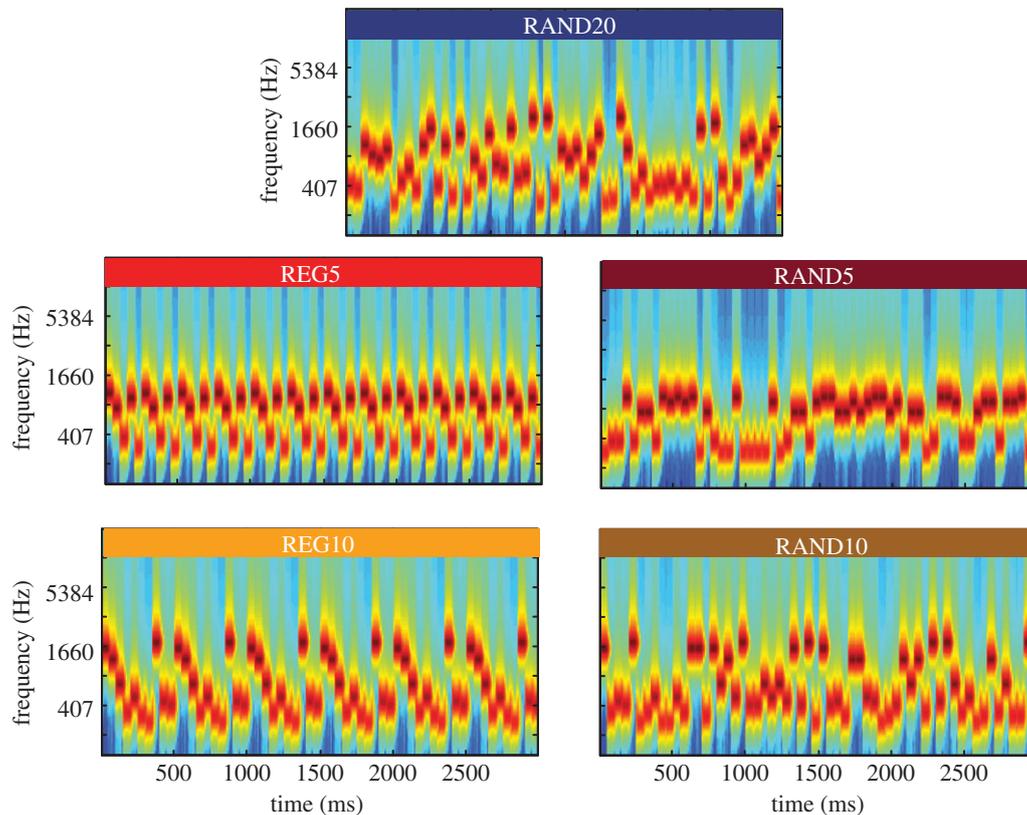
### (a) Methods

#### (i) Stimuli

Stimuli (figure 1) were 3000 ms long sequences of 50 ms tone pips (60 tone pips altogether; each ramped on and off with a 5 ms raised cosine ramp). Tone frequencies were drawn from a pool of 20 logarithmically spaced values between 222–2000 Hz. A unique sequence was presented on each trial. Sequences were defined by two parameters: *Rcyc* (alphabet size)—the number of frequencies chosen (at random, with replacement) from the pool, and *regularity* (REG or RAND). In regular (REG) sequences, a sub-pool of *Rcyc* frequencies were chosen from the full pool, and arranged in repeating cycles of length *Rcyc*. Random (RAND) sequences were generated by drawing each tone at random from the sub-pool of *Rcyc* frequencies. REG and RAND sequences of the same *Rcyc* were generated in pairs, using the same sub-pool, such that conditions were matched for the occurrence of each frequency (figure 1). REG conditions used *Rcyc* = 5, 10 and 15; RAND included an additional condition of *Rcyc* = 20 (using the whole frequency pool), yielding 7 conditions (REG5, REG10, REG15, RAND5, RAND10, RAND15 and RAND20). These sequences are too rapid to allow deliberate reasoning of the order of individual tones; nevertheless, the repetitions in REG sequences lead to a strong, ‘pop-out’ percept of a pattern [19]. Examples of the stimuli used are provided as the electronic supplementary material.

#### (ii) Procedure

The procedure was similar to the MEG experiment described in [19]. Subjects were engaged in an incidental visual task



**Figure 1.** Example spectrograms of the RAND and REG stimuli used. RAND20 (top) contains all 20 frequencies from the pool, in random order. REG5 and REG10 (left) consist of a regularly repeating pattern of 5 or 10 tones, with frequencies chosen at random from the pool (REG15 was also used but not shown). For each REG stimulus, a matching RAND stimulus, consisting of the same frequencies but in random order, was generated. All stimuli were unique (never repeated) and generated anew for each participant.

and were naive about the nature of the auditory stimuli. Auditory stimuli were presented binaurally with the Psychophysics Toolbox extension in MATLAB [55]. In total, subjects heard 700 unique stimuli (100 for each condition). The inter-stimulus interval (ISI) was jittered between 1100 and 1500 ms. The visual task was displayed on a separate computer using Cogent 2000 in MATLAB ([www.vislab.ucl.ac.uk/cogent.php](http://www.vislab.ucl.ac.uk/cogent.php)). The timing was not correlated with that of the auditory stimuli. For each trial, three colour photographs of landscapes were shown for 5 s each, and images faded gradually from one image to the next to minimize visual transients. Subjects were instructed to press a keyboard button if the first and third image within a trial were identical (10% of trials), and to withhold a response otherwise. Inter-trial interval was jittered between 2 and 5 s. The session was split into four consecutive blocks. Feedback (number of hits, misses and false alarms) for the visual task was provided at the end of each block.

### (iii) Recording and data preprocessing

EEG signals were recorded using a Biosemi system (Biosemi Active Two AD-box ADC-17, Biosemi, Netherlands) with 64 electrodes; at a sampling rate of 2048 Hz. Recording was restarted at each block. Data were analysed with SPM12 (Statistical Parametric Mapping; <http://www.fil.ion.ucl.ac.uk/spm/>) and Fieldtrip (<http://www.fieldtriptoolbox.org/>; [56]) toolboxes for MATLAB (2015a, MathWorks). All filtering was performed with a two-pass, Butterworth, fifth order filter. Data were low-pass filtered at 110 Hz, downsampled at 256 Hz, high-pass filtered at 0.1 Hz, re-referenced to the average, divided into 5000 ms epochs (with 1000 ms pre stimulus onset and 1000 ms post-offset) and baseline-

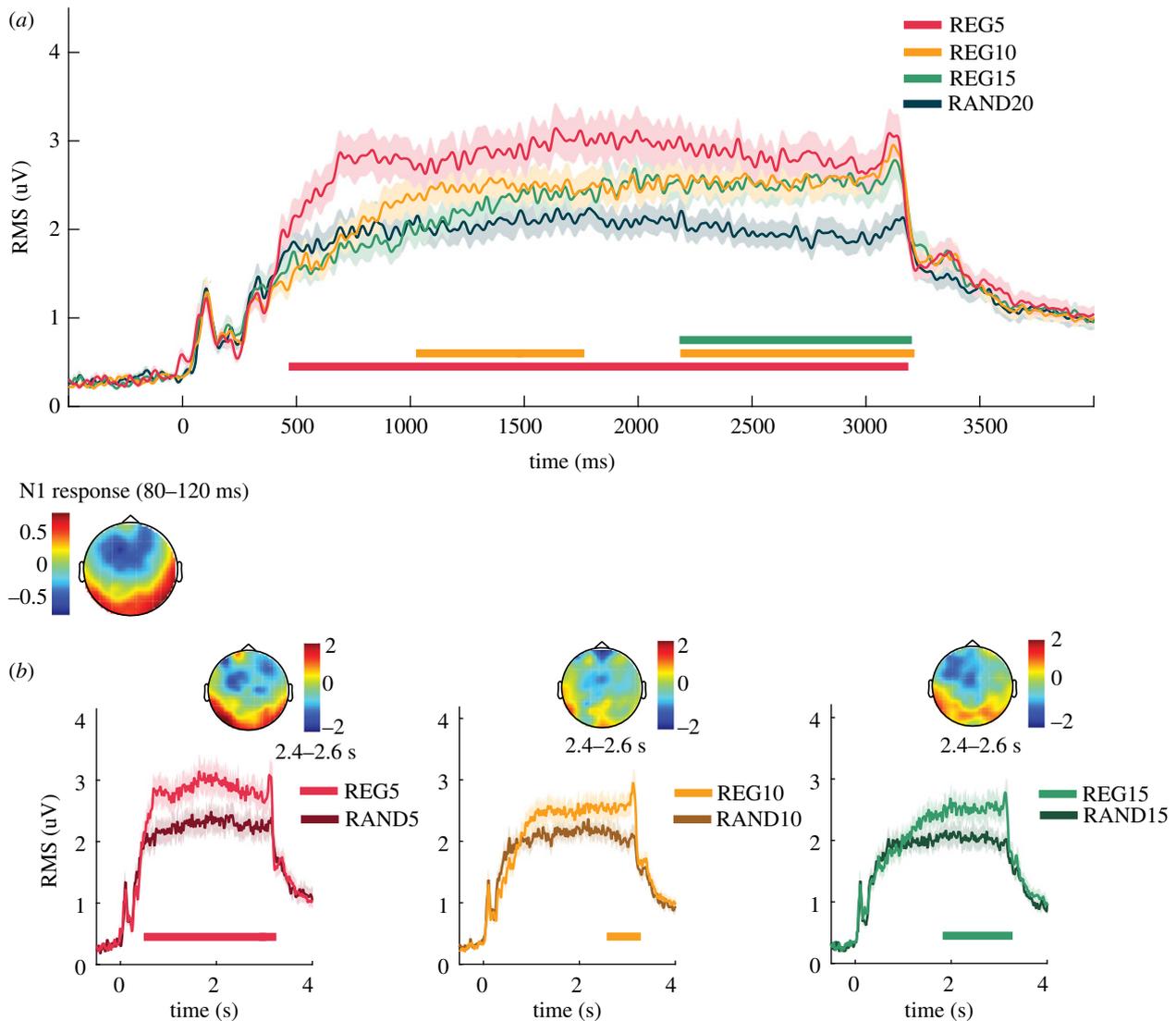
corrected relative to the pre-onset interval. Outlier epochs were removed, if the average power over all time samples and channels exceeded 2 s.d. from the mean over trials; on average, 76% of epochs were retained. Subsequently, data were low-pass filtered at 30 Hz and de-noising source separation (DSS; [57,58]) was applied to maximize reproducibility across epochs, keeping the first five components and projecting back into sensor space. Finally, data were averaged over epochs for each channel, condition and subject.

### (iv) Data analysis

For each participant and condition, the root-mean-square (RMS) over channels was calculated at each time sample in the epoch. This was used as a measure of brain activation over time. The distribution of RMS (mean, s.e.) was then estimated for each condition using bootstrap resampling across subjects (1000 iterations; [59]). This was used to calculate the group-level *t*-statistic of the difference between pairs of conditions at each time-point. *T*-tests (two-tail) were performed using *t*-statistics computed on clusters in time, and controlled for a family-wise error rate of 0.05 [60]. Additionally, a repeated-measures ANOVA with factors of regularity and alphabet size was performed on the mean RMS power between 1000 and 3000 ms, including all conditions except RAND20 to give a balanced design.

### (v) Subjects

In total, 23 paid subjects took part (mean age 23.3 years, range 20–29 years; 11 female). Two subjects were excluded due to exceptionally noisy EEG data. None reported a history of hearing impairment or neurological disorder.



**Figure 2.** Group results of Experiment 1 (EEG). Shaded error margins show  $\pm 2$  s.e.m. (a) EEG evoked responses (group RMS over all channels) for RAND20 and REG of different  $R_{cyc}$ , over the entire epoch. Horizontal bars below plots indicate time intervals where cluster-level statistics showed a significant difference between each of the REG conditions and RAND20. (b) The responses to REG are re-plotted alongside their respective RAND controls. Horizontal bars below plots indicate time intervals where cluster-level statistics showed a significant difference between each pair of conditions. The scalp voltage map of the difference between each pair of REG and RAND conditions, calculated between 2.4 and 2.6 s post-onset, are plotted above each trace. For comparison to a standard scalp voltage distribution in response to a simple sound, the scalp topography of the N1 (calculated over a 40 ms window centred on 100 ms following stimulus onset) is also provided.

## (b) Results and discussion

Group RMS (RMS of all subjects' RMSs) for the three regular conditions, REG5, REG10 and REG15, alongside RAND20 as a common control, are shown in figure 2a. The brain response shows an N1 peak (at 100 ms post-onset) before rising gradually and reaching a sustained level, which persists until stimulus offset. An offset response is visible about 100 ms after sequence cessation. The sustained evoked response is characterized by regular fluctuations at 20 Hz reflecting responses to individual tones. All three REG conditions show an increased sustained response when compared with RAND20. The timing at which the group RMS for REG conditions diverge from RAND20 (taken to reflect the time required by the brain to discover the regularity) increases with cycle length: 406 ms (8 tones), 750 ms (15 tones), 1067 ms (21 tones); for REG with  $R_{cyc} = 5, 10, 15$ , respectively. This is during the second cycle in each case (1.6, 1.5 and 1.4 cycles, respectively), before the pattern has repeated completely, although stable statistical significance is reached somewhat later (horizontal lines

beneath the RMS plot). As discussed in [19], this demonstrates the operation of a rapid, automatic process of regularity detection. Group RMS for REG and RAND of matched  $R_{cyc}$  are shown in figure 2b. The response to REG is consistently higher than its matched RAND. The scalp voltage map of the difference between REG and matched RAND conditions, calculated between 2.4 and 2.6 s post-onset, is shown in figure 2b. For comparison to a standard scalp voltage distribution in response to sound, the scalp topography of the N1 onset response (calculated over a 40 ms window centred on 100 ms following stimulus onset) is also provided.

The RMS, over the interval 1000–3000 ms post-onset, extracted from each condition, was subjected to a repeated-measures ANOVA with regularity (REG versus RAND) and  $R_{cyc}$  (alphabet size of 5, 10 or 15 tones) as within subject factors. This yielded significant main effects of regularity ( $F = 23.2$ ,  $p < 0.001$ ) and of  $R_{cyc}$  ( $F = 5.4$ ,  $p = 0.014$ ), with no interactions. The relative mean increase between RAND and REG of matched  $R_{cyc}$  was 29.7%, 36.7% and 27.5% for  $R_{cyc} = 5, 10$  and 15, respectively.

Experiment 1 demonstrated a rapid differentiation in brain response between regular and random sequences, replicating the MEG results [19]. However, the EEG responses are somewhat noisier than those measured with MEG, probably influenced by a number of factors including ambient electric noise and the smaller number of sensors used here (64, versus 274 in MEG), which impairs the efficiency of de-noising. Electrode voltage drifts, which introduce low frequency noise, may also have affected the robustness of the observed sustained responses.

Overall, the data demonstrate that regular, predictable sequences lead to a dramatic power increase of some 30%, which is remarkable for evoked responses, and suggests a large change in the underlying neural activity. This finding is surprising in the light of previous work consistently reporting reduced evoked responses to predictable patterns and interpreted as reflecting reduced prediction error (e.g. [47]; for a review see [51]). The discrepancy with previous work may be due to much of the existing work using sound patterns, which confound repetition with predictability, making it difficult to dissociate adaptation effects from those purely due to prediction. The present paradigm, using wide-band signals and complex sound patterns allows us to control for simple effects of adaptation (see additional discussion in [19]). Furthermore, we use very rapid sequences, where the perception of patterns pops-out spontaneously rather than being consciously trackable. It is likely that the neural processes involved in extracting the regularity are different from those implicated in work using slower temporal patterns (e.g. [52,53]), which allow high-level (conscious or mnemonic) prediction of future events.

One possible explanation for the activation pattern observed here is that it reflects automatic, bottom-up-driven attentional capture by REG patterns. This attentional process will be the focus of the rest of this paper. The behavioural experiments below investigate the hypothesis that the large, sustained amplitude shift that was observed for the REG stimuli may reflect increased perceptual salience [36,37]. That is to say, the more reliable REG stimuli trigger an automatic (exogenous) attentional bias. This proposition leads to the testable prediction that REG and RAND will have different effects on behaviour, reflecting an attentional bias towards regularity in REG sequences, even when task-irrelevant.

### 3. Experiment 2

This experiment aimed to measure the (assumed) behavioural consequences of attentional capture by regular sounds. We evaluated performance on a demanding listening task, with REG or RAND sequences presented concurrently, as task-irrelevant distractors. If REG patterns spontaneously capture exogenous attention, we predicted that REG sequences will prove more detrimental to performance than RAND sequences. This prediction is in line with previous behavioural experiments, whereby task-irrelevant stimuli outside the focus of attention can result in attentional capture manifest as degradation in performance in a behavioural task [61,62].

The main task was based on an auditory change-detection paradigm [63,64]. Stimuli were artificial acoustic scenes, comprised of multiple simultaneous tone-pip streams, each characterized by a distinct, constant frequency and amplitude-modulation (AM) rate. The task required listeners

to detect occasional changes (appearance or disappearance of one stream) within these scenes. This simulates the challenges faced by listeners in natural acoustic scenes, in which many concurrent sound sources must be processed and monitored simultaneously.

We presented the change-detection task and REG-RAND distractor sequences concurrently to different ears, such that they competed directly throughout the trial (figure 3a). If REG patterns spontaneously capture exogenous attention, we predicted that REG sequences will prove more detrimental to performance than RAND sequences.

## (a) Methods

### (i) Stimuli

The stimuli and experimental approach for the change-detection paradigm are described in detail in a previous study [65]. In brief, stimuli were artificial acoustic scenes consisting of eight concurrent streams of tone pips, each with a unique frequency (between 200 and 4000 Hz) and AM rate (3 to 35 Hz). In total, 50% of the stimuli contained a change partway through the scene: appearance (CA) or disappearance (CD) of a stream. Scene changes occurred between 1000 and 2000 ms post-onset. The overall stimulus duration was between 2000 and 4000 ms.

On each trial, a scene stimulus and a REG10 or RAND10 distractor sequence (with equal probability) were presented concurrently at the same dB level, to different ears, such that they competed directly throughout the trial (figure 3a). Each REG-RAND sequence consisted of between 40 and 80 tones. For REG sequences, this constituted between 4 and 8 cycles; i.e. sufficient for the regularity to become perceptually established.

### (ii) Procedure

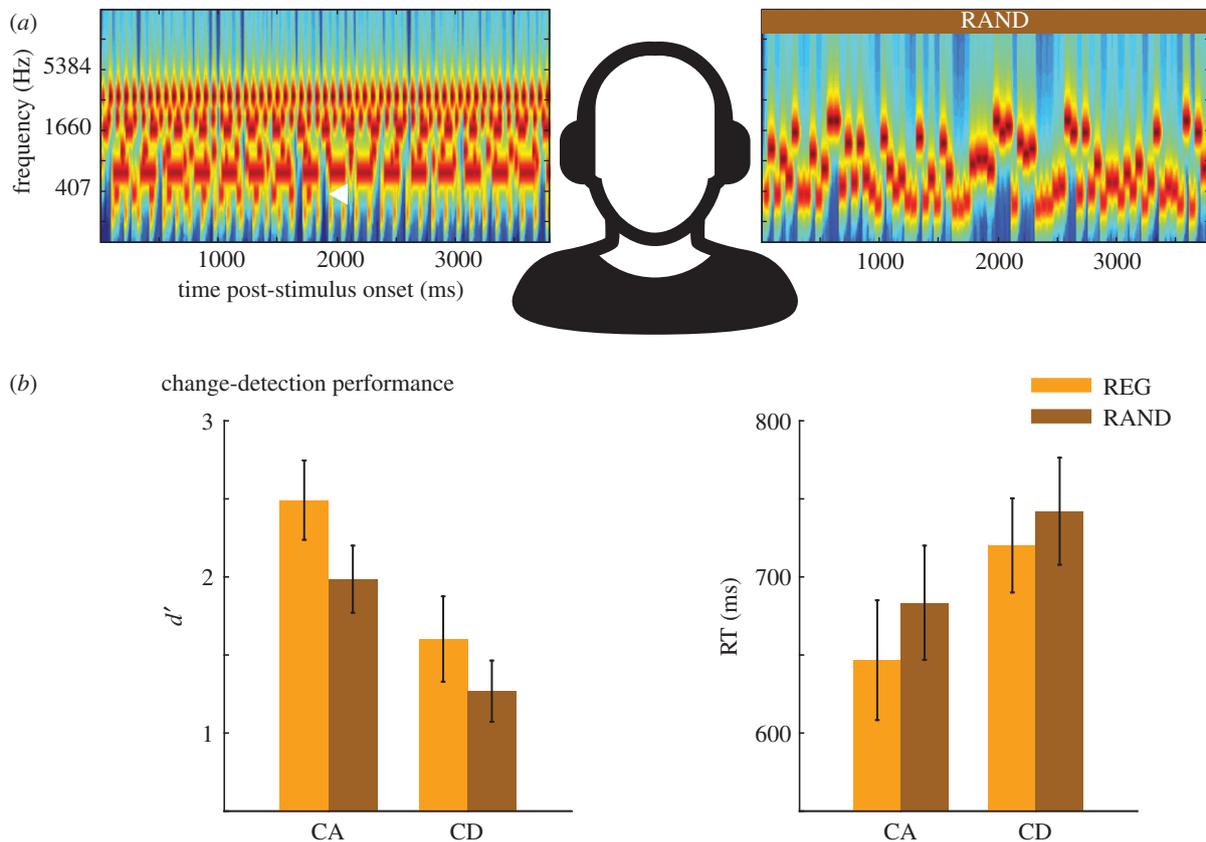
Stimuli were blocked by change type (CA or CD), with 50% of the trials in each block (160 overall) containing a change. The ISI was randomized between 700 and 2000 ms. REG or RAND sequences were randomly paired with each scene stimulus. Subjects were instructed to attend to the ear containing the scene and respond by button-press as soon as they heard a change. To avoid confusion, the ear of presentation was fixed throughout the experiment, but counterbalanced across subjects. Subjects were naive to the structure of the REG-RAND sequences, and told these were simply distractors to the main change-detection task. The block order was counterbalanced between subjects, and a break was allowed after every 40 trials. The session began with a short training block where feedback was given on each trial.

### (iii) Subjects

Ten subjects participated in this experiment (mean age 24.0 years; 7 female).

## (b) Results and discussion

Figure 3b shows RT and sensitivity ( $d'$ ) scores for detection of scene changes with REG or RAND distractors presented concurrently. A repeated-measures ANOVA was performed on RT and  $d'$ ; with change type (CA/CD) and distractor regularity (REG/RAND) as factors. RT showed main effects of change type ( $F = 16.299$ ;  $p = 0.003$ ) and regularity ( $F = 8.064$ ;  $p = 0.019$ ) with no interaction. Similarly,  $d'$  showed



**Figure 3.** Experiment 2. (a) A schematic representation of the stimulus paradigm. Scene stimuli and REG or RAND sequences (RAND in this example) were presented concurrently to different ears (counterbalanced across subjects). All scenes contained eight streams. This example shows a scene with a disappearing (CD) stream, which is indicated with a white arrowhead. (b) Behavioural results;  $d'$  (left) and reaction time (right) for detection of change events. Error bars show  $\pm 1$  s.e.m.

a main effect of change type ( $F = 18.244$ ;  $p = 0.002$ ) and regularity ( $F = 9.786$ ;  $p = 0.012$ ) with no interaction.

The results reveal that, contrary to our hypothesis, RAND is more detrimental to performance than REG. The data are consistent with the interpretation that RAND is harder to ignore, at least when in direct competition with a concurrent, task-relevant, auditory stream (see also [23]; more discussion below).

## 4. Experiment 3

Rather than using REG or RAND sequences as task-irrelevant distractors, Experiment 3 placed REG and RAND sequences in direct competition as task-relevant streams. Here, REG and RAND were presented concurrently and both actively monitored for targets (a silent gap). This is in contrast with Experiment 2, where performance on the task required ignoring REG or RAND stimuli. We predicted that when both sequence types are monitored simultaneously, gaps in REG sequences should be more readily detectable than gaps in RAND sequences. This design is similar to that used in [44], who demonstrated that targets embedded within regularly repeating visual streams are more easily detected, even though the regularity of the stream was not itself goal-relevant.

### (a) Methods

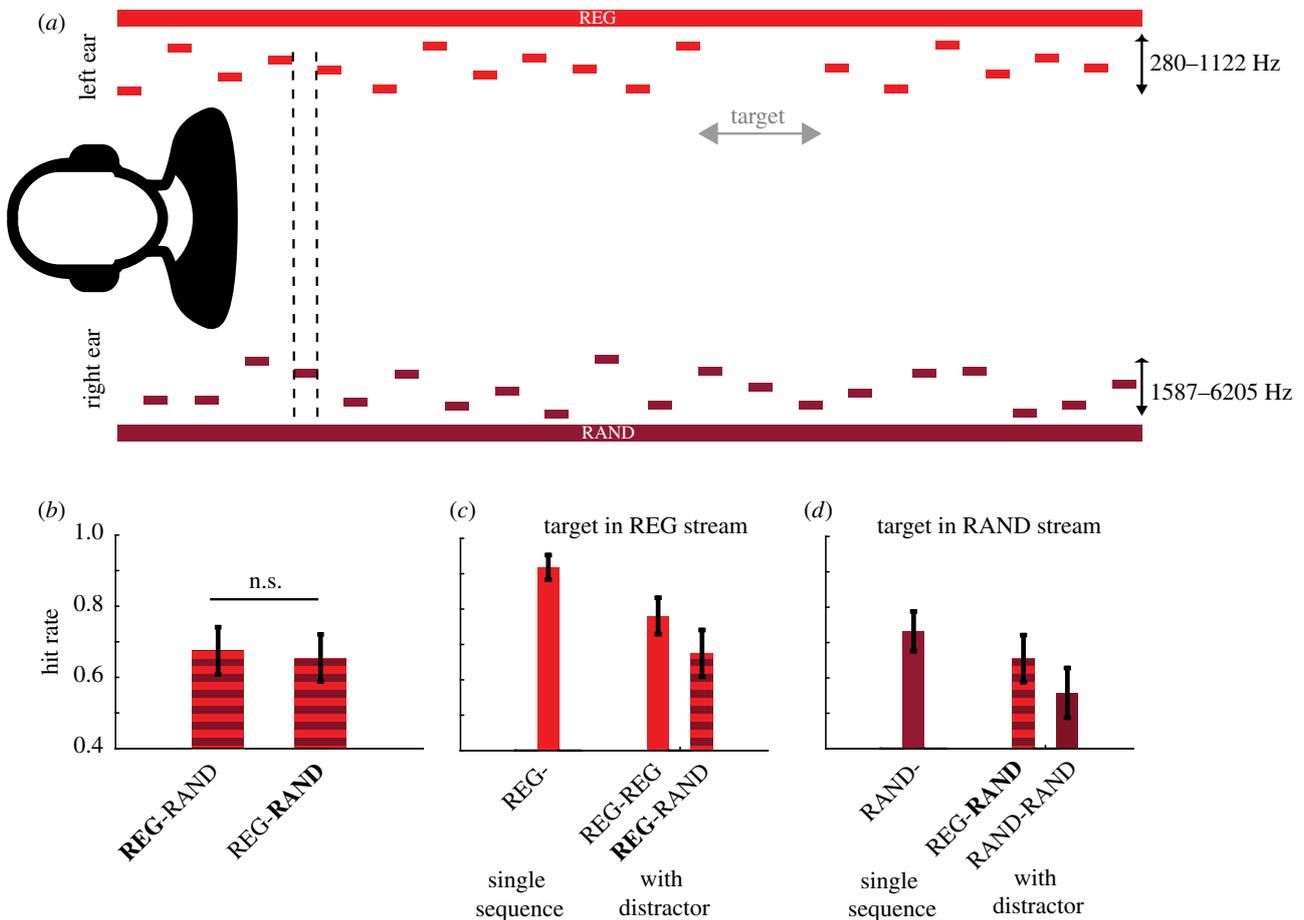
#### (i) Stimuli

This experiment used REG5 and RAND5 sequences consisting of 50 ms tone pips interspersed with 50 ms gaps. Trials involved the presentation of two concurrent sequences, one

in each ear (figure 4a). Sequences could be both REG or both RAND, or one of each. In order to facilitate the perception of the two sequences as independent concurrent streams, the sequences were staggered by 50 ms, such that tones occurred in alternation between the ears. In addition, sequences were spectrally separated, such that the sequence in the right ear was always a higher pitch. The tones were chosen from a pool of 13 logarithmically spaced frequencies between 1587 and 6205 Hz for the right ear and between 280 and 1122 Hz for the left.

On 50% of trials, one of the sequences contained a target (an omission of two consecutive tones). Stimuli were 6000 ms long. When present, the target occurred at least 2000 ms after stimulus onset (following four REG cycles); i.e. at a point in the REG stimulus when the regular pattern has been established.

The main experiment consisted of the following conditions: (i) RAND sequences in both ears (RAND-RAND, 25% of trials); (ii) REG sequences in both ears (REG-REG, 25% of trials); (iii) REG sequence in one ear and RAND in the other (REG-RAND, 50% of trials). The target occurred in one of the two sequences with equal probability. For REG-RAND, we denote the stream containing the target using bold type; thereby sub-dividing this condition into (iii(a)) **REG**-RAND and (iii(b)) REG-**RAND**. Each condition was counterbalanced across the two ears, such that target occurrence and REG versus RAND sequences were equally likely in each ear. Before the main session, subjects also completed a block where only a single-sequence (REG or RAND) was presented to one of the ears with equal probability; a target was present on 50% of trials. These conditions are denoted as REG- and RAND-



**Figure 4.** Experiment 3. (a) A schematic representation of the stimulus paradigm. On each trial, two concurrent sequences (each either REG or RAND) were presented, one to each ear. Individual tones were interleaved as demonstrated with the black dashed line. On 50% of the trials, one of the sequences contained a 200 ms gap (target, shown in grey, top). The sequences presented to the right ear were of a higher pitch than to the left, such that the frequency ranges did not overlap. (b) Hit rate for target detection when REG was presented to one ear and RAND to the other. The sequence containing the target is indicated in bold. (c) Hit rate for all conditions with a target in the REG stream; single-sequence REG (left) and dual-sequence (right). (d) As for C; except all conditions had a target in the RAND stream. Error bars show  $\pm 1$  s.e.m.

### (ii) Procedure

Subjects were instructed to press a keyboard button as soon as they heard a gap in a sequence. Feedback was provided after each trial. The single-sequence block contained 80 stimuli (40 for each of REG- and RAND-). The main experiment contained 120 trials for each of the four dual-sequence conditions, presented in a randomized order. Subjects were given short breaks every 10 min.

### (iii) Subjects

Eleven subjects participated in this experiment (mean age = 25.7; 7 female)

### (b) Results and discussion

We used hit rates as our primary outcome measure, as most subjects produced low false positive rates (2.1% for single-sequence, 2.0% for dual-sequence) yielding 'artificially' high  $d'$  scores. Hit rates were deemed the most unambiguous and representative outcome measure of performance in this task.

In order to test the main hypothesis that regular sequences would 'pop-out' and attract attention, we initially compared performance for REG-RAND and REG-RAND, as both contain simultaneously presented random and regular sequences (figure 4b). We postulated that regularity would bias

attention, leading to improved performance when targets were embedded in regular streams (REG-RAND) as opposed to random (REG-RAND). A repeated-measures ANOVA showed no significant difference between the average hit rate values for REG-RAND and REG-RAND (hit rates were 0.65 and 0.67, respectively;  $F = 1.2$ ;  $p = 0.3$ ). These results suggest that regular sound patterns do not bias attention.

Figure 4c,d shows the hit rate for all conditions, separated by whether the target is in the REG or the RAND stream—for both single-sequence and dual-sequence stimuli. The hit rates for targets in the single-sequence (REG- and RAND-) conditions are the left-most bars in each plot. In the single-sequence condition, subjects were better at identifying targets in regular streams compared with random streams ( $F = 20.6$ ;  $p = 0.001$ ). These findings are consistent with previous work showing that performance is improved when targets are embedded in temporally regular sequences [66–68]. This is the case even when the dimension in which the regularity expressed is independent of the dimension along which targets differ.

A repeated-measures ANOVA was conducted on hit rates in the dual-sequence conditions, with factors for regularity of the target stream (REG-RAND) and the parallel stream (REG-RAND). There was a main effect of target stream ( $F = 51.48$ ;  $p < 0.01$ ); here again subjects were overall better

at detecting targets in REG. We also found a main effect of parallel stream ( $F = 26.97$ ;  $p < 0.01$ ); revealing overall poorer target detection when the parallel stream was RAND relative to when it was REG. This pattern is in agreement with the outcomes of Experiment 2, and consistent with the interpretation that RAND patterns incur increased demand on processing resources (discussed further below). However, there was no interaction between the two factors, suggesting that a RAND parallel stream was equally costly to target-detection performance in a RAND or a REG stream.

## 5. Discussion

Brain responses measured with functional magnetic resonance imaging (fMRI), MEG [19] and EEG (as seen here) show consistently increased activation to regular acoustic patterns, relative to matched random stimuli. One interpretation of these systematic, pronounced effects is that they indicate large differences in attentional capture between regular and random patterns, such that regular patterns automatically and involuntarily attract more attention. This account of the imaging data is consistent with previous behavioural work in the visual modality [44] and is broadly in line with the fact that sensitivity to predictable patterns in the natural environment is a major pre-requisite for survival. Organisms produce regular, periodic motor sequences, such as locomotion and vocalizations, which are expressed as a pattern in the temporal succession of sensations. The ability to automatically orient towards such patterns within a crowded, noisy scene is often critical for continued existence.

### (a) Attentional capture by regularity?

The behavioural experiments reported here aimed to identify a behavioural correlate for the observed brain response effects. Using two tasks designed to probe different aspects of attentional capture, we consistently find no evidence for the exogenous capture of attention by regular acoustic patterns. Despite the sizeable change in the EEG signal associated with processing regular (REG), relative to random (RAND) tone-pip patterns, REG sequences were not more distracting than matched random sequences when task irrelevant, and were also no more perceptually salient when participants were actively monitoring REG and RAND streams concurrently.

While there is no evidence for REG sequences being more perceptually salient than RAND sequences (or vice versa), Experiments 2 and 3 suggest that RAND sequences are more computationally demanding, and hence more distracting, than REG sequences (e.g. [69]; the discussion below disentangles these issues).

The paradigm in Experiment 2 shared key similarities with the EEG experiment (Experiment 1). Participants were naive to the nature of the distracting REG or RAND patterns, and focused on a different task. A change-detection task, rather than a visual task similar to that in the EEG experiment, was chosen because: (i) its rapid nature allowed us to probe behaviour more frequently, and hence efficiently; and (ii) a competing auditory (rather than a visual) task is more likely to reveal effects of distraction, because it poses more competition for shared resources (see review in [70,71]). It is therefore unlikely that failure to observe effects of attentional capture is due to the difference in task *per se*. Furthermore, by removing the decoy task altogether,

Experiment 3 constitutes a stricter test for a possible attentional bias. When REG and RAND are monitored concurrently, we observe equal gap-detection performance whether the target is in REG or RAND, which suggests that they do not differ in their perceptual salience.

The reasons for the discrepancy with results from vision [44], where effects of attentional capture by regularity have been reported, are unclear and may be due to many factors, perhaps including a genuine difference in the mechanisms of attentional allocation in the visual and auditory domain. Further work directly comparing the two modalities is required to resolve this issue.

### (b) Processing of regular versus random sequences

The results of Experiment 3, demonstrating increased sensitivity to targets in REG relative to RAND sequences when presented alone, are consistent with many previous demonstrations that regularity facilitates behavioural performance. These studies, albeit mostly using regularity in the temporal dimension rather than in frequency as we do here, consistently show that regularity facilitates behavioural performance. Expected events are detected and assessed more rapidly and accurately than unexpected events [61–64,66,67,72–75]. This occurs, as is the case here, even when the task dimension is orthogonal to the feature dimension over which the regularity is defined (e.g. [68]) and hypothesized to arise due to the ‘pre-activation’ of the relevant neural machinery for processing-predicted events [63,76].

The same processes have been demonstrated to contribute to the *suppression* of regular streams when they are not behaviourally relevant. For example, Andreou *et al.* [23] demonstrated that it is easier to ignore a temporally regular sequence, relative to a temporally irregular sequence (see also [24,77,78]). Similarly, in Experiment 2, we show that REG sequences are less distracting than RAND sequences when participants are required to ignore those sequences and focus on a competing change-detection task. A potential mechanism for this effect is supplied by predictive coding [15,16], whereby predictable inputs are attenuated by top-down predictions, and the resulting prediction error triggers a process of updating the internal predictive model. Regularity allows the derivation of a predictive rule; therefore, it becomes easier to ‘explain away’ the irrelevant stimulus, by suppressing the prediction error with a closely matching top-down prediction. Irregular stimuli demand more resources for processing as they elicit a constant stream of prediction errors and thus constantly trigger model updating. This may be taken to suggest that (unpredictable) RAND sequences are more perceptually salient. From the point of view of predictive coding, this is sensible because RAND sequences are characterized by higher information content than REG sequences. Friston *et al.* [79] define salience in terms of the ability to reduce uncertainty or to inform hypotheses about the sensory scene being sampled. In the visual domain, this is usually measured in terms of Bayesian surprise [33] and more generally as information gain or epistemic value [80]. However, whether the theoretically information-rich RAND signals are useful in reducing uncertainty about high-level representations is an open question. In other words, is unpredictability itself salient? The lack of a bias in performance between REG and RAND sequences presented in direct competition (Experiment 3) suggests not.

### (c) If not attentional capture, what is the source of the electroencephalography effect?

EEG and MEG measurements demonstrate an increase in the amplitude of the sustained response for REG relative to RAND stimuli. In fMRI, this is associated with greater activation for REG relative to RAND sequences across a large portion of the superior temporal gyrus, including Heschl's gyrus and planum temporale [19]. The behavioural results reported above suggest that this increased activation is not associated with attentional bias towards (or increased perceptual salience of) REG sequences.

Critically, the above explanation for why RAND sequences were, in some cases, more detrimental to performance than REG sequences implies *more* activation (increased demand for computational resources) for random patterns relative to regular ones. This may seem contradictory to the brain-level effects. However, it is possible, that the increased auditory cortical activation for regular patterns observed in M/EEG and fMRI reflects increased *inhibition*. It is difficult to dissociate excitatory and inhibitory activation with standard non-invasive brain imaging techniques; rather future computational and electrophysiological tools would be critical for exploring this possibility. Indeed, recent findings in animal models demonstrate a critical role for inhibition in shaping the response of primary auditory cortex neurons to regularly repeating sounds in the context of an oddball paradigm [81].

Another potential explanation for the larger response to REG is that regularity detection is associated with heightened sensitivity (increased gain) of the sensory units activated by the regular pattern. According to this 'precision-weighting' account, precise, i.e. highly predictable, sensory streams are preferentially weighted by increasing the post-synaptic gain of the relevant (prediction error) units [14]. Importantly, this can occur within the remit of automatic processing, so does not entail attentional capture [31].

Lastly, it is possible that the increased sustained response we observe is due to another process (or indeed multiple processes) such as learning, working memory or recognition of a match to a memory of previous stimulation (see [19] for further discussion). This interpretation is consistent with the diffuse source network including auditory cortex, hippocampus and inferior frontal gyrus identified in [19] as contributing to the brain response to structured sequences.

To summarize, a picture emerges from these results in which regularity in non-attended items does not capture attention. In fact, as demonstrated in Experiment 2, random stimuli can be more distracting than regular ones. Consistent with the literature, we found that regularity does, however, aid in scene analysis by being easier to ignore (Experiment 2) and requiring fewer resources to process (Experiment 3). Collectively, the behavioural and brain imaging findings can be reconciled by considering both to result from mechanisms that minimize surprise and uncertainty about the world [10,25].

**Ethics.** All experimental procedures reported in this paper were approved by the research ethics committee of University College London, and written informed consent was obtained from each participant.

**Data accessibility.** The behavioural and EEG data are available online, alongside MATLAB code for analysis and plotting. This is available through the Dryad repository <http://doi.org/10.5061/dryad.j17h3> [82].

**Authors' contributions.** R.S., A.B., C.G., N.B. and M.C. designed the experiment. R.S., A.B. and C.G. acquired and analysed the data. R.S., A.B., C.G., N.B., K.F. and M.C. interpreted the results, as well as writing and revising the article. R.S., A.B., C.G., N.B., K.F. and M.C. approved the final version of the manuscript.

**Competing interests.** The authors have no competing financial interests to declare.

**Funding.** This work was supported by a Brain Research Trust PhD studentship to R.S. and a BBSRC project grant no. (BB/K003399/1) to M.C. K.F. is funded by a Wellcome Trust Principal Research Fellowship (Ref: 088130/Z/09/Z).

## References

- Saffran JR, Johnson EK, Aslin RN, Newport EL. 1999 Statistical learning of tone sequences by human infants and adults. *Cognition* **70**, 27–52. (doi:10.1016/S0010-0277(98)00075-4)
- Chait M, Poeppel D, de Cheveigné A, Simon JZ. 2007 Processing asymmetry of transitions between order and disorder in human auditory cortex. *J. Neurosci.* **27**, 5207–5214. (doi:10.1523/JNEUROSCI.0318-07.2007)
- 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. (doi:10.1162/jocn.2009.21131)
- Barakat BK, Seitz AR, Shams L. 2013 The effect of statistical learning on internal stimulus representations: predictable items are enhanced even when not predicted. *Cognition* **129**, 205–211. (doi:10.1016/j.cognition.2013.07.003)
- 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. (doi:10.1016/j.cub.2015.06.035)
- Fiser J, Aslin RN. 2001 Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychol. Sci.* **12**, 499–504. (doi:10.1111/1467-9280.00392)
- Conway CM, Christiansen MH. 2005 Modality-constrained statistical learning of tactile, visual, and auditory sequences. *J. Exp. Psychol. Learn. Memory Cogn.* **31**, 24–39. (doi:10.1037/0278-7393.31.1.24)
- Gebhart AL, Newport EL, Aslin RN. 2009 Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. *Psychon. Bull. Rev.* **16**, 486–490. (doi:10.3758/PBR.16.3.486)
- Bendixen A, Schröger E, Ritter W, Winkler I. 2012 Regularity extraction from non-adjacent sounds. *Front. Psychol.* **3**, 143. (doi:10.3389/fpsyg.2012.00143)
- Bendixen A. 2014 Predictability effects in auditory scene analysis: a review. *Front. Neurosci.* **8**, 1–16. (doi:10.3389/fnins.2014.00060)
- Paavilainen P. 2013 The mismatch-negativity (MMN) component of the auditory event-related potential to violations of abstract regularities: a review. *Int. J. Psychophysiol.* **88**, 109–123. (doi:10.1016/j.ijpsycho.2013.03.015)
- Garrido MI, Sahani M, Dolan RJ. 2013 Outlier Responses Reflect Sensitivity to Statistical Structure in the Human Brain. *PLoS Comput. Biol.* **9**, e1002999. (doi:10.1371/journal.pcbi.1002999)
- Agus TR, Thorpe SJ, Pressnitzer D. 2010 Rapid formation of robust auditory memories: insights from noise. *Neuron* **66**, 610–618. (doi:10.1016/j.neuron.2010.04.014)
- Friston K. 2005 A theory of cortical responses. *Phil. Trans. R. Soc. B* **360**, 815–836. (doi:10.1098/rstb.2005.1622)
- Clark A. 2013 Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* **36**, 181–204. (doi:10.1017/S0140525X12000477)
- Hohwy J. 2013 *The predictive mind*. Oxford, UK: Oxford University Press.

17. Nordby H, Roth WT, Pfefferbaum A. 1988 Event-related potentials to breaks in sequences of alternating pitches or interstimulus intervals. *Psychophysiology* **25**, 262–268. (doi:10.1111/j.1469-8986.1988.tb01239.x)
18. Baldeweg T. 2006 Repetition effects to sounds: evidence for predictive coding in the auditory system. *Trends Cogn. Sci.* **10**, 93–94. (doi:10.1016/j.tics.2006.01.010)
19. Barascud N, Pearce MT, Griffiths TD, Friston KJ, Chait M. 2016 Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. *Proc. Natl Acad. Sci. USA* **113**, E616–E625. (doi:10.1073/pnas.1508523113)
20. Winkler I, Denham SL, Nelken I. 2009 Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends Cogn. Sci.* **13**, 532–540. (doi:10.1016/j.tics.2009.09.003)
21. Nobre AC, Correa A, Coull JT. 2007 The hazards of time. *Curr. Opin Neurobiol.* **17**, 465–470. (doi:10.1016/j.conb.2007.07.006)
22. Denham SL, Winkler I. 2006 The role of predictive models in the formation of auditory streams. *J. Physiol.* **100**, 154–170. (doi:10.1016/j.jphysparis.2006.09.012)
23. Andreou L-V, Kashino M, Chait M. 2011 The role of temporal regularity in auditory segregation. *Hear. Res.* **280**, 228–235. (doi:10.1016/j.heares.2011.06.001)
24. Bendixen A, Denham SL, Gyimesi K, Winkler I. 2010 Regular patterns stabilize auditory streams. *J. Acoust. Soc. Am.* **128**, 3658–3666. (doi:10.1121/1.3500695)
25. Schröger E, Bendixen A, Denham SL, Mill RW, Böhm TM, Winkler I. 2013 Predictive regularity representations in violation detection and auditory stream segregation: from conceptual to computational models. *Brain Topogr.* **27**, 565–577. (doi:10.1007/s10548-013-0334-6)
26. Barascud N, Griffiths TD, McAlpine D, Chait M. 2014 ‘Change deafness’ arising from inter-feature masking within a single auditory object. *J. Cogn. Neurosci.* **26**, 514–528. (doi:10.1162/jocn\_a\_00481)
27. Summerfield C, de Lange FP. 2014 Expectation in perceptual decision making: neural and computational mechanisms. *Nature* **15**, 745–756. (doi:10.1038/nm3838)
28. Summerfield C, Egner T. 2016 Feature-based attention and feature-based expectation. *Trends Cogn. Sci.* **20**, 401–404. (doi:10.1016/j.tics.2016.03.008)
29. Kok P, Rahnev D, Jehee JFM, Lau HC, de Lange FP. 2012 Attention reverses the effect of prediction in silencing sensory signals. *Cereb. Cortex* **22**, 2197–2206. (doi:10.1093/cercor/bhr310)
30. Chennu S, Noreika V, Gueorguiev D, Blenkmann A, Kochen S, Ibanez A, Owen AM, Bekinschtein TA. 2013 Expectation and attention in hierarchical auditory prediction. *J. Neurosci.* **33**, 11 194–11 205. (doi:10.1523/JNEUROSCI.0114-13.2013)
31. Feldman H, Friston KJ. 2010 Attention, uncertainty, and free-energy. *Front. Hum. Neurosci.* **4**, 1–23. (doi:10.3389/fnhum.2010.00215)
32. Pearce JM, Hall G. 1980 A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* **87**, 532–552. (doi:10.1037/0033-295X.87.6.532)
33. Itti L, Baldi P. 2009 Bayesian surprise attracts human attention. *Vis. Res.* **49**, 1295–1306. (doi:10.1016/j.visres.2008.09.007)
34. Kaya EM, Elhilali M. 2017 Modelling auditory attention. *Phil. Trans. R. Soc. B* **372**, 20160101. (doi:10.1098/rstb.2016.0101)
35. Mackintosh NJ. 1975 A theory of attention: variations in the associability of stimuli with reinforcement. *Psychol. Rev.* **82**, 276–298. (doi:10.1037/h0076778)
36. Itti L, Koch C. 2000 A saliency-based search mechanism for overt and covert shifts of visual attention. *Vis. Res.* **40**, 1489–1506. (doi:10.1016/S0042-6989(99)00163-7)
37. Itti L, Koch C. 2001 Feature combination strategies for saliency-based visual attention systems. *J. Electron. Imaging* **10**, 161–169. (doi:10.1117/1.1333677)
38. Pearce JM, Mackintosh NJ. 2010 Two theories of attention: a review and a possible integration. In *Attention and Associative Learning: From Brain to Behaviour* (eds CJ Mitchell, ME Le Pelley), pp. 11–39. Oxford, UK: Oxford University Press.
39. Dayan P, Kakade S, Montague PR. 2000 Learning and selective attention. *Nat. Neurosci.* **3**(Suppl), 1218–1223. (doi:10.1038/81504)
40. Esber GR, Haselgrove M. 2011 Reconciling the influence of predictiveness and uncertainty on stimulus saliency: a model of attention in associative learning. *Proc. R. Soc. B* **278**, 2553–2561. (doi:10.1098/rspb.2011.0836)
41. Chun MM, Jiang Y. 1999 Top-down attentional guidance based on implicit learning of visual covariation. *Psychol. Sci.* **10**, 360–365. (doi:10.1111/1467-9280.00168)
42. Summerfield JJ, Lepsien J, Gitelman DR, Mesulam MM, Nobre AC. 2006 Orienting attention based on long-term memory experience. *Neuron* **49**, 905–916. (doi:10.1016/j.neuron.2006.01.021)
43. Chalk M, Seitz AR, Seriès P. 2010 Rapidly learned stimulus expectations alter perception of motion. *J. Vis.* **10**, 2. (doi:10.1167/10.8.2)
44. Zhao J, Al-Aidroos N, Turk-Browne NB. 2013 Attention is spontaneously biased toward regularities. *Psychol. Sci.* **24**, 667–677. (doi:10.1177/0956797612460407)
45. Haenschel C, Vernon DJ, Dwivedi P, Gruzeliel JH, Baldeweg T. 2005 Event-related brain potential correlates of human auditory sensory memory-trace formation. *J. Neurosci.* **25**, 10 494–10 501. (doi:10.1523/JNEUROSCI.1227-05.2005)
46. Alink A, Schwiedrzik CM, Kohler A, Singer W, Muckli L. 2010 Stimulus predictability reduces responses in primary visual cortex. *J. Neurosci.* **30**, 2960–2966. (doi:10.1523/JNEUROSCI.3730-10.2010)
47. Costa-Faidella J, Baldeweg T, Grimm S, Escera C. 2011 Interactions between ‘what’ and ‘when’ in the auditory system: temporal predictability enhances repetition suppression. *J. Neurosci.* **31**, 18 590–18 597. (doi:10.1523/JNEUROSCI.2599-11.2011)
48. Meyer T, Olson CR. 2011 Statistical learning of visual transitions in monkey inferotemporal cortex. *Proc. Natl Acad. Sci. USA* **108**, 19 401–19 406. (doi:10.1073/pnas.1112895108)
49. Todorovic A, van Ede F, Maris E, de Lange FP. 2011 Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: an MEG study. *J. Neurosci.* **31**, 9118–9123. (doi:10.1523/JNEUROSCI.1425-11.2011)
50. Todorovic A, de Lange FP. 2012 Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *J. Neurosci.* **32**, 13 389–13 395. (doi:10.1523/JNEUROSCI.2227-12.2012)
51. Baldeweg T. 2008 ERP repetition effects and mismatch negativity generation. *J. Psychophysiol.* **21**, 204–213. (doi:10.1027/0269-8803.21.34.204)
52. Hsu Y-F, Hämäläinen JA, Waszak F. 2014 Both attention and prediction are necessary for adaptive neuronal tuning in sensory processing. *Front. Hum. Neurosci.* **8**, 14608. (doi:10.3389/fnhum.2014.00152)
53. Auzztulewicz R, Friston K. 2015 Attentional enhancement of auditory mismatch responses: a DCM/MEG study. *Cereb. Cortex* **25**, 4273–4283. (doi:10.1093/cercor/bhu323)
54. Ahlfors SP, Han J, Belliveau JW, Hämäläinen MS. 2010 Sensitivity of MEG and EEG to source orientation. *Brain Topogr.* **23**, 227–232. (doi:10.1007/s10548-010-0154-x)
55. Kleiner M, Brainard D, Pelli D, Ingling A, Murray R, Broussard C. 2007 What’s new in Psychtoolbox-3. *Perception* **36**, 1–16.
56. Oostenveld R, Fries P, Maris E, Schoffelen J-M. 2010 FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* **2011**, 1–9. (doi:10.1155/2011/156869)
57. de Cheveigné A, Simon JZ. 2008 Denoising based on spatial filtering. *J. Neurosci. Methods* **171**, 331–339. (doi:10.1016/j.jneumeth.2008.03.015)
58. de Cheveigné A, Parra LC. 2014 Joint decorrelation, a versatile tool for multichannel data analysis. *Neuroimage* **98**, 487–505. (doi:10.1016/j.neuroimage.2014.05.068)
59. Efron B, Tibshirani RJ. 1993 *An introduction to the bootstrap, monographs on statistics and applied probability*, vol. 57. New York, NY: Chapman and Hall/CRC.
60. Maris E, Oostenveld R. 2007 Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190. (doi:10.1016/j.jneumeth.2007.03.024)
61. Correa Á, Nobre AC. 2008 Neural modulation by regularity and passage of time. *J. Neurophysiol.* **100**, 1649–1655. (doi:10.1152/jn.90656.2008)
62. Rohenkohl G, Cravo AM, Wyart V, Nobre AC. 2012 Temporal expectation improves the quality of

- sensory information. *J. Neurosci.* **32**, 8424–8428. (doi:10.1523/JNEUROSCI.0804-12.2012)
63. Bendixen A, Schröger E, Winkler I. 2009 I heard that coming: event-related potential evidence for stimulus-driven prediction in the auditory system. *J. Neurosci.* **29**, 8447–8451. (doi:10.1523/JNEUROSCI.1493-09.2009)
  64. Leaver AM, Van Lare J, Zielinski B, Halpern AR, Rauschecker JP. 2009 Brain activation during anticipation of sound sequences. *J. Neurosci.* **29**, 2477–2485. (doi:10.1523/JNEUROSCI.4921-08.2009)
  65. Cervantes Constantino F, Pinggera L, Paranamana S, Kashino M, Chait M. 2012 Detection of appearing and disappearing objects in complex acoustic scenes. *PLoS ONE* **7**, e46167. (doi:10.1371/journal.pone.0046167)
  66. Barnes R, Jones MR. 2000 Expectancy, ATTENTION, AND TIME. *Cognit. Psychol.* **41**, 254–311. (doi:10.1006/cogp.2000.0738)
  67. Jones MR, Moynihan H, MacKenzie N, Puente J. 2002 Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychol. Sci.* **13**, 313–319. (doi:10.1111/1467-9280.00458)
  68. Jones MR, Johnston HM, Puente J. 2006 Effects of auditory pattern structure on anticipatory and reactive attending. *Cognit. Psychol.* **53**, 59–96. (doi:10.1016/j.cogpsych.2006.01.003)
  69. Falk R, Konold C. 1997 Making sense of randomness: implicit encoding as a basis for judgment. *Psychol. Rev.* **104**, 301–318. (doi:10.1037/0033-295X.104.2.301)
  70. Chait M, Ruff CC, Griffiths TD, McAlpine D. 2012 Cortical responses to changes in acoustic regularity are differentially modulated by attentional load. *Neuroimage* **59**, 1932–1941. (doi:10.1016/j.neuroimage.2011.09.006)
  71. Masutomi K, Barascud N, Kashino M, McDermott JH, Chait M. 2016 Sound segregation via embedded repetition is robust to inattention. *J. Exp. Psychol. Hum. Percept. Perform.* **42**, 386–400. (doi:10.1037/xhp0000147)
  72. Jaramillo S, Zador AM. 2011 The auditory cortex mediates the perceptual effects of acoustic temporal expectation. *Nat. Neurosci.* **14**, 246–251. (doi:10.1038/nn.2688)
  73. Lange K. 2013 The ups and downs of temporal orienting: a review of auditory temporal orienting studies and a model associating the heterogeneous findings on the auditory N1 with opposite effects of attention and prediction. *Front. Hum. Neurosci.* **7**, 263. (doi:10.3389/fnhum.2013.00263)
  74. Ellis RJ, Jones MR. 2010 Rhythmic context modulates foreperiod effects. *Atten Percept Psychophys* **72**, 2274–2288. (doi:10.3758/APP.72.8.2274)
  75. Geiser E, Notter M, Gabrieli JDE. 2012 A corticostriatal neural system enhances auditory perception through temporal context processing. *J. Neurosci.* **32**, 6177–6182. (doi:10.1523/JNEUROSCI.5153-11.2012)
  76. Hughes HC, Darcey TM, Barkan HI, Williamson PD, Roberts DW, Aslin CH. 2001 Responses of human auditory association cortex to the omission of an expected acoustic event. *Neuroimage* **13**, 1073–1089. (doi:10.1006/nimg.2001.0766)
  77. Rimmele J, Schröger E, Bendixen A. 2012 Age-related changes in the use of regular patterns for auditory scene analysis. *Hear. Res.* **289**, 98–107. (doi:10.1016/j.heares.2012.04.006)
  78. Devergie A, Grimault N, Tillmann B, Berthommier F. 2010 Effect of rhythmic attention on the segregation of interleaved melodies. *J. Acoust. Soc. Am.* **128**, EL1–EL7. (doi:10.1121/1.3436498)
  79. Friston K, Adams RA, Perrinet L, Breakspear M. 2012 Perceptions as hypotheses: saccades as experiments. *Front. Psychol.* **3**, 1–20. (doi:10.3389/fpsyg.2012.00151)
  80. Friston K, Rigoli F, Ognibene D, Mathys C, Fitzgerald T, Pezzulo G. 2015 Active inference and epistemic value. *Cogn. Neurosci.* **6**, 187–214. (doi:10.1080/17588928.2015.1020053)
  81. Natan RG, Briguglio JJ, Mwilambwe-Tshilobo L, Jones SI, Aizenberg M, Goldberg EM, Geffen MN. 2015 Complementary control of sensory adaptation by two types of cortical interneurons. *Elife* **4**, 7359. (doi:10.7554/eLife.09868)
  82. Southwell R, Baumann A, Gal C, Barascud N, Friston K, Chait M. 2017 Data from: Is predictability salient? A study of attentional capture by auditory patterns. Dryad Digital Repository. (<http://dx.doi.org/10.5061/dryad.j17h3>)