The influence of temporal context on vision over multiple time scales

Kacie Lee[1] & Reuben Rideaux[1,2]

[1]School of Psychology, The University of Sydney, Camperdown, Australia
[2]Queensland Brain Institute, The University of Queensland, St Lucia, Australia
corresponding author: reuben.rideaux@sydney.edu.au

Conflict of interest statement: The authors declare no competing interests.

Acknowledgements: We thank Arnaldo Bisbal, Leon Zhong, and Cecelia Chenh for assistance with data collection. We also thank Dr William Harrison for their feedback on an earlier version of the manuscript. This work was supported by Australian Research Council (ARC) Discovery Early Career Researcher Awards awarded to RR (DE210100790) a National Health and Medical Research Council (NHMRC; Australia) Investigator Grant (2026318).

Keywords: vision, predictive coding, serial dependence, sequential dependencies, statistical learning, temporal effects, EEG, pupillometry.

ABSTRACT

Past sensory experiences influence perception of the present. Multiple research subfields have emerged to study this phenomenon at different temporal scales. These fall into three categories: the influence of immediately preceding sensory events (micro), short sequences of events (meso), and regularities over long sequences of events (macro). In a single paradigm, we examined the influence of temporal context on perception at each scale. We identify two distinct mechanisms that operate across all scales. The first is moderated by attention and supports rapid motor responses to expected events. The second is independent of task-demands and dampens the feedforward neural responses to expected events, leading to unexpected events eliciting earlier and more precise neural representations. We further show that perceptual recall exclusively reflects neural representations during this initial feedforward stage and that serial dependence (recall biases towards previous events) is explained by expectation of sensory stability over time.

INTRODUCTION

The capacity to adapt to patterns in the environment supports biological function from sensory processing to motor action. The temporal context in which sensory events are embedded can be leveraged to more effectively process and respond to this information. For example, expert tennis players predict the trajectory of the ball from the movements of their opponent and use this information to prepare their return[1]. Influential theories of normative brain function such as predictive coding posit that temporal context serves to improve representational fidelity and reduce neurometabolic expenditure[2–4].

Humans are sensitive to the temporal context of events across multiple time scales. At the shortest scale, each event influences processing of the next event (**Fig. 1a**, micro). For example, stimulus reproductions are biased towards previous stimuli. This phenomenon is referred to as *serial dependency* and has been demonstrated for a range of visual and auditory features[5–7]. At intermediate scales, short sequences of events form patterns that uniquely influence responses to subsequent events (**Fig. 1a**, meso). For instance, events that satisfy regular patterns produce attenuated neural responses (*oddball effect*)[8] and are responded to more quickly and accurately than events that violate them (*sequential dependencies*)[9,10]. At longer time scales, the relative frequency of past events can alter how those in the present are processed (**Fig. 1a**, macro). These effects are often referred to as *statistical learning* and are thought to reflect adaptation to regularities in the environment[11,12].

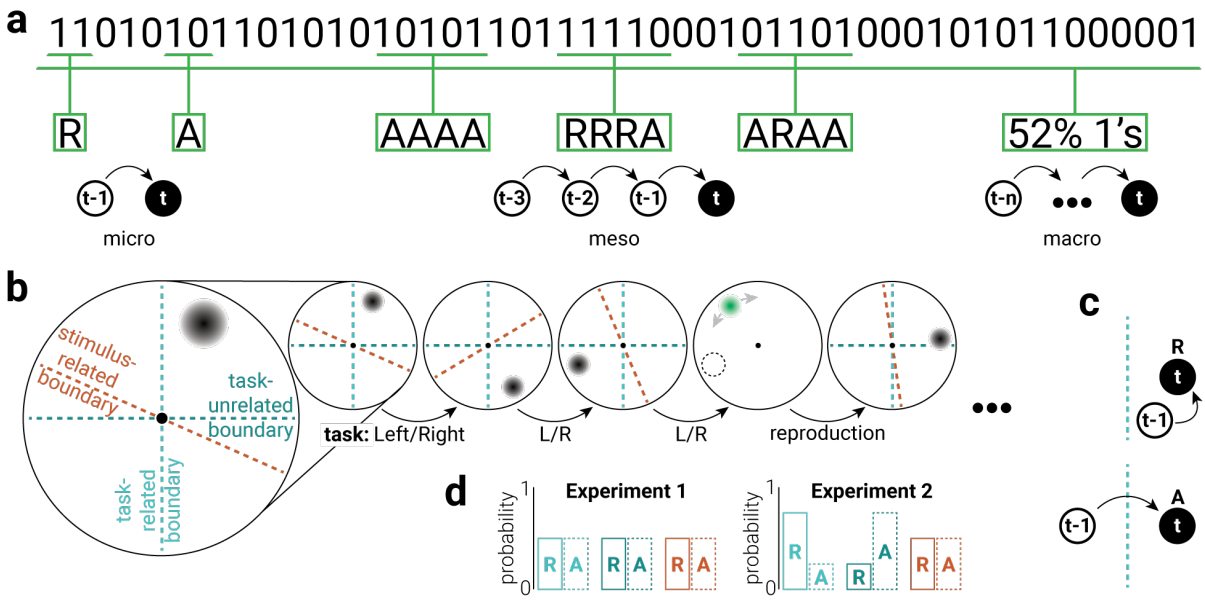


Figure 1. Multiple scales of temporal context. a) Illustration of three scales of temporal context in a binary time series. At the micro scale, events can either stay the same over time (repeat; R) or change (alternate; A). At the meso scale, short sequences of events (e.g., 5) form patterns that vary in their regularity (e.g., four repeats of the same event seems more regular than a mixture of repeats and alternations between events). At the macro scale, general trends regarding the relative frequency of different events are formed over longer time periods. **b**) To test the influence of temporal context on visual perception across different scales, participants were instructed to indicate the location of serially presented targets, which were randomly positioned on an imaginary circle centered on fixation. Participants performed a speeded binary judgement (e.g., left/right of fixation) on each trial and additionally reproduced the location of the target on 10% of trials. Trials were categorized as either (**c**, top) repeat (R) or (**c**, bottom) alternate (A) based on the location of the target relative to the previous target, according to three spatial reference planes: task-related (light cyan), task-unrelated (dark cyan), and stimulus-related (orange). **d**) In Experiment 1, the location of targets was uniformly sampled such that repeat and alternate trials were equally likely. In Experiment 2, the probability of repeat and alternate trials was biased across the task-related and unrelated planes.

Incoming sensory information is shaped by its temporal context across time scales that range many orders of magnitude. Previous work has developed a variety of experimental designs to isolate the influence of temporal context at each scale and examine its behavioural and neural consequences. This approach has propagated multiple fields of research and implicitly asserts qualitive differences, however, it limits comparison between temporal context at different scales. It is possible that temporal context serves perception differently at each time scale and is thus associated with unique adaptive influences on sensory processing. For example, perception may be best supported by one mechanism that promotes a set of outcomes from moment to moment and another mechanism that promotes a different set of outcomes in response to trends formed over long periods. Alternatively, a common rule may be

applied at all scales of temporal context, signalling a unifying adaptive mechanism. Predictive coding theory offers a normative framework with which to understand the influence of temporal context across scales, but contradictory evidence[3,13–15] (i.e., sharpening or dampening of expected events) from distinct experimental paradigms has challenged this unification.

We designed a novel paradigm to investigate how temporal context shapes perception across multiple scales. We further used electroencephalography (EEG) and pupillometry recordings to characterize the neural mechanisms associated with these perceptual consequences. We identify two distinct mechanisms that operate across all scales. The first is moderated by attention and supports rapid motor responses to expected events. The second is independent of task-demands and dampens the feedforward neural responses to expected events, leading to unexpected events eliciting earlier and more precise neural representations. Together, these canonical adaptive mechanisms explain a raft of temporal sensory phenomena, including serial dependence, and adjudicate competing accounts of predictive coding.

RESULTS

To measure changes in visual perception associated with different scales of temporal context, we tasked participants with indicating the location of serially presented visual stimuli (gaussian blobs randomly positioned at a fixed distance around a central fixation point). To assess response time and accuracy, participants performed a speeded binary judgement (e.g., left or right of fixation) on each trial. On 10% of trials, participants additionally reproduced the location of the stimulus, providing a measure of recall precision. To test the influence of attention, trials were sorted according to two spatial reference planes, based on the location of the stimulus: task-related and task-unrelated (**Fig. 1b**). The task-related plane corresponded to participants' binary judgement (e.g., left or right of fixation) and the task-unrelated plane was orthogonal to this (e.g., above or below fixation).

Micro-scale temporal context. At the shortest temporal scale, the speed with which events are responded to can be influenced by those that immediately precede them[16]. We investigated the influence of the previous stimulus on the current stimulus by comparing responses between trials in which the previous stimulus was presented either on the same side (repeat) or the other side (alternate) of the reference plane (**Fig. 1c**, **2a**). Consistent with previous work[16], we found faster and more accurate responses to repeat stimuli along the task-related plane (speed: z_{75}=2.33, p=.020; accuracy: z_{72}=2.41, p=.016; **Fig. 2b-c**, light cyan). However, in contrast, we found that repeat stimuli were recalled less precisely than alternate stimuli (z_{72}=3.74, p=2.00e^{-4}; **Fig. 2d**, light cyan). For the task-unrelated plane, we found faster responses for repeat stimuli, but no difference in accuracy, and marginally higher precision for alternate stimuli (speed: z_{75}=6.66, p=2.76e^{-11}; accuracy: z_{72}=0.92, p=.357; precision: z_{76}=1.62, p=.105; **Fig. 2b-d**, dark cyan).

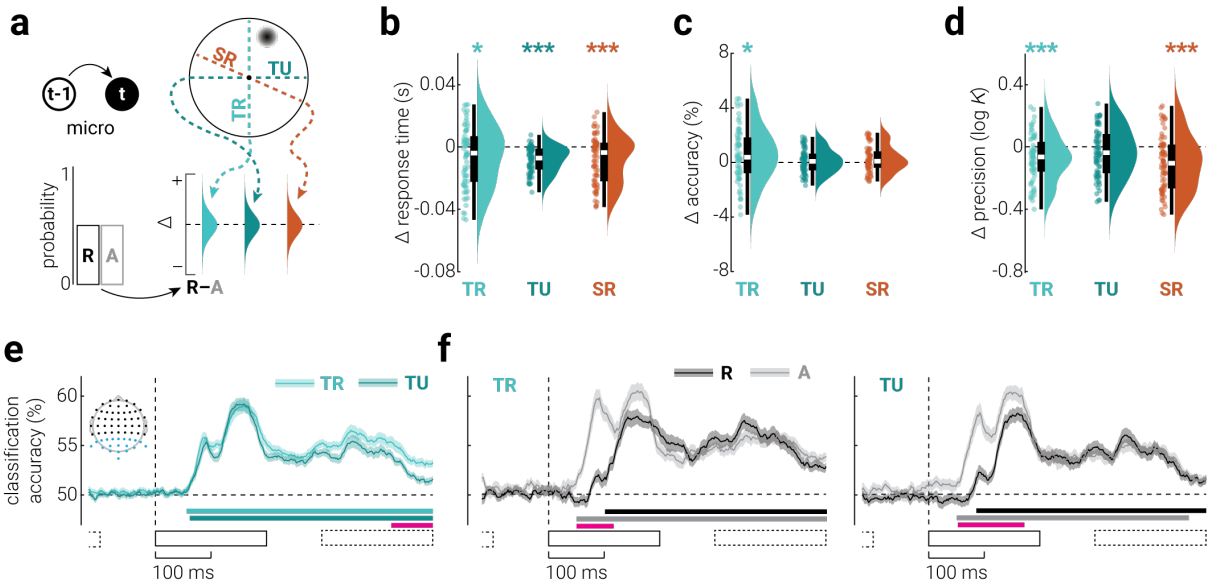


Figure 2. The influence of micro-scale temporal context on visual processing. a) Micro temporal context refers to the influence of the last event on the current event. We assessed its influence by comparing task performance for repeat and alternate presentations along task-related (light cyan), task-unrelated (dark cyan), and stimulus-related (orange) reference planes. **b-d**) The difference (repeat – alternate) in (**b**) response time, (**c**) accuracy, and (**d**) precision for the three references planes. Asterisks indicate significant differences (**p*<.05, ***p*<.01, ****p*<.001). **e**) Classification accuracy of stimuli presented on different sides of task-related and unrelated planes, from re-analysis of previously published EEG data[17]. **f**) Same as (**e**), but split into repeat and alternate stimuli, along (left) task-related and (right) unrelated planes. Inset in (**e**) shows the EEG sensors included in the analysis (blue dots). Black rectangles indicate the timing of stimulus presentations (solid: target stimulus, dashed: previous and subsequent stimuli). Shaded regions indicate ±SEM. Horizontal bars indicate cluster corrected periods of significance (cyan and greyscale: above chance accuracy, pink: difference).

While the difference in response time for repeat and alternate stimuli were of similar magnitude between task-related and unrelated planes (z_{71}=1.15, p=.249), the *variance* across participants was considerably higher in the former condition ($F_{1,140}$=46.26, p=2.76e^{-10}). The increased variability of response time differences across the task-related plane likely reflects individual differences in attention and prioritization of responding either quickly or accurately. On each trial, the correct response (e.g., left or right) was equally probable. So, to perform the task accurately, participants were motivated to respond without bias, i.e., without being influenced by the previous stimulus. We would expect this to reduce the difference in response time for repeat and alternate stimuli across the task-related plane, but not the task-unrelated plane. However, attention may amplify the bias towards making faster responses for repeat stimuli, by increasing awareness of the identity of stimuli as either repeats or alternations[17]. These two opposing forces vary with task engagement and strategy and thus would be expected produce increased variability

across the task-related plane. Unlike response time, we would not expect accuracy to be modulated by task-unrelated micro temporal context, as the tasks were orthogonal. That is, while a repeat along the task-unrelated plane may increase the speed with which an (in/correct) response is made along the task-related plane, it would not be expected to change the type of response (e.g., left or right). Precision differences were similar between task-related and unrelated planes, both in terms of magnitude (z_{70}=1.77, p=.076) and variability ($F_{1,138}$=0.66, p=.417), suggesting that modulation of representational fidelity may be operationalized by a distinct mechanism to that which modifies response speed, and is not moderated by attention.

These results could be interpreted as a general expectation that sensory events, e.g., the image of an object on the retina, are more likely to stay the same than to change[7]. Through this lens, these findings would indicate that micro temporal context leads to faster and more accurate responses to expected events, but more precise encoding of unexpected events. As we will show, these two outcomes resonate across all levels of temporal context and point to two distinct mechanisms that unify the influence of past experiences on current perception.

To test for neural correlates of micro temporal context, we re-analyzed a previously published EEG dataset in which human participants viewed visual (arc) stimuli presented at random angles around fixation, while monitoring for targets (stimuli presented at a specific location)[17]. We used multivariate linear-discriminant analysis to classify the location of stimuli according to task-related (< or > |90°| from target angle) and unrelated planes (orthogonal to task-related plane), from parietal and occipital sensors (**Supplementary Fig. 1a**). We found that classification accuracy was higher across the task-related plane (**Fig. 2e**).

In separately assessing classification accuracy for repeat and alternate stimuli, we found temporal dependencies that were unrelated to stimulus properties, as evidenced by their presence in results produced from data with shuffled labels (**Supplementary Fig. 2**). We removed these general temporal dependencies by subtracting the bias in the shuffled data from the original accuracy. For both task-related and unrelated planes, we found that classification accuracy was higher for alternate stimuli from ~60-160 ms (i.e., the first moments of reliable stimulus

decoding) and that alternate stimuli could be reliably classified ~20 ms earlier than repeat stimuli (**Fig. 2f**). The difference in decoding accuracy between repeat and alternate stimuli was the same for task-related and unrelated conditions. This neural phenomenon may explain the increased behavioural recall precision observed for alternate stimuli. The timing and task-independence of this phenomenon points to a passive, massively parallel, mechanism that operates on feedforward sensory signals, by supressing responses to repeat stimuli.

Another phenomenon associated with micro temporal context is *serial dependence*, i.e., reproductions of events from memory are biased towards the events that immediately preceded them[5,7]. To assess serial dependence in the behavioural data, we began by comparing behavioural performance across a third, stimulus-related, plane (**Fig 2a**, orange); whereas task planes were constant, the stimulus-related plane was defined by the location of the stimulus on the previous trial, and thus varied from trial to trial. We found a similar pattern of results as those for task-related and unrelated planes: responses were faster for stimuli that repeated (z_{77}=3.73, p=2.00e^{-4}; **Fig. 2b**, orange) but more precise for those that alternated (z_{76}=4.51, p=6.34e^{-6}; **Fig. 2d**, orange). The latter result seems inconsistent with previous work on serial dependence, which has reported more precise judgements when subsequently presented stimuli are similar[18].

To further investigate the influence of micro temporal context, we computed task performance as a function of the angular distance from the previous stimulus (*Δ location*). We calculated the bias of reproduction responses as a function of Δ location and found the archetypal pattern of attractive biases associated with serial dependence (**Fig. 3a**). Consistent with our previous binary analysis, we also found that responses were faster and more accurate when Δ location was small (**Fig. 3b**, **c**). We further found that the precision of responses peaked at both large and small Δ locations, with the worst precision for stimuli with around ±60° Δ location (**Fig. 3d**). The peak in precision for large Δ locations is consistent with alternate events being encoded more precisely, while the peak for small offsets may be explained by integration of location and Δ location representations (**Supplementary Fig. 3**).

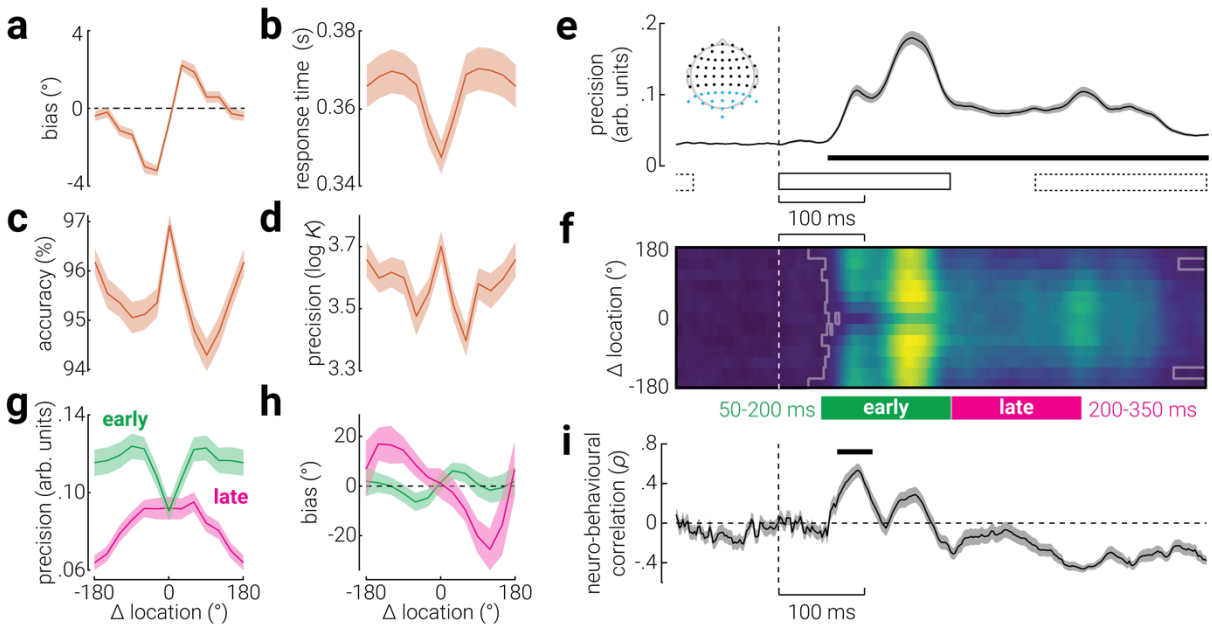


Figure 3. The influence of serial dependence on visual processing. Serial dependence is associated with sensory stimuli being perceived as more similar to previous stimuli and is typically assessed by measuring the perception of stimuli as a function of their distance to previous stimuli (Δ location). **a**) Location reproduction bias, (**b**) binary task response time and (**c**) accuracy, and (**d**) reproduction precision as a function of distance from the previous stimulus. **e**) Decoding precision for stimulus location, from re-analysis of previously published EEG data [17]. Inset shows the EEG sensors included in the analysis (blue dots), and black rectangles indicate the timing of stimulus presentations (solid: target stimulus, dashed: previous and subsequent stimuli). **f**) Decoding precision for location, as a function of time and Δ location. Bright colours indicate higher decoding precision; absolute precision values can be inferred from (**e**). Cluster corrected periods of decoding precision that is significantly above values from a shuffled comparison are bordered by semi-transparent white lines. Early (50-200 ms) and late (200-350 ms) periods are illustrated in green and pink, respectively. **g-h**) Average (**g**) location decoding precision and (**h**) bias during the early (green) and late (pink) period. **i**) Correlation between behavioural bias and location decoding bias as a function of time. Horizontal bars in (**e** & **i**) indicate cluster corrected periods of significance. Note, the temporal abscissa is aligned across (**e**, **f**, & **i**). Shaded regions indicate ±SEM.

To further investigate the influence of serial dependence, we applied inverted encoding modelling to the EEG recordings to decode the angular location of stimuli. We found that decoding precision of stimulus location was reliably above chance from ~60 ms following stimulus onset (**Fig. 3e**; **Supplementary Fig. 1b**, **c**). Note, we found a similar pattern of results for decoding *accuracy*, but decoding accuracy was confounded by the same general temporal dependency observed in the previous linear classification analysis, while decoding *precision* was not (**Supplementary Fig. 3a-l**). To understand how serial dependence influences the representation of these features, we inspected decoding precision for location as a function of both time and Δ location (**Fig. 3f**). We found distinct patterns of decoding precision between locations over time. Initially (early: 50-200 ms following stimulus onset; **Fig. 3g**,

green), location decoding precision was worst for small Δ locations, but then later became best (late: 200-350 ms; **Fig. 3g**, pink). We calculated the average decoded locations and found a pattern of attractive biases towards the previous stimulus during the early period (**Fig. 3h**, green). In contrast, during the late period, the average decoded locations exhibited a broad repulsive bias away from the previous stimulus (**Fig. 3h**, pink). Whereas the early attractive bias in consistent with serial dependence, the later repulsive bias is consistent with previous fMRI work on serial dependence[19], and likely reflects general temporal dependency in neural recordings, rather than adaptation (**Supplementary Fig. 3m**).

To assess the relationship between the neural representation and perception, we computed the correlation between the average recall bias (**Fig. 3a**) and the bias in the decoded location over time. We found a single significant period of neuro-behavioural correspondence that aligned with the emergence of stimulus decoding (60-100 ms following stimulus onset; **Fig. 3i**), suggesting a primacy of feedforward processing in shaping perceptual recall.

Meso-scale temporal context. The influence of recent events can accrue over time, producing unique changes in behavioural responses associated with short sequences of stimuli (**Fig. 4a**). Higher-order sequential dependences are an example of how stimuli (at least) as far back as five events in the past can shape the speed and accuracy of responses to the current stimulus[9,10]. For five binary events (e.g., left or right), there are 16 possible sequences of repeat and alternate stimuli. Typically, sequential dependencies are assessed by computing the average response time and/or accuracy for stimuli, binned according to which of these 16 sequences reflects their recent history. The differences observed between sequences have been interpreted as indicative of the expectation of the current stimulus, based on previous events[9].

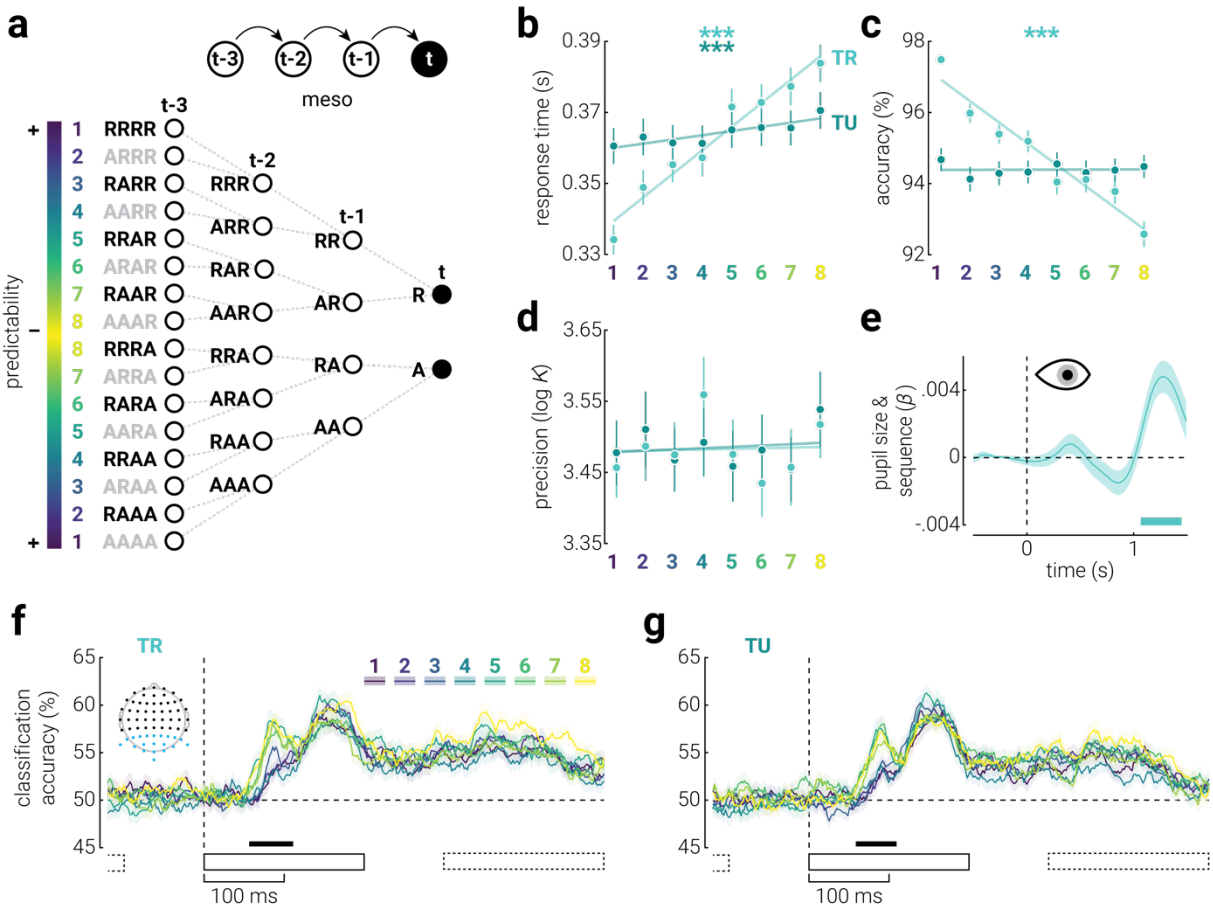


Figure 4. The influence of meso-scale temporal context on visual processing. a) Illustration of all 16 possible sequences of repeat and alternate events for a series of five binary events. Sequences are arranged symmetrically such that those for which the final event is thought to be most expected are at the top and bottom and those for which the final event is least expected are in the middle. To control for differences between repeat and alternate events, we combined symmetric pairs, resulting in a total of eight sequences. **b**) Response time, (**c**) accuracy, and (**d**) precision as a function of sequence, across task-related and unrelated reference planes. Note, lower numbers on the abscissa are associated with sequences in which the final target stimulus is more expected. Asterisks indicate significant main effects of sequence (***p<.001). Error bars indicate ±SEM; semi-transparent lines indicate linear fits to the data. **e**) The correlation between pupil size and sequence as a function of time. Note, we did not analyze pupillometry data for micro temporal context due the confounding effect of differential foveal luminance between repeat and alternate stimuli. **f**) Classification accuracy of stimuli presented on different sides of task-related planes as a function of time, for each of the eight sequences, from re-analysis of previously published EEG data [17]. **g**) Same as (**f**), but for the task-unrelated plane. Inset in (**f**) shows the EEG sensors included in the analysis (blue dots). Black rectangles indicate the timing of stimulus presentations (solid: target stimulus, dashed: previous and subsequent stimuli). Shaded regions indicate ±SEM. Horizontal bars indicate cluster corrected periods of significant relationships between (**e**) pupil size or (**f, g**) classification accuracy and sequence order.

We tested the influence of higher-order sequential dependencies on response time, accuracy, and recall precision, across task-related and unrelated reference planes. However, we first isolated the influence of meso temporal context from that of micro context (i.e., difference in performance between repeat and alternate stimuli), by binning responses into eight sequences formed by combining the symmetric

partners in the 16-sequence array (e.g., RRRR with AAAA, RRRA with AAAR, etc.). Thus, the frequency of repeat and alternate stimuli was equal across the eight sequences and any differences observed in performance are unlikely to be due to differential responses to these events. We found that responses were faster for sequences with expected targets across both reference planes (task-related: $\chi^2_{7,469}$=375.86, p=3.56e^{-77}; task-unrelated: $\chi^2_{7,483}$=110.94, p=5.85e^{-21}; **Fig. 4b**). Responses were also more accurate for expected targets across the task-related plane ($\chi^2_{7,399}$=208.37, p=2.03e^{-41}; **Fig. 4c**). By contrast, we found no such clear relationship between precision and expectancy (both p>.05; **Fig. 4d**).

We found that micro temporal context influenced recall precision, whereas we did not detect significant modulation of precision for meso context. This may be because meso temporal effects reflect motor priming and do not precipitate visual predictions. When predictions are violated, the pupils dilate in what has been interpreted as an arousal response[20,21]. We found that participants' pupils dilated more in response to unexpected sequences (**Fig. 4d**), indicating that meso temporal context establishes sensory predictions. Another possible explanation why we did not detect an effect on recall precision is that we lacked statistical power. The micro analyses split data between two conditions (repeat and alternate), whereas in the meso analysis, data is split between eight (sequences). Indeed, in our second experiment, we reproduced the response time and accuracy effects of meso temporal context and further found that unexpected sequences were recalled more precisely (**Supplementary Fig. 4**).

To further test whether meso temporal context effects were qualitatively different from those at the micro scale, we separated the EEG classification accuracies (**Fig. 2e**) into the eight sequence categories and corrected for shuffled accuracy (**Supplementary Fig. 5**). We found a strikingly similar pattern of results to those observed at the micro scale: there was increased accuracy for unexpected sequences from ~60-100 ms along both task-related and unrelated reference planes (**Fig. 4h**, **g**). This finding provides further evidence of a passive mechanism that suppresses feedforward sensory activity elicited by expected stimuli and leads to more precise recall of those that are unexpected. That the effect was reproduced

when repeat and alternate stimuli were matched also shows that it cannot be explained by sensory adaptation[22].

Macro-scale temporal context. Events within micro and meso temporal contexts occur within the limits of short-term memory. Most events that occur further into the past cannot be recalled; yet, they continue to influence perception[11,12]. To test the influence of macro temporal context on perception, we ran a second experiment in which the probability of stimulus presentation was biased towards repeats or alternates along task-related and unrelated reference planes (**Fig. 1d**, Experiment 2). The biases were counterbalanced across participants such that we could compare performance between relatively frequent (expected) and infrequent (unexpected) events, while controlling for differences between repeat and alternate stimuli (**Fig. 5a**). Consistent with micro and meso scale effects, we found that expected events were responded to faster and more accurately, while unexpected events were reproduced more precisely, across both task-related (speed: z_{78}=7.67, p=1.68e^{-14}; accuracy: z_{73}=7.42, p=1.13e^{-13}; precision: z_{79}=3.26, p=.001) and unrelated planes (speed: z_{71}=6.69, p=2.19e^{-11}; accuracy: z_{78}=6.76, p=1.32e^{-11}; precision: z_{75}=2.72, p=.007; **Fig. 5b-d**). This effect was larger along the task-related plane for response time and accuracy, but not precision (speed: z_{71}=7.24, p=4.42e^{-13}; accuracy: z_{78}=7.27, p=3.52e^{-13}; precision: z_{75}=0.93, p=.350). Pupillometry analysis further revealed that participants' pupils dilated more in response to unexpected stimuli ~1 s after stimulus onset (**Fig. 5e**).

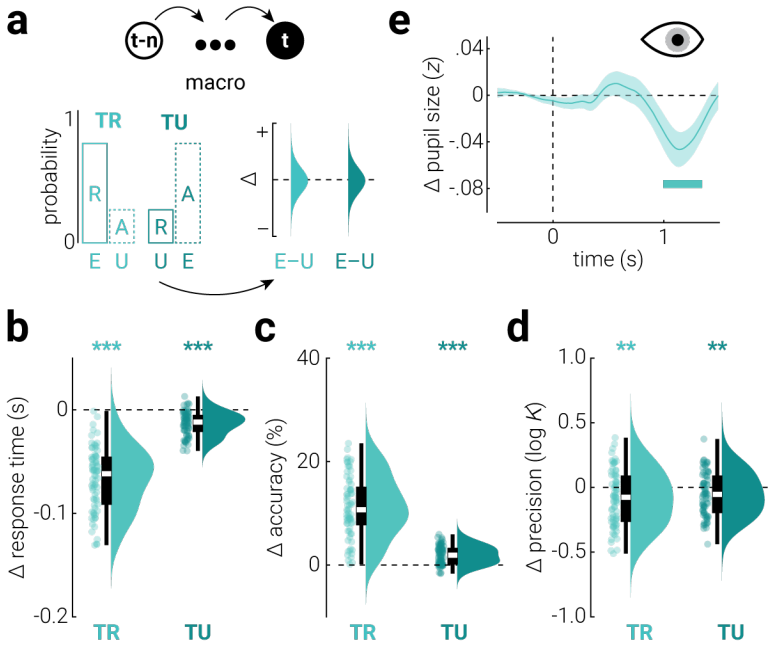


Figure 5. The influence of macro temporal context on visual processing. a) Repeat (R) and alternate (A) stimuli were presented with unequal probabilities (counterbalanced across participants) along task-related and unrelated reference planes. The difference in performance between relatively frequent (expected; E) and infrequent (unexpected; U) stimuli was calculated to assess the influence of macro temporal context on visual processing. **b-d**) The difference (expected – unexpected) in (**b**) response time, (**c**) accuracy, and (**d**) reproduction precision between expected and unexpected stimuli, along task-related (TR) and unrelated (TU) planes. Note that we observed accuracy differences associated with macro temporal context across the task-unrelated plane. In the micro and meso analyses, the expected outcome along the task-unrelated plane was never in conflict with that in the task-related plane. For example, if a stimulus is expected on the same (right) side of the display (repeat; task-related plane), then regardless of whether the stimulus appeared above or below fixation, there was a quadrant of the display that would satisfy both the task-related and unrelated expectations (bottom- or top-right). This is because the expectation was either for repeat stimuli (micro) or balanced between repeat and alternate stimuli (meso). By contrast, in the macro condition, where stimuli could be expected to alternate, this produced trials on which the task-related and unrelated expectations conflicted, such that there was no location that could satisfy both. Accuracy along the task-unrelated plane was reduced in these instances where there was conflict. Asterisks indicate significant differences (**p<.01, ***p<.001). **e**) The difference (expected – unexpected) in pupil size as a function of time from stimulus presentation, along the task-related (TR) plane. Shaded region indicates ±SEM and the horizontal bar indicates a cluster corrected period of significant difference.

In line with previous work, stimuli in our experiment were displayed until participants responded. To test how variable presentation duration influenced the results, we ran a third experiment that was the same as the second, except with a fixed stimulus duration (200 ms). Experiment 3 replicated both the behavioural and pupillometry results from Experiment 2 (**Supplementary Fig. 6**), indicating that the effects observed were not related to variable stimulus duration.

DISCUSSION

Evidence from multiple, separate, fields of research show that humans are sensitive to the temporal context in which sensory stimuli are embedded across a broad range of time scales. Here we used a novel experimental paradigm to characterise and compare how past sensory events occurring at different temporal scales shape perception of the present. By combining behavioural, neuroimaging, and physiological approaches to assess the influence of temporal context at the micro, meso, and macro scale, we identified new properties of these phenomena in isolation and further revealed two distinct adaptive mechanisms that unify them.

Micro-scale temporal context. At the micro scale, stimuli were responded to faster and more accurately when preceded by those that were similar, but recalled more precisely when preceded by those that were different. This may reflect a general expectation that objects in the environment are more likely to stay the same than to change[7]. Under this interpretation, expected events are rapidly responded to while unexpected events are precisely encoded. As will be discussed shortly, this interpretation resonates with the influence of temporal context at longer scales.

Analysis of EEG recordings showed that during the initial stage of stimulus processing in the cortex (60 ms after onset), alternating stimuli were represented both (20 ms) earlier and more robustly than those that repeat. This finding provides a neural correlate of the improved recall precision for these stimuli and, consistent with recent psychophysical work[23], suggests that unexpected (alternating) events are prioritized in terms of both encoding fidelity and processing speed. Given that unexpected stimuli were responded to more slowly than expected stimuli, despite benefitting from sensory processing priority, this may indicate that the latter is supported by anticipatory motoric activity[24].

Effects on binary judgements associated with micro temporal context, such as those described above, have been referred to as first-order sequential dependencies[16]. By contrast, those associated with continuous report tasks are typically referred to as serial dependence[5,7]. We examined behavioural responses for continuous effects and found the stereotypical pattern of response biases associated with serial dependence, i.e., attraction towards the preceding stimulus.

Our corresponding EEG analyses revealed two distinct stages of serial dependence: an early stage (50-200 ms), where the neural representation was least precise for stimuli preceded by those that were similar and biased towards the previous stimulus, and a later stage (200-350 ms), where the representation was most precise for successive similar stimuli and biased away from the previous stimulus. Our neuro-behavioural correlations showed that the bias observed in the neural representation uniquely matched the bias observed in behavioural responses during a period aligned with the initial decoding peak at ~60 ms.

These findings indicate that serial dependence and first-order sequential dependencies can be explained by the same underlying principle. In particular, sensory events are generally expected to stay the same, rather than change. Behavioural and neural recordings show a predisposition for expected outcomes, which manifest as faster, biased, behavioural responses. By contrast, unexpected outcomes are encoded earlier and more precisely, leading to increased recall fidelity for these events.

More broadly, our findings provide converging lines of evidence that perceptual reports more closely reflect the neural representation during the initial feedforward stage of the sensory processing cascade. In particular, the difference in recall precision for repeat and alternate stimuli was only evident during this period (60-100 ms following stimulus onset), and the behavioural bias associated with serial dependence matched corresponding measures of the neural representation at this time. However, there is stimulus-related information that can be reliably decoded from brain activity after this period. For instance, in the serial dependency analysis, we found a repulsive pattern of biases during the late stage, which was inconsistent with behaviour. While more work is needed to understand the relevance of the neural representation during this later period, its existence may explain contradictory evidence from previous fMRI work showing a repulsive bias in BOLD activity associated with serial dependence[19], like that observed in the late stage here. Our findings suggest that these results may reflect stimulus-unrelated temporal context.

Meso-scale temporal context. At the meso scale, consistent with previous work[9,10], we found that short sequences of stimuli were responded to faster and more accurately when they were expected. Extending these findings, and matching the

influence of micro-scale temporal context, we also found that responses were more precise for unexpected sequences (in Experiment 2, and non-significantly in Experiment 1). Pupillometry analyses confirmed that participants were surprised by stimuli that were categorized as unexpected, and EEG analyses revealed a strikingly similar pattern of results as was found at the micro scale. In particular, unexpected stimuli were encoded earlier and more robustly. The similarity between the influence at the micro and meso scale is particularly remarkable given that the influence of micro temporal context was controlled for in the meso-scale analyses.

Theoretical and empirical work has implicated both motor and sensory priming in the modulation of response time and accuracy associated with higher-order sequential dependencies[24,25]. Decoupling the motoric component of experimental designs used to probe the phenomenon has been an obstacle to assessing how sensory processing is influenced by meso-scale temporal context. In the EEG experiment that was re-analyzed here, participants viewed serially presented stimuli and were tasked with reporting the number of targets after 30 presentations. Thus, there were no motor actions generated during the recording period used to perform the analysis and our results can only be attributed to differences in visual processing between sequences. Indeed, the timing of the improved encoding of unexpected sequences, like that for micro temporal context, can only be accounted for by a mechanism that prioritizes the encoding speed and fidelity of stimuli during the initial feedforward cascade of sensory processing.

Macro temporal context. Results from our second experiment, in which biases in the probability of repeat and alternate stimuli were introduced, showed that responses are faster and more accurate for expected stimuli, and that recall precision is better for unexpected stimuli. Pupillometry analysis confirmed that participants were surprised by unexpected stimuli. We replicated these results in a third experiment, demonstrating that the same outcome for stimuli of fixed duration. These findings parallel those observed at the micro and meso scales and point to two canonical adaptive mechanisms that operate across a wide range of temporal contexts to support a common sensory outcome.

The role of attention. At each scale, we found that attention (operationalized through task-relatedness) amplified the influence of temporal context for response time and

accuracy, but not precision. These findings suggest that perception and behaviour are modulated by expectation through two distinct mechanisms. One mechanism is sensitive to top-down goals (attention) and supports rapid responses to expected stimuli, while the other mechanism automatically detects mismatches between expected and actual outcomes and prioritizes the encoding speed and fidelity of unexpected stimuli. In line with this, some studies have reported interactions between attention and predictive mechanisms[26,27], while others have shown that oddball event-related potentials can be detected in coma patients[28].

Consistent with previous work[17], we found that the influence of attention on the neural representation only emerged after ~200 ms. By contrast, we found that effects of expectation, for both micro- and meso-scale temporal context, were present during the earliest moments of cortical sensory processing, i.e., during the feedforward cascade. These findings suggest that predictive mechanisms are operational during feedforward processing, while attention modulates sensory processing at a later stage.

Predictive coding. Predictive coding theories argue that *prediction errors* are generated when bottom-up sensory inputs deviate from top-down expectations[3]. There is broad consensus that prediction errors are associated with increased neural activation[13,29–33], which is typically observed during the initial processing cascade in sensory cortices associated with the unexpected stimulus[34,35], but has also been reported in subcortical regions[20]. By contrast, whether this modulation of neural activation results from 'dampening' or 'sharpening' of the neurons tuned to expected features remains highly contested[3,13–15]. Whether through increased selectively (sharpening) or reduced sensitivity (dampening), both accounts predict attenuated activation in response to expected events. However, according to the sharpening model, expected events are represented more precisely, while the dampening model posits the opposite, i.e., unexpected events are more precisely represented. At every scale of temporal context and across behavioural and neural recordings, we consistently found that unexpected events were encoded more precisely than expected events. These findings provide compelling evidence that predictive mechanisms reduce the sensitivity of neurons tuned to expected features, likely in order to reduce metabolic energy expenditure[3], rather than increasing their selectively.

The expectation of stimuli shaped their neural representation in a robust and striking manner. In particular, the peak in decoding accuracy from 50-150 ms following stimulus onset, which represents the initial stage of (feedforward) cortical representation, was almost entirely accounted for by stimulus expectation such that expected stimuli were only minimally represented during this period. This modulation cannot be attributed to adaptation, as we found the same result at the meso scale, where the change in location between subsequent stimuli was matched between conditions. Further, we found an attractive bias during this period, as opposed to the repulsive bias associated with adaptation. There was relatively weak evidence for a difference between the neural representation of expected and unexpected stimuli following this initial peak. Given that unexpected stimuli were recalled more precisely in the behavioural task, this provides further support to the notion that recall of these stimuli primarily reflects neural activity during this initial stage of sensory processing; however, more work is needed to understand how this relationship changes with increasing stimulus complexity. These findings provide empirical evidence that predictive mechanisms operate to reduce the feedforward sensory activity associated with expected stimuli in a manner that reduces their representational fidelity (i.e., dampening/cancellation).

Prioritizing the speed and fidelity with which unexpected events are encoded likely serves to rapidly and precisely update internal predictive models when there is a mismatch between predictions and incoming sensory information. By contrast, expected events may benefit from anticipatory motoric activation, which facilitates rapid behavioural responses, and do not require encoding prioritization.

Conclusion. Past sensory information shapes perception of the present. Multiple research areas have emerged to investigate the isolated influence of temporal context at different scales. In a single experimental paradigm, we compared the influence of past stimuli at micro, meso, and macro temporal scales, and found two common rules that characterized the relationship between the past and present across all ranges. While the neural mechanisms that implement the influence of temporal context across different scales may vary (e.g., tonic inhibition/excitation, synaptic weight), they appear to promote a two unifying outcomes: 1) rapid motor responses to expected events, and 2) prioritized encoding speed and fidelity of unexpected events.

508 **METHODS**

509 ***Participants.*** Eighty neurotypical human adults participated in Experiments 1 and 2
510 (mean±standard deviation age, Experiment 1: 20.9±2.9 years, 14 males, 65 females,
511 1 non-binary, Experiment 2: 21.0±5.8 years, 18 males, 62 females), and 40
512 neurotypical human adults participated in Experiment 3 (24.7±8.3 years, 10 males, 30
513 females. The data of one participant was omitted from Experiment 1 because they
514 were unable to complete the experiment. Sample size was informed by previous
515 studies using similar psychophysical methods[23]. Participants were recruited from The
516 University of Sydney and had normal or corrected-to-normal vision (assessed using a
517 standard Snellen eye chart). All participants were naïve to the aims of the experiment
518 and gave informed written consent. The experiment was approved by The University
519 of Sydney Human Research Ethics Committee.

520 ***Apparatus.*** The experiment was conducted in a dark acoustically shielded room. The
521 stimuli were presented on a 41.5-inch ASUS ROG Swift OLED monitor with 3840 x
522 2160 resolution and a refresh rate of 120 Hz. Viewing distance was maintained at 1 m
523 using a chin and head rest. Stimuli were generated in MATLAB v2020a (The
524 MathWorks, Inc., Matick, MA) using Psychophysics Toolbox[36,37] v3.0.18.13 (see
525 http://psychtoolbox.org/). Gaze direction and pupil diameter were recorded
526 monocularly (right eye) at 1 kHz using an EyeLink 1000 (SR Research Ltd., Ontario,
527 Canada).

528 ***Stimuli, task, and procedure.*** The stimuli comprised white Gaussian blobs (standard
529 deviation=0.3°, contrast=0.5) positioned 4° of visual angle from fixation (location
530 randomly selected between 0-360°) on a black background. A centrally positioned
531 white fixation dot (radius=0.25°) was presented to reduce eye movements. Stimuli
532 were presented until the participant responded. Half the participants were instructed
533 to use the "z" and "/" keys to indicate whether the stimulus appeared on the left or right
534 of fixation, and the other half used the "b" and "t" keys to indicate whether it appeared
535 above or below fixation; the keys were selected to map the relative positions on the
536 keyboard to the orientation of the task. Participants performed 15 to 20 blocks of 100
537 trials (~90 min), receiving feedback on their response time and accuracy at the end of
538 each block. On 10% of randomly selected trials, after participants responded, the

target blob was replaced by a green blob at a random location (same size as target blob), and participants were instructed to use the response keys to rotate the location of the green blob around to the location of the previous target blob and press the "spacebar" key to confirm their response.

Experiment 2 was the same as Experiment 1, except that biases were introduced to the location of the target blob. In particular, sequentially presented target stimuli were more likely to be presented on the same side (75%; repeat) than on the other side (25%; alternate), or vice versa, along each of the task reference planes (left/right and above/below). Thus, a 2 (task reference plane) × 2 (left/right bias) × 2 (above/below bias) design was used and the eight conditions were counterbalanced across participants. On each trial, based on the location of the target on the previous trial and the unique bias condition, the quadrant of the next target was probabilistically determined and then the location was randomly selected within the quadrant. Critically, the biases manipulated the probability of repeat and alternate stimuli, not the spatial location (e.g., left or right). Experiment 3 was the same as Experiment 2, but stimuli were presented for a fixed duration (0.2 s). This duration was selected because it was below the normal range of response times and was intended to allow the stimulus to be presented for the full duration before participants responded.

Previous work on first-order sequential dependencies has shown that at relatively long response-to-stimulus-intervals (e.g., 2 s), alternations are responded to faster that repeats[9]. This is thought to be due to an irrational belief that when two states are equally probable, the occurrence of one state increases the probability of the subsequent occurrence of the other state, i.e., the gambler's fallacy. Here we deliberately employed a short response-to-stimulus-interval to target automatic sensory processes (i.e., *automatic facilitation*) and avoid these high-level cognitive biases (i.e., *strategic expectancy*)[25].

Eye tracking. Pupil size recordings were epoched to between -0.5 to 1.5 s around target stimulus presentation. Pupillometry data were preprocessed by removing blinks (0.1 s buffer window), removing outliers (median absolute deviation>2.5), interpolated, bandpass filtered with a first-order Butterworth filter (highpass=15.6 Hz, lowpass=500 Hz), *z*-scored, baselined to the 0.5 s period before stimulus presentation, and downsampled to 125 Hz, respectively. Due to hardware issues, pupil data were not

collected from one participant in Experiment 1, two in Experiment 2, and three in Experiment 3.

Behavioural analyses. Trials were grouped according to whether stimuli repeated (the previous stimulus was presented on the same side) or alternated (the previous stimulus was presented on the opposite side) according to three reference planes: task-related, task-unrelated, and stimulus-related. The task-related plane mapped onto the participants' task condition (left-right or above-below) and the task-unrelated plane was orthogonal to the task-related plane. For example, if a participant was assigned the left-right task condition, the task-related plane was defined by the vertical meridian and the task-unrelated plane was defined by the horizontal meridian. The stimulus-related plane was defined by the line bisecting the fixation point at an angle orthogonal to the stimulus location on the previous trial, and thus changed over time.

For micro analyses, the difference in responses to repeat and alternate stimuli was compared. For meso analyses, trials were further sorted into eight pairs of unique sequences, based on the (repeat/alternate) identity of the previous five trials (**Fig. 4a**). For macro analyses in Experiment 2 and 3, repeat and alternate trials were grouped according whether they were expected (75% probability) or unexpected (25% probability), based on participants' experimental condition. For all analyses, sequences that were disrupted by a reproduction task or block advancement were omitted. For each participant, after removal of trials where response time exceeded 2 s, median response times and average accuracy were calculated. Mixture modelling was used to estimate the precision and guess rate of reproduction responses, based on the concentration (κ) and height of von Mises and uniform distributions, respectively [38]. Precision values were positively skewed across participants, so a logarithmic transformation was applied to normalize their distribution ($\log_{10} \kappa$). For serial dependence analyses, the location (μ) of the von Mises distribution was used to estimate response bias.

EEG. EEG data from a previously published dataset were re-analyzed [17]. The stimuli comprised coloured arcs (inner ring, 0.25°; outer ring, 4.25°) extending 45° polar angle on a mid-grey background. Trials consisted of 30 arc stimuli (location and colour randomly selected between 0-360°) presented for 0.2 s each, separated by a blank 0.1 s inter-stimulus-interval. Participants were instructed to indicate the number of

target stimuli that were presented during the preceding trial. At the beginning of each block, target stimuli were identified as either those appearing at a defined location or colour, counterbalanced across the session. EEG data from blocks in which the target was defined by spatial location were included in micro and meso linear discriminant analyses; all data were included in the sequential dependency forward encoding analysis.

The EEG recordings were digitised at 1024 Hz sampling rate with a 24-bit A/D conversion. The 64 active scalp Ag/AgCl electrodes were arranged according to the international standard 10–20 system for electrode placement[39] using a nylon head cap and with an online reference of FCz. Offline EEG pre-processing was performed using EEGLAB v2021.1[40]. The data were initially downsampled to 512 Hz and subjected to a 0.1 Hz highpass filter to remove slow baseline drifts and a 45 Hz lowpass filter to remove high-frequency noise/artifacts. Data were then re-referenced to the common average before being epoched into segments around each stimulus (-0.2 s to 1.0 s from the stimulus onset). To reduce the influence of signals produced by eye movements, blinks, and non-sensory cortices, we only included data from the parietal, parietal-occipital, and occipital sensors.

Neural Decoding. For micro and meso analyses, EEG recordings (epochs) were grouped according to which side of the task-related or unrelated reference plane they were presented. As the task in the EEG experiment was to detect stimuli presented at a specific location, rather than perform a binary judgement (e.g., left or right of fixation), we constructed task boundaries based on the spatial coordinates that were indicative to the target detection task. Thus, the task-related reference plane was defined by the line bisecting the fixation point at an angle orthogonal to the target location, because this plane was most informative of whether a stimulus was a target. By contrast, the task-unrelated plane was the line orthogonal to the task-related line, as this plane was uninformative of whether a stimulus was a target. Ten-fold cross-validation linear discriminant analysis was performed with the MATLAB *classify* function (discriminant function, diaglinear) to calculate the two-way classification accuracy separately at each time point and reference plane. Classification accuracy was then separately assessed for repeat and alternate trials (micro analysis) and the eight pairs of five-event sequences (meso analysis).

As shown in **Supplementary Figure 2**, we found that there were differences in decoding accuracy for repeat and alternate stimuli in the EEG data, even when stimulus labels were shuffled. This is likely due to temporal autocorrelations within the EEG data that are unrelated to the decoded stimulus dimension. This signal causes the decoder to classify temporally proximal stimuli as the same class, leading to a bias towards repeat classification. For example, in general, the EEG signal during trial one will be more similar to that during trial two than during trial ten, because of low frequency trends in the recordings. If the decoder has been trained to classify the signal associated with trial one as a leftward stimulus, then it will be more likely to classify trial two as a leftward stimulus too. These autocorrelations are unrelated to stimulus features; thus, to isolate the influence of stimulus-specific temporal context, we subtracted the accuracy produced by shuffling the stimulus labels from the unshuffled accuracy.

For serial dependency analyses, we used an inverted modelling approach to reconstruct either the location or Δ location (angular distance from previous stimulus) of the stimuli[41]; note, the results of Δ location analyses are shown in **Supplementary Figure 3**. A theoretical (forward) model was nominated that described the measured activity in the EEG sensors given the location/Δ location of the presented stimulus. The forward model was then used to obtain the inverse model that described the transformation from EEG sensor activity to stimulus location/Δ location. The forward and inverse models were obtained using a ten-fold cross-validation approach.

Similar to previous work[22,42,43], the forward model comprised five hypothetical channels, with evenly distributed idealized location/Δ location preferences between 0° and 360°. Each channel consisted of a half-wave rectified sinusoid raised to the fifth power. The channels were arranged such that a tuning curve of any location/Δ location preference could be expressed as a weighted sum of the five channels. The observed EEG activity for each presentation could be described by the following linear model:

$$\mathbf{B} = \mathbf{WC} + \mathbf{E} \qquad (1)$$

where $\mathbf{B}$ indicates the (*m* sensors × *n* presentations) EEG data, $\mathbf{W}$ is a weight matrix (*m* sensors × 5 channels) that describes the transformation from EEG activity to stimulus location/Δ location, $\mathbf{C}$ denotes the hypothesized channel activities (5 channels × *n* presentations), and $\mathbf{E}$ indicates the residual errors.

To compute the inverse model, we estimated the weights that, when applied to the data, would reconstruct the underlying channel activities with the least error. In line with previous magnetencephalography work[44,45], when computing the inverse model, we deviated from the forward model proposed by Brouwer and Heeger (2009)[42] by taking the noise covariance into account to optimize it for EEG data, given the high correlations between neighbouring sensors. We then estimated the weights that, when applied to the data, would reconstruct the underlying channel activities with the least error. Specifically, $\mathbf{B}$ and $\mathbf{C}$ were demeaned such that their average over presentations equalled zero for each sensor and channel, respectively. The inverse model was then estimated using a subset of data selected through cross-fold validation. The hypothetical responses of each of the five channels were calculated from the training data, resulting in the response row vector $\mathbf{c}_{train,i}$ of length n_{train} presentations for each channel i. The weights on the sensors $\mathbf{w}_i$ were then obtained through least squares estimation for each channel:

$$\mathbf{w}_i = \boldsymbol{B}_{train}\mathbf{c}_{train,i}^{\mathrm{T}}\left(\mathbf{c}_{\text{train,i}}\mathbf{c}_{train,i}^{\mathrm{T}}\right)^{-1} \qquad (2)$$

where $\boldsymbol{B}_{train}$ indicates the (m sensors × n_{train} presentations) training EEG data. Subsequently, the optimal spatial filter $\mathbf{v}_i$ to recover the activity of the ith channel was obtained as follows[45]:

$$\mathbf{v}_i = \frac{\widetilde{\sum}_i^{-1}\mathbf{w}_i}{\mathbf{w}_i^{\mathrm{T}}\widetilde{\sum}_i^{-1}\mathbf{w}_i} \qquad (3)$$

where $\widetilde{\sum}_i$ is the regularized covariance matrix for channel i. Incorporating the noise covariance in the filter estimation leads to the suppression of noise that arises from correlations between sensors. The noise covariance was estimated as follows:

$$\widehat{\sum}_i = \frac{1}{n_{\text{train}} - 1}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i^{\mathrm{T}} \qquad (4)$$

$$\boldsymbol{\varepsilon}_i = \mathbf{B}_{train} - \mathbf{w}_i\mathbf{c}_{train,i} \qquad (5)$$

where n_{train} is the number of training presentations. For optimal noise suppression, we improved this estimation by means of regularization by shrinkage using the analytically determined optimal shrinkage parameter[45], yielding the regularized covariance matrix $\widetilde{\sum}_i$.

For each presentation, we decoded location and Δ location by converting the channel responses to polar form:

$$z = \mathbf{c} \cdot e^{2i\boldsymbol{\varphi}} \quad (6)$$

692 and calculating the estimated angle:

$$\hat{\theta} = \frac{\arg(z)}{2} \quad (7)$$

693 where $\mathbf{c}$ is a vector of channel responses and $\boldsymbol{\varphi}$ is the vector of angles at which the
694 channels peak. Decoding accuracy, which represents the similarity of the decoded
695 location/colour to the presented location and Δ location[44], was expressed by projecting
696 the mean resultant (presentations averaged across 12 evenly distributed Δ location
697 bins) of the difference between decoded and arc location onto a vector with 0°:

$$\hat{r}_\theta = \mathrm{Re}[\bar{R}], \quad \bar{R} = \frac{1}{n}\sum_{j=1}^{n} \exp\left(\mathrm{i}\left(\hat{\theta}_j - \theta\right)\right) \quad (8)$$

698 Precision was estimated by calculating the angular deviation[46] of the decoded
699 orientations within each orientation bin:

$$\hat{\sigma}_\theta = \sqrt{2(1 - |\bar{R}|)} \quad (9)$$

700 and normalized, such that values ranged from 0 to 1, where 0 indicates a uniform
701 distribution of decoded orientations across all orientations (i.e., chance-level
702 decoding) and 1 represents perfect consensus among decoded orientations:

$$\hat{p}_\theta = 1 - \frac{2\hat{\sigma}_\theta}{\sqrt{2}} \quad (10)$$

703 Bias was estimated by computing the circular mean of angular difference between the
704 decoded and presented orientation:

$$\hat{b}_\theta = \arg(\bar{R}) \quad (11)$$

705 As shown in **Supplementary Figure 3**, we found the same general temporal
706 dependencies in the decoding accuracy computed using inverted encoding that were
707 found using linear discriminant classification. However, while the pattern of results for
708 decoding *accuracy* and *precision* were similar, decoding precision was not
709 contaminated by these general temporal dependencies and thus we used this metric
710 in the results presented in **Figure 3**.

711 ***Statistical analyses.*** Statistical analyses were performed in MATLAB v2020a and
712 CircStat Toolbox v1.12.0.0[47]. Prior to statistical testing of behavioural estimates,
713 within-participant average values of response time, accuracy, and precision estimates
714 >1.5 times the interquartile range of the group distribution were removed. Non-

715 parametric inferential tests (Wilcoxon signed rank test[48] and Friedman's ANOVA[49])
716 were used to assess the significance of paired differences and main effects. Levene's
717 test of homogeneity was used to compare the difference in variance between
718 conditions[50]. For pupillometry and linear discriminant analyses, one-dimensional
719 cluster correction was applied to remove spurious significant differences [51]. For
720 inverted modelling analyses of location and Δ location as a function of Δ location and
721 time, a two-dimensional circular-linear cluster correction was applied. First, at each Δ
722 location and/or time point, the effect size of the dependent variable (e.g., decoding
723 accuracy) was calculated. Next, we calculated the summed value of these statistics
724 (separately for positive and negative values) within contiguous featural and/or
725 temporal clusters of significant values. We then simulated the null distribution of the
726 maximum summed cluster values using permutation (n=5000) of the sign or condition
727 labels, for single and paired t-test comparisons, respectively, from which we derived
728 the 95% percentile threshold value. Clusters identified in the data with a summed
729 effect-size value less than the threshold were considered spurious and removed. Prior
730 to cluster analysis, we smoothed the decoding accuracy estimates in the inverted
731 modelling analyses along the feature dimension using a uniform kernel (size, 4).

732 ***Data and Code Availability.*** The data and analysis code generated in this study have
733 been deposited in the following OSF database: https://osf.io/9qfdk/

REFERENCES

1. Wolpert, D. M. & Flanagan, J. R. Motor prediction. *Current Biology* **11**, R729–R732 (2001).
2. Barlow, H. B. Possible Principles Underlying the Transformations of Sensory Messages. in *Sensory Communication* (ed. Rosenblith, W. A.) 216–234 (The MIT Press, 1961). doi:10.7551/mitpress/9780262518420.003.0013.
3. Friston, K. A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**, 815–836 (2005).
4. Rao, R. P. N. & Ballard, D. H. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* **2**, 79–87 (1999).
5. Cicchini, G. M., Anobile, G. & Burr, D. C. Compressive mapping of number to space reflects dynamic encoding mechanisms, not static logarithmic transform. *Proceedings of the National Academy of Sciences* **111**, 7867–7872 (2014).
6. Cicchini, G. M., Mikellidou, K. & Burr, D. C. Serial Dependence in Perception. *Annual Review of Psychology* **75**, 129–154 (2024).
7. Fischer, J. & Whitney, D. Serial dependence in visual perception. *Nat Neurosci* **17**, 738–743 (2014).
8. Squires, N. K., Squires, K. C. & Hillyard, S. A. Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology* **38**, 387–401 (1975).
9. Kirby, N. H. Sequential effects in two-choice reaction time: Automatic facilitation or subjective expectancy? *Journal of Experimental Psychology: Human Perception and Performance* **2**, 567–577 (1976).

10. Remington, R. J. Analysis of sequential effects on choice reaction times. *Journal of Experimental Psychology* **82**, 250 (1969).
11. Schapiro, A. & Turk-Browne, N. Statistical Learning. in *Brain Mapping* 501–506 (Elsevier, 2015). doi:10.1016/B978-0-12-397025-1.00276-1.
12. Simoncelli, E. P. & Olshausen, B. A. Natural Image Statistics and Neural Representation. *Annu. Rev. Neurosci.* **24**, 1193–1216 (2001).
13. Kok, P., Jehee, J. F. M. & de Lange, F. P. Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex. *Neuron* **75**, 265–270 (2012).
14. Lange, F. P. de, Heilbron, M. & Kok, P. How Do Expectations Shape Perception? *Trends in Cognitive Sciences* **22**, 764–779 (2018).
15. Richter, D., Heilbron, M. & de Lange, F. P. Dampened sensory representations for expected input across the ventral visual stream. *Oxford Open Neuroscience* **1**, kvac013 (2022).
16. Bertelson, P. Sequential redundancy and speed in a serial two-choice responding task. *Quarterly Journal of Experimental Psychology* **13**, 90–102 (1961).
17. Rideaux, R. Task-related modulation of event-related potentials does not reflect changes to sensory representations. *Imaging Neuroscience* **2**, 1–13 (2024).
18. Cicchini, G. M., Mikellidou, K. & Burr, D. C. The functional role of serial dependence. *Proceedings of the Royal Society B: Biological Sciences* **285**, 20181722 (2018).
19. Sheehan, T. C. & Serences, J. T. Attractive serial dependence overcomes repulsive neuronal adaptation. *PLOS Biology* **20**, e3001711 (2022).

20. Mazancieux, A. *et al.* Brainstem fMRI signaling of surprise across different types of deviant stimuli. *Cell Reports* **42**, 113405 (2023).
21. Preuschoff, K., ’t Hart, B. & Einhauser, W. Pupil Dilation Signals Surprise: Evidence for Noradrenaline’s Role in Decision Making. *Frontiers in Neuroscience* **5**, (2011).
22. Rideaux, R., West, R. K., Rangelov, D. & Mattingley, J. B. Distinct early and late neural mechanisms regulate feature-specific sensory adaptation in the human visual system. *Proceedings of the National Academy of Sciences* **120**, e2216192120 (2023).
23. Rideaux, R. *et al.* Violated predictions enhance the representational fidelity of visual features in perception. 2024.03.27.587109 Preprint at https://doi.org/10.1101/2024.03.27.587109 (2024).
24. Jentzsch, I. & Sommer, W. Functional localization and mechanisms of sequential effects in serial reaction time tasks. *Perception & Psychophysics* **64**, 1169–1188 (2002).
25. Gao, J., Wong-Lin, K., Holmes, P., Simen, P. & Cohen, J. D. Sequential Effects in Two-Choice Reaction Time Tasks: Decomposition and Synthesis of Mechanisms. *Neural Computation* **21**, 2407–2436 (2009).
26. Auksztulewicz, R. & Friston, K. Attentional Enhancement of Auditory Mismatch Responses: a DCM/MEG Study. *Cerebral Cortex* **25**, 4273–4283 (2015).
27. Smout, C. A., Tang, M. F., Garrido, M. I. & Mattingley, J. B. Attention promotes the neural encoding of prediction errors. *PLOS Biology* **17**, e2006812 (2019).
28. Guérit, J.-M., Verougstraete, D., de Tourtchaninoff, M., Debatisse, D. & Witdoeckt, C. ERPs obtained with the auditory oddball paradigm in coma and

altered states of consciousness: clinical relationships, prognostic value, and origin of components. *Clinical Neurophysiology* **110**, 1260–1269 (1999).

29. Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W. & Muckli, L. Stimulus Predictability Reduces Responses in Primary Visual Cortex. *J. Neurosci.* **30**, 2960–2966 (2010).
30. Den Ouden, H., Kok, P. & De Lange, F. How Prediction Errors Shape Perception, Attention, and Motivation. *Frontiers in Psychology* **3**, (2012).
31. Meyer, T. & Olson, C. R. Statistical learning of visual transitions in monkey inferotemporal cortex. *Proceedings of the National Academy of Sciences* **108**, 19401–19406 (2011).
32. Richter, D., Ekman, M. & Lange, F. P. de. Suppressed Sensory Response to Predictable Object Stimuli throughout the Ventral Visual Stream. *J. Neurosci.* **38**, 7452–7461 (2018).
33. Todorovic, A., Ede, F. van, Maris, E. & Lange, F. P. de. Prior Expectation Mediates Neural Adaptation to Repeated Sounds in the Auditory Cortex: An MEG Study. *J. Neurosci.* **31**, 9118–9123 (2011).
34. Tang, M. F., Smout, C. A., Arabzadeh, E. & Mattingley, J. B. Prediction error and repetition suppression have distinct effects on neural representations of visual information. *eLife* **7**, e33123 (2018).
35. Tang, M. F. *et al.* Expectation violations enhance neuronal encoding of sensory information in mouse primary visual cortex. *Nat Commun* **14**, 1196 (2023).
36. Brainard, D. H. The Psychophysics Toolbox. *Spatial Vision* **10**, 433–436 (1997).

37. Pelli, D. G. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision* vol. 10 437–442 Preprint at https://doi.org/10.1163/156856897X00366 (1997).
38. Bays, P. M., Catalao, R. F. G. & Husain, M. The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision* **9**, 7 (2009).
39. Oostenveld, R. & Praamstra, P. The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology* 7 (2001).
40. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods* **134**, 9–21 (2004).
41. Brouwer, G. J. & Heeger, D. J. Cross-orientation suppression in human visual cortex. *Journal of Neurophysiology* **106**, 2108–2119 (2011).
42. Brouwer, G. J. & Heeger, D. J. Decoding and reconstructing color from responses in human visual cortex. *Journal of Neuroscience* **29**, 13992–14003 (2009).
43. Harrison, W. J., Bays, P. M. & Rideaux, R. Neural tuning instantiates prior expectations in the human visual system. *Nat Commun* **14**, 5320 (2023).
44. Kok, P., Mostert, P. & de Lange, F. P. Prior expectations induce prestimulus sensory templates. *Proceedings of the National Academy of Sciences* **114**, 10473–10478 (2017).
45. Mostert, P., Kok, P. & de Lange, F. P. Dissociating sensory from decision processes in human perceptual decision making. *Sci Rep* **5**, 18253 (2015).
46. Zar, J. H. *Biostatistical Analysis*. (Pearson Education India, 1999).

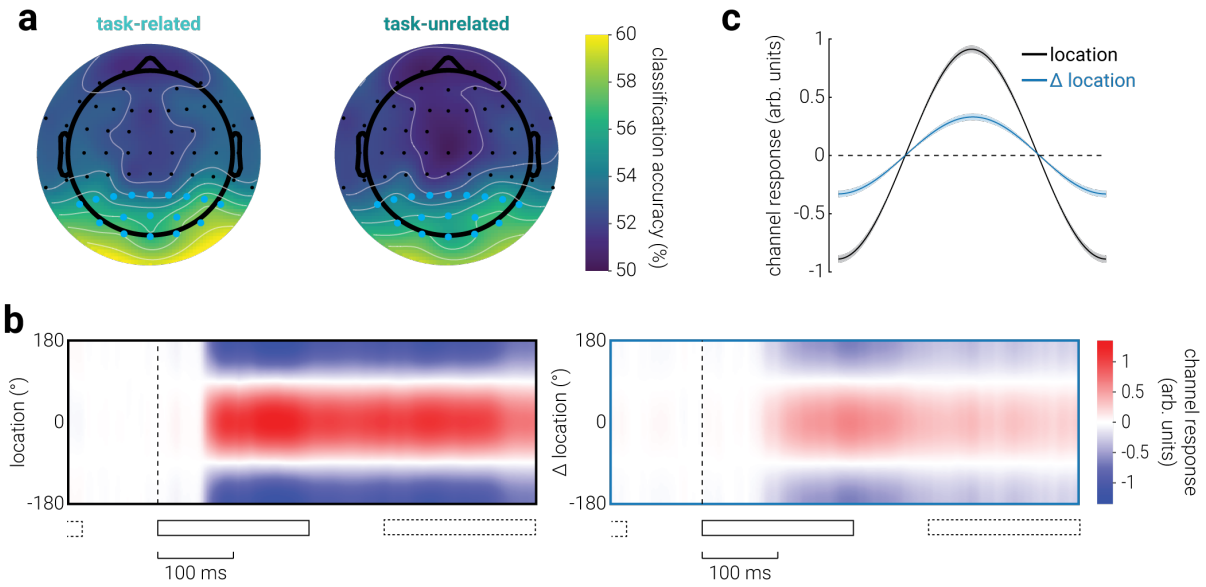
47. Berens, P. CircStat: A MATLAB Toolbox for Circular Statistics. *J. Stat. Soft.* **31**, (2009).
48. Wilcoxon, F. Individual Comparisons by Ranking Methods. in *Breakthroughs in Statistics: Methodology and Distribution* (eds. Kotz, S. & Johnson, N. L.) 196–202 (Springer, New York, NY, 1992). doi:10.1007/978-1-4612-4380-9_16.
49. Friedman, M. The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. *Journal of the American Statistical Association* **32**, 675–701 (1937).
50. Levene, H. Robust tests for equality of variances. (1960).
51. Pernet, C. R., Latinus, M., Nichols, T. E. & Rousselet, G. A. Cluster-based computational methods for mass univariate analyses of event-related brain potentials/fields: A simulation study. *Journal of Neuroscience Methods* **250**, 85–93 (2015).

[insert publisher name here]

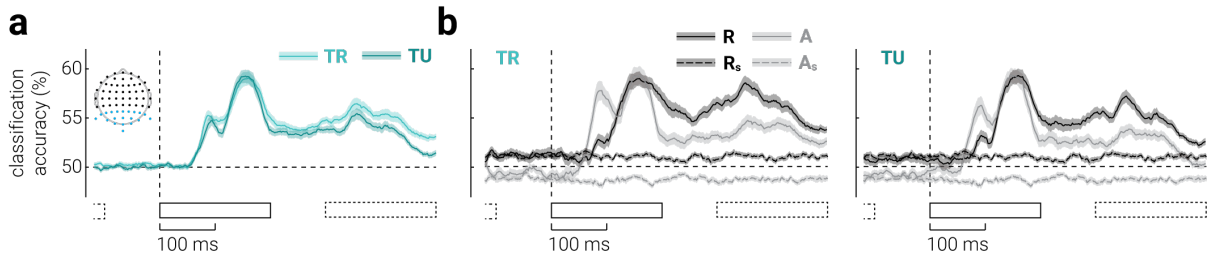
Supplemental Information

The influence of temporal context on visual perception over multiple time scales

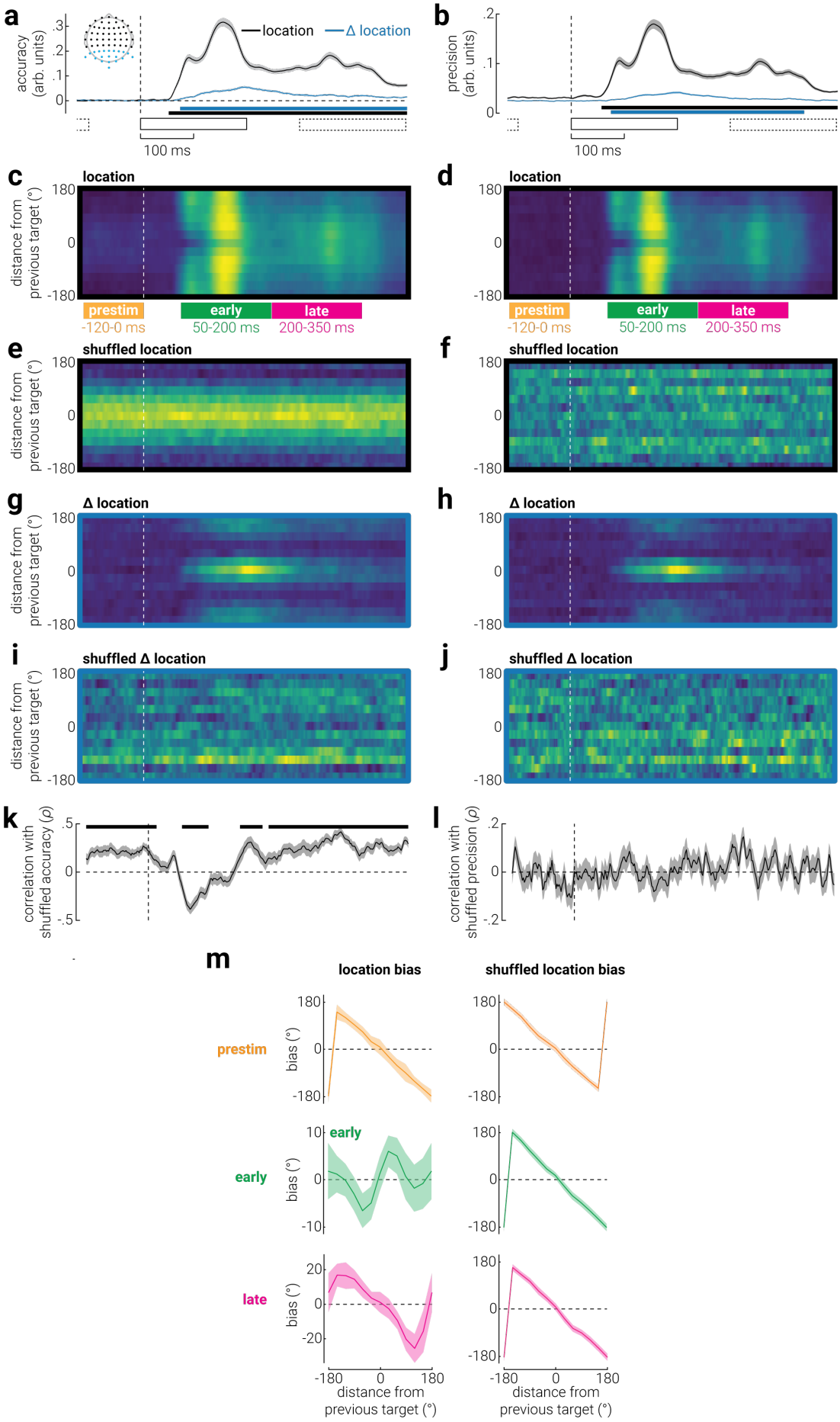
Kacie Lee & Reuben Rideaux



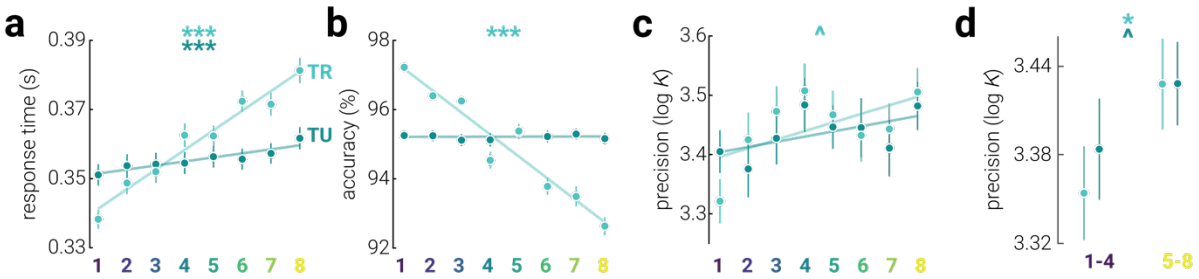
Supplementary Figure 1. Decoding stimulus location and Δ location from EEG recordings. a) Topographic representation of task-related and unrelated stimulus location information produced using linear discriminant analysis to classify stimulus location, separately for each sensor, with time as the multivariate dimension. Blue dots indicate the sensors that were subsequently used in all other EEG analyses. **b**) Channel responses for location and Δ location, produced by computing the dot product between the inverted model channel responses and the forward model, as a function of time. **c**) The average channel response for location and Δ location between 0-500 ms following stimulus onset.



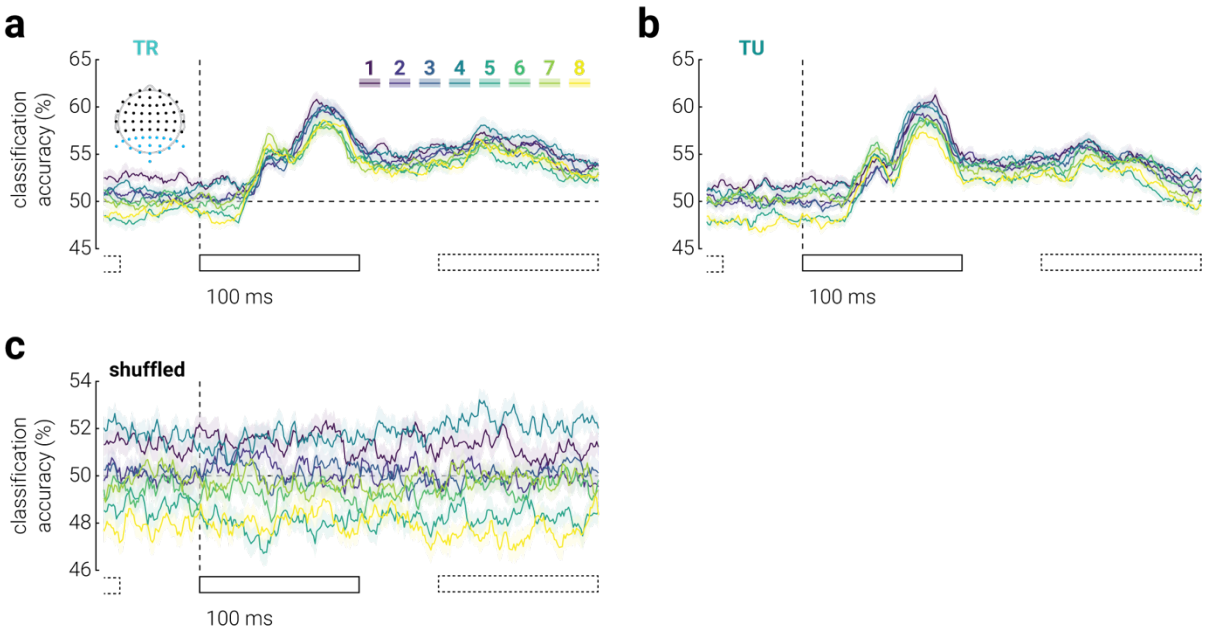
Supplementary Figure 2. Removal of general micro temporal dependencies in EEG responses. We found that there were differences in decoding accuracy for repeat and alternate stimuli in the EEG data, even when stimulus labels were shuffled. This is likely due to temporal autocorrelations within the EEG data that are unrelated to the decoded stimulus dimension. This signal causes the decoder to classify temporally proximal stimuli as the same class, leading to a bias towards repeat classification. For example, in general, the EEG signal during trial one will be more similar to that during trial two than during trial ten, because of low frequency trends in the recordings. If the decoder has been trained to classify the signal associated with trial one as a leftward stimulus, then it will be more likely to classify trial two as a leftward stimulus too. These autocorrelations are unrelated to stimulus features; thus, to isolate the influence of stimulus-specific temporal context, we subtracted the accuracy produced by shuffling the stimulus labels from the unshuffled accuracy (as presented in **Figure 2e, f**). **a**) Shows the uncorrected decoding accuracy along task related and unrelated planes. Note that these results are the same as the corrected version shown in **Figure 2e**, because the confound is only apparent when accuracy is grouped according to temporal context. **b**) Same as (**a**), but split into repeat and alternate stimuli, along (left) task-related and (right) unrelated planes. Decoding accuracy when labels are shuffled is also shown. Inset in (**a**) shows the EEG sensors included in the analysis (blue dots). Black rectangles indicate the timing of stimulus presentations (solid: target stimulus, dashed: previous and subsequent stimuli). Shaded regions indicate ±SEM.


a
accuracy (arb. units)
.3
.2
.1
0
location
Δ location
100 ms
b
precision (arb. units)
.2
.1
0
100 ms
c
location
distance from previous target (°)
180
0
-180
prestim
-120-0 ms
early
50-200 ms
late
200-350 ms
d
location
e
shuffled location
f
shuffled location
g
Δ location
h
Δ location
i
shuffled Δ location
j
shuffled Δ location
k
correlation with shuffled accuracy (ρ)
.5
0
-.5
l
correlation with shuffled precision (ρ)
.2
0
-.2
m
location bias
shuffled location bias
prestim
early
late
bias (°)
10
-10
20
-20
distance from previous target (°)

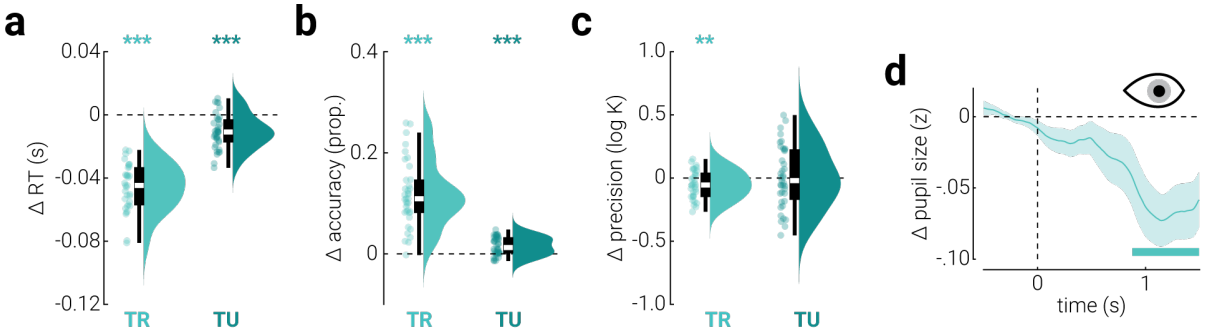
Supplementary Figure 3. Decoding accuracy, but not precision, is confounded by general serial dependencies. As described in *Methods - Neural Decoding,* we used inverted encoding modelling of EEG recordings to estimate the decoding accuracy and precision of stimulus location and Δ location. As shown in (**a**, **b**), there was a high correspondence between the relative magnitude of accuracy and precision, for both location and Δ location, over time. We also found that these parameters were very similar when further examined as a function of the distance of between current and previous stimulus location (**c**, **d**, **g**, **h**). To assess the influence of the general temporal dependencies found in the previous linear classification analysis between repeat and alternate stimuli, we calculated accuracy and precision on after shuffling the labels of the dataset. For decoding accuracy, we found a clear bias towards similar stimuli, i.e., those with small inter-trial offsets, for stimulus location (**e**), and less clear evidence for Δ location (**i**). Indeed, correlation analysis between the original and shuffled data decoding accuracy confirmed a significant positive relationship between the two for period before and after stimulus presentation (**k**). By contrast, for decoding precision, we found no evidence of these general temporal dependencies for either location or Δ location (**f**, **j**, **l**). Based on these findings, we assessed stimulus-specific serial dependence in the EEG recordings using decoding precision, which is uncontaminated by general serial dependence. However, this general dependence may explain the findings of a previous study. In particular, [19] found a repulsive pattern of biases associated with serial dependency in fMRI BOLD responses. The authors suggested that this may be due to adaptation, and provided a model in which this bias is reversed in higher cortical areas to produce the attractive bias associated with serial dependence. However, we find the same repulsive pattern of biases in the shuffled data (**m**, right), and in the pre-stimulus and late post-stimulus periods of the original data (**m**, top-left and bottom-left). By contrast, we found an attractive bias in the early post-stimulus period of the original data, consistent with serial dependence (**m**, middle-left). Thus, a more parsimonious explanation for the repulsive bias found by [19] is that they were observing the bias introduced by general temporal dependencies. Pre-stimulus (-120-0 ms), early (50-200 ms), and late (200-350 ms) periods are illustrated in orange, green, and pink, respectively. Note, the temporal abscissa is aligned across (**a**, **c, e**, **g, l, k** & **b, d, f, h, j, l**). Shaded regions indicate ±SEM. In our inverted encoding analysis, the precision of the early neural representation of stimulus location matched what we would expect from the previous binary analyses of micro temporal context, that is, highest precision when there was a large Δ location (i.e., angular distance between the target stimulus and the previous stimulus). However, behavioural precision showed an additional peak around small Δ location. This may be explained by the integration of location and Δ location representations, as Δ location, which was also peaked around small and large offsets, as shown in (**g**).



Supplementary Figure 4. Meso-scale temporal context effects in Experiment 2. a) Response time, (**b**) accuracy, and (**c**) precision as a function of sequence, across task-related and unrelated reference planes. Note, lower numbers on the abscissa are associated with more expected sequences. **d**) Same as (**c**), but pooled across the four most expected (1-4) and unexpected sequences (1-8). Asterisks indicate significant correlations with sequence order (*p<.05, **p<.01, ***p<.001), hats indicate marginal significance (^p<.06). Error bars indicate ±SEM.



Supplementary Figure 5. Removal of general meso temporal dependencies in EEG responses. To remove any potential general sequential dependences from the meso-scale classification analysis of EEG recordings, we performed the same subtraction method as used in the micro-scale analysis. (**a**) Classification accuracy of stimuli presented on different sides of task-related planes as a function of time, for each of the eight sequences, from re-analysis of previously published EEG data [17]. **b**) Same as (**a**), but for the task-unrelated plane. (**c**) Same as (**a**), but with shuffled labels. **Figure 4f, g** shows (**a**, **b**) after removing shuffled accuracy (**c**). Inset in (**a**) shows the EEG sensors included in the analysis (blue dots). Black rectangles indicate the timing of stimulus presentations (solid: target stimulus, dashed: previous and subsequent stimuli). Shaded regions indicate ±SEM.



Supplementary Figure 6. Macro-scale temporal context effects in Experiment 3. a-c) The difference (expected – unexpected) in (**a**) response time, (**b**) accuracy, and (**c**) reproduction precision between expected and unexpected stimuli, along task-related (TR) and unrelated (TU) planes. Asterisks indicate significant differences (**p<.01, ***p<.001). **d**) The difference (expected – unexpected) in pupil size as a function of time from stimulus presentation, along the task-related (TR) plane. Shaded region indicates ±SEM and the horizontal bar indicates a cluster corrected period of significant difference.