City Research Online

City, University of London Institutional Repository

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/23319/

Link to published version: http://dx.doi.org/10.3758/s13414-019-01899-x

City Research Online: http://openaccess.city.ac.uk/   publications@city.ac.uk

Predictable events elicit less visual and temporal information uptake in an oddball paradigm.

Blake W. Saurels[1], Ottmar V. Lipp[1,2], Kielan Yarrow[3] & Derek H. Arnold[1]

1. *School of Psychology, The University of Queensland, Australia.*

2. *School of Psychology, Curtin University, Australia.*

3. *Department of Psychology, City, University of London, United Kingdom.*

Running Title: Predictable events elicit less information uptake.

Submission Date: January 2019

Corresponding Author: Blake W. Saurels

e-mail: blake.saurels@uq.net.au

Abstract

In the visual oddball paradigm, surprising inputs can seem expanded in time relative to unsurprising repeated events. A horizontal input embedded in a train of successive vertical inputs can, for instance, seem relatively protracted in time, even if all inputs are presented for an identical duration. It is unclear if this effect results from surprising events becoming apparently protracted, or from repeated events becoming apparently contracted in time. To disambiguate, we used a non-relative duration reproduction task, in which several standards preceded a test stimulus that had to be reproduced. We manipulated the predictability of test content over successive presentations. Overall, our data suggest that predictable stimuli induce a contraction of apparent duration (Experiments 1, 3 & 4). We also examine sensitivity to test content, and find that predictable stimuli elicit less uptake of visual information (Experiments 2 & 3). We discuss these findings in relation to the predictive coding framework.

Our perception of the passage of time is malleable (Augustine, 1960; Wittmann, 2013). Several means of altering apparent duration have been discovered and used experimentally (Block & Grondin, 2014; Grondin, 2010; Matthews & Meck, 2016). One is the oddball paradigm, where a surprising 'odd' event seems longer than a train of predictable, repeated events (the *temporal oddball effect*; Pariyadath & Eagleman, 2007; Schindel, Rowlands, & Arnold, 2011; Tse, Intriligator, Rivest, & Cavanagh, 2004; Ulrich, Nitschke, & Rammsayer, 2006). Disagreement concerning how to interpret oddball results has led to two opposing accounts: the duration *dilation* account (Tse et al., 2004) and the duration *contraction* account (Pariyadath & Eagleman, 2007).

The duration dilation account put forward by Tse et al. (2004) posits that the temporal oddball effect (TOE) results from an *increase* in subjective duration for surprising events. These investigators found that the TOE only occurred for stimuli presented for longer than ~120ms (also see Ulrich et al., 2006). This led them to suggest the effect works via attentional modulation, citing the similarity in timing between this and the time required for attention to be orientated to new stimuli (~120-150ms; Hikosaka, Miyauchi, & Shimojo, 1993; Nakayama & Mackeben, 1989). They linked this idea to Treisman's (1963) model of interval timing, suggesting that: *"If attending to a stimulus boosts information processing of that stimulus, the counter would count more units, and subjective time would expand."* (Tse et al., 2004, p. 1172). This could link the TOE to the effect arousal has been shown to have on apparent timing (e.g. Mella, Conty, & Pouthas, 2011).

Pariyadath and Eagleman (2007) suggested instead that the TOE is driven by a *contraction* of perceived duration for repeated stimuli (i.e., the oddball is judged as longer because the perceived duration of repeated inputs is shortened). Their investigation centred on the debut

effect, a related finding wherein the first of a train of repeated events can seem longer than all subsequent repeats (Kanai & Watanabe, 2006). Others have demonstrated this repetition effect with single stimulus repetitions (Birngruber, Schröter, & Ulrich, 2015; Matthews, 2011). Pariyadath and Eagleman (2007) suggested that both the TOE and the debut effects result from a contraction of perceived duration for repeated events. They further suggested that the magnitude of the brains' response to inputs was an index of perceived duration, and that this was reduced for repeated stimuli via low-level sensory adaptation.

This last suggestion was refuted by Schindel et al. (2011), who found that switching the eye to which tests were presented eliminated Troxler fading (a phenomenon known to result from low-level monocular adaptation), but this had no discernible impact on the TOE. Schindel et al. (2011) also showed that a protracted continuous stimulus presentation resulted in Troxler fading for a short test subsequently presented to the same eye, but had no impact on the apparent duration of these tests. In combination, these data speak against a link between the TOE and visual adaptation at early monocular stages of visual processing (see Clarke & Belcher, 1962). These results, however, leave open the possibility that higher-level adaptation or expectation effects might impact the magnitude of the brains' response to inputs, and that this might impact perceived duration.

Opposing accounts of the TOE remain tenable, as a fundamental question has not been adequately resolved: do oddballs seem longer, or do repeated events seem shorter in duration? Both accounts remain viable because most oddball experiments have used a relative judgement of perceived duration (e.g., 'was the duration of the oddball longer or shorter relative to other/previous stimuli?'). One of the aims of this study is to differentiate these two accounts, by moving away from using relative measures of duration. To accomplish this, we adopted a

time reproduction task (see Mioni, Stablum, McClintock, & Grondin, 2014). This allowed us to measure reproduced durations for repeated (a final input, identical to a preceding train of repeating images) and oddballs (a final input, differing in orientation to a preceding train repeating images, see Figure 1) tests, comparing these to a 'random' baseline condition (see Figure 1).
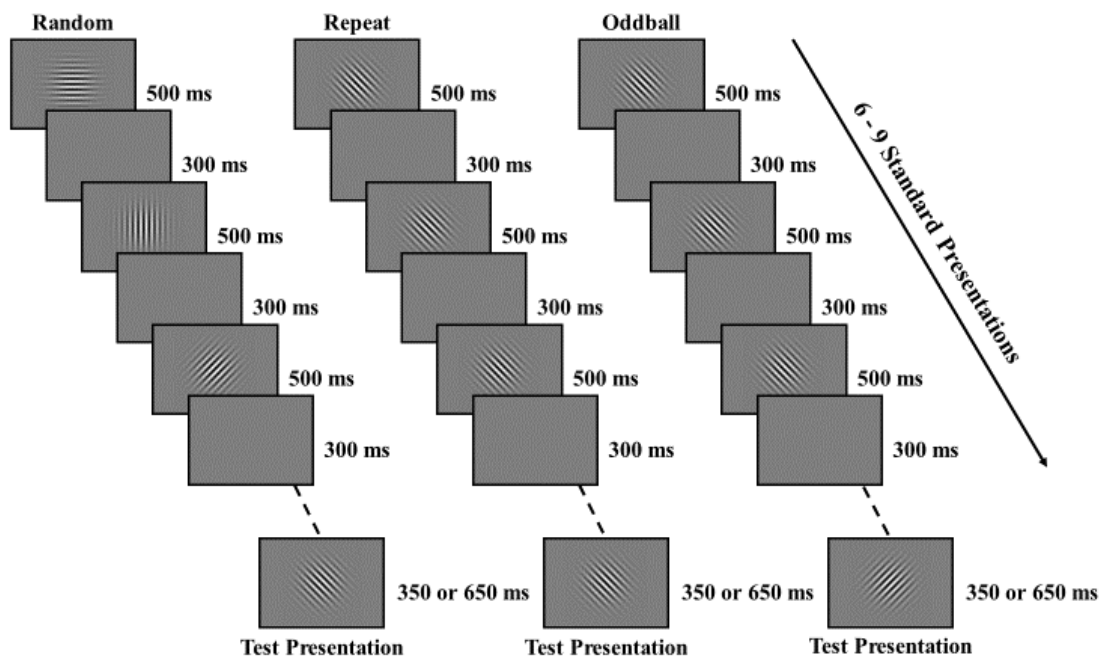

Figure 1. Stimulus presentation protocols for Experiment 1. Three types of trial are depicted: 1) Random, wherein the orientation of each Gabor differed from the last, 2) Repeat, wherein the orientation of each Gabor was identical, and 3) Oddball, wherein all standard Gabors were identical, but the test Gabor orientation differed. Standard presentations were 500ms long, ISIs were 300ms, and tests were 350 or 650ms.

The random baseline condition is designed to minimise orientation-tuned adaptation from successive presentations of the same input (which could contribute to the TOE – see Pariyadath and Eagleman, 2007), while maintaining 'surprise', by having a train of events that each differ in orientation from the last (the key characteristic of an oddball presentation; later in the project we came to question this rationale, see Experiments 3 & 4 and General Discussion). We hypothesised that if the TOE is produced by a duration contraction for repeated stimuli, repeat tests should engender *shorter* reproductions than random tests. Alternatively, if the TOE is

produced by a duration dilation for surprising stimuli, we hypothesised that oddballs should engender *longer* reproductions than random tests. If both accounts are correct, both differences should manifest. We tested these predictions in Experiment 1.

The second aim of this project is to measure visual sensitivity to test content. The duration of physical exposure can limit sensitivity to visual input (Barlow, 1958). This motivated us to see if the presentation protocols that elicit the TOE would also modulate sensitivity. To assess this, in Experiment 2 we modify the oddball paradigm. Instead of participants making duration reproductions, participants are asked to discern which side of a split-half test Gabor has a higher spatial frequency, permitting an objective measure of sensitivity to visual information. If the TOE is driven by a more rapid uptake of visual information (as suggested by Tse et al., 2004), we reasoned that people should be more sensitive to content embedded in oddball tests, relative to random and repeated tests. We test this hypothesis in Experiment 2.

Experiment 1

Methods

Experimental procedures, participant numbers, exclusion criteria and analyses for this experiment were pre-registered (https://aspredicted.org/74je2.pdf).

Participants. Twelve participants (4 males; pre-registered target *N*) were recruited for testing via a research participation scheme at the University of Queensland. All reported having normal or corrected-to-normal visual acuity, and there were no exclusions based on pre-registered criteria (see results section). This *N* was considered appropriate in light of previous investigations of subjective timing (e.g. Arnold, Nancarrow, & Yarrow, 2012, *N* = 10; Keane, Spence, Yarrow, & Arnold, 2015, *N* = 8; Marinovic & Arnold, 2012, *N* = 8), which suggested

a very strong effect size (Cohen's d = ~0.9; used throughout). Thus, we estimated an N of 12 would provide sufficient power (~0.81) to detect the TOE. Ages ranged from 17 to 37 (M = 23, SD = 6). This experiment was approved by the University of Queensland ethics committee and was conducted in accordance with the principles of the Declaration of Helsinki.

Stimuli and apparatus. Stimuli consisted of Gabor patches, with a diameter subtending 10 degrees of visual angle (dva), a Michelson luminance contrast of 100%, and a spatial constant of 1.7 dva, presented on a grey background. Individual Gabors were oriented +/- 0, 45, 90, or 135º from vertical. Stimuli were presented on a 20" CRT SyncMaster 1100p-Plus monitor, driven by a Cambridge Research Systems ViSaGe stimulus generator and custom Matlab R2007b (The MathWorks, Natick, MA) software. The monitor had a resolution of 1024 x 768 pixels and a refresh rate of 100 Hz. Participants viewed stimuli from 57cm, from directly in front of the monitor with their chin placed on a chin rest.

Procedure. Each trial involved six to nine 500ms standard Gabor presentations (the precise number was determined at random on a trial-by-trial basis) separated by 300ms inter-stimulus intervals (ISIs). Following standards, a single test Gabor was presented for either 350ms (short) or 650ms (long; see Figure 1). After each test presentation, participants were asked to reproduce the duration of the test, by pressing and holding down the left mouse button for a matched duration.

There were three types of trial: random, repeat, and oddball. On random trials, standard Gabor orientations were pseudo-randomly set to one of four candidate orientations (0, 45, 90, or 135 º), with the proviso that the orientation of each Gabor differed from the preceding Gabor by ≥45º. On these trials, test Gabors were oriented +/-90º relative to the last standard Gabor. On

repeat trials, all standard and test Gabors had the same orientation, selected randomly from the four candidate orientations on a trial-by-trial basis. On oddball trials, standard Gabors had the same orientation, selected randomly from the four candidate orientations on a trial-by-trial basis, and test Gabors were rotated +/-90° relative to the repeats. Each type of trial was presented 40 times for each test duration, for a total of 240 individual trials, all completed in a random order. The dependent variable was the median reproduced test duration, calculated for each of the six experimental conditions.

Results

Exclusion criteria. It was pre-registered that participants who failed to reproduce longer durations for long tests (650ms), relative to short tests (350ms), regardless of test condition, would be excluded from further analysis. *T*-tests were conducted for individual participant data, and no participants were excluded.

Reproduced durations. The basis for our analyses were the median durations reproduced by each individual for each test condition (see Table 1 for mean of medians).

Readers may note from Table 1 that participants tended to reproduce physically *shorter* intervals (~500ms) for long presentations (650ms). This may speak to motor planning times being added in perception to button presses – similar to the stopped clock illusion after a saccade (Yarrow & Rothwell, 2003). Additionally, there is some indication of a Vierordt (1868) effect, with longer reproductions than physical tests for short presentations, and shorter reproductions than physical tests for long presentations. These differences relative to physical times are not, however, the focus of our project. Instead, we are concerned with distortions of

perceived duration resulting from repetition or surprise, which are addressed by examining

differences in performance across experimental conditions.

Table 1   Mean of median durations (ms) reproduced by each participant, for each test condition for each test duration, along with standard deviations between individual medians.

Condition	Short (350ms)		Long (650ms)	
	M	*SD*	*M*	*SD*
Repeat	372	120	497	128
Random	374	110	502	109
Oddball	365	107	544	124

Difference scores. As per our pre-registered analysis plan, individual reproduced duration

difference scores were calculated for repeat and oddball conditions, relative to the random

baseline condition. These were calculated by taking differences between median reproduced

test durations for the pertinent test conditions. Positive difference scores indicate that

participants reproduced *longer* durations relative to random baseline, whereas negative scores

indicate that participants reproduced *shorter* durations.

Difference scores for the repeat condition (M = -4ms, SD = 60ms) indicated that reproduced

durations for repeat and random baseline tests were not different (t_{11} = 0.22, p = .832).

Difference scores for the oddball condition (M = 17ms, SD = 26ms) indicated that while

participants reproduced longer durations in response to oddballs, relative to the random

baseline tests, this difference was not statistically significant (t_{11} = 2.19, p = .051). Nor were

the sets of difference scores reliably different from one another (t_{11} = 1.57, p = .145).

Statistical analyses described up to this point were pre-registered. Hereafter we describe

additional follow-up analyses.

The results of pre-registered analyses were surprising in light of existing literature, which had suggested a robust TOE. We suspected this apparent discrepancy was due to motor planning noise contaminating duration estimates. Any such noise would disproportionately impact the start and finish of reproduced time intervals, and this should have been particularly apparent for short (350ms) relative to longer (650ms) tests (due to the shortened interval in-between endpoints). This disproportionate impact, of motor planning noise on short interval reproductions, was supported by a post-hoc analysis of SDs between individual time reproductions. If these were due to interval timing error, they should have increased in proportion to Weber's law (Treisman, 1963) – by a factor of ~1.86 for our long (650ms) relative to short (350ms) tests. Instead, the increase was markedly less (1.13; 350ms condition SDs: M = 114ms, SD = 48ms; 650ms condition SDs: M = 129ms, SD = 53ms; t_{11} = 2.60, p = .025), suggesting time reproductions were subject to motor planning errors that were disproportionate for short tests. We thus repeated the statistical analyses described above but excluded data for short tests.

For analyses restricted to long trials, difference scores for the repeat condition (M = -5ms, SD = 67ms) indicated that reproduced durations for repeat and random baseline tests were still similar (t_{11} = 0.26, p = .800; see Figure 2). Difference scores for the oddball condition (M = 42ms, SD = 30ms), however, indicated that participants had reproduced *longer* durations for oddballs, relative to random baseline tests (t_{11} = 4.90, p < .001). There was also a difference between these two sets of difference scores (t_{11} = 2.38, p = .037).
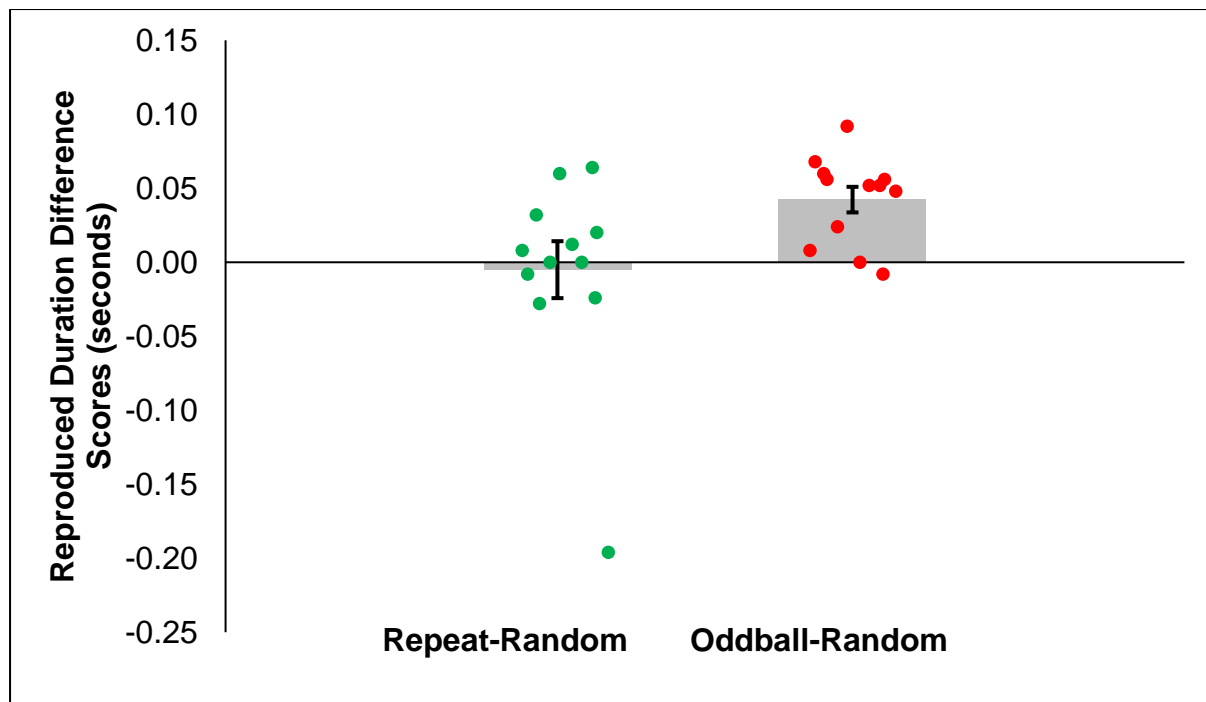
Figure 2. Reproduced duration difference scores for the long (650ms) trials. Positive difference scores indicate longer reproduced durations relative to random baseline tests, whereas negative difference scores indicate shorter reproduced durations. Grey bars depict average difference scores. Black error bars depict +/- 1 standard error. Coloured circles depict individual difference scores.

Readers may note an outlier repeat-random difference score in Figure 2. Exclusion of this does not impact the pattern of results; repeat and random baseline tests are still similar ($t_{10} = 1.36$, $p = .203$), longer times are still reproduced for oddball relative to random tests ($t_{10} = 4.42$, $p = .001$), and there is still a difference between these two sets of difference scores ($t_{10} = 3.06$, $p = .012$).

Bayes' factors were estimated using JASP software (2015), with a default Cauchy prior width of 0.707, to assess the weight of evidence for the null (that reproduced durations would be equivalent for tests and baseline trials) and alternate (that reproduced durations would differ for tests and baseline trials) hypotheses. The results of these analyses, restricted to long trials, were consistent with frequentist results, with very strong evidence for the alternate hypothesis for oddballs ($BF_{10} = 75.439$). These analyses also revealed moderate evidence for the null

hypothesis for repeat tests ($BF_{10} = 0.296$), and anecdotal evidence for the alternate hypothesis in relation to differences between the individual sets of difference scores ($BF_{10} = 2.100$).

Experiment 2

Details were as for Experiment 1, with the following exceptions.

Methods

Participants. We pre-determined target N of 30 participants. This number is greater than Experiment 1, as we anticipated a weaker effect size (~0.55) based on previous studies that had used similarly large sample sizes to examine spatial frequency discrimination (see: Astle, Webb, & McGraw, 2010; Patel, Maurer, & Lewis, 2010; van der Boomen & Peters, 2017). We estimated that an N of 30 would provide sufficient power (0.83). In total, 34 participants were recruited for testing via a research participation scheme at the University of Queensland. None had participated in Experiment 1. All reported having normal or corrected-to-normal visual acuity. Four were removed from analysis based on a pre-determined exclusion criteria (see results section), leaving a total of 30 participants (the target N; 6 males). Ages for the final sampled ranged from 17 to 39 ($M = 20.90$, $SD = 5.42$).

Stimuli and apparatus. Gabor patches were split in half, via a 1 dva wide vertical grey bar (matched to the display background; see Figure 3). Standard Gabors had the same spatial frequency (1.5 cycles/degree of visual angle; herein, c/dva) on either side, whereas test Gabors had different spatial frequencies, adjusted +/- relative to the standard value, on either side. The side of the test with the higher spatial frequency was determined at random, on a trial-by-trial basis. Due to this minor spatial frequency difference, the repeat test image wasn't identical to

preceding images. However, with the exception of the first few trials of each staircase, differences were very slight, with sampling converging on a threshold noticeable difference.
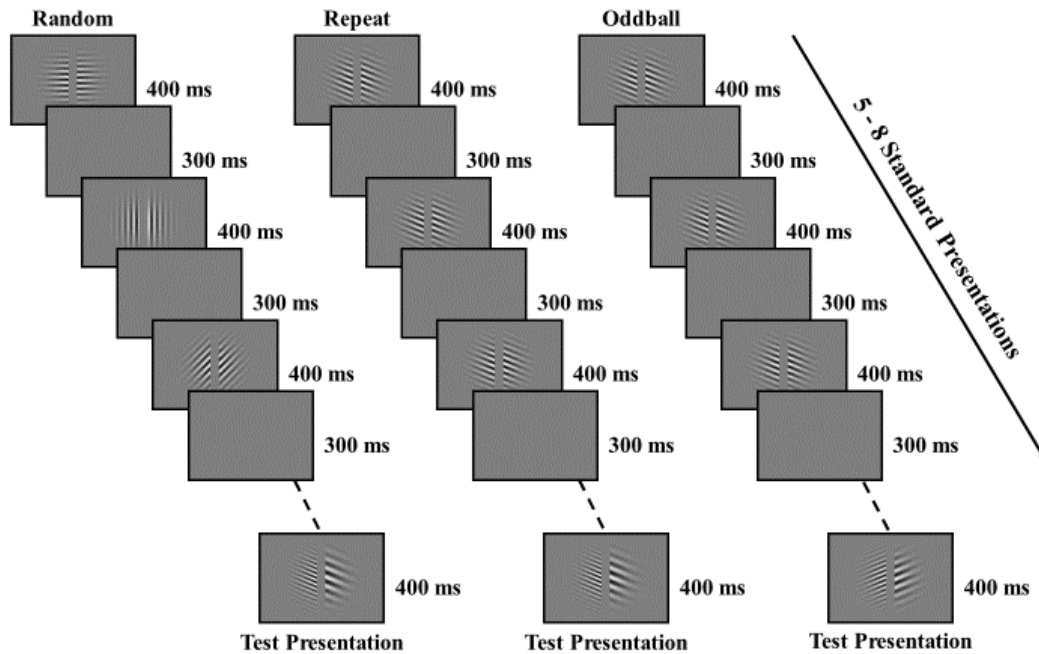

Figure 3. Stimulus presentation protocols for Experiment 2. The three types of trial depicted follow the same sequence as Experiment 1.

Procedure. Each trial involved five to eight 400ms standard Gabor presentations, followed by a 400ms test presentation. We used a smaller average number of presentations per trial, and shorter standard test presentations (both relative to Experiment 1), in order to obtain more experimental observations per session. We had confirmed anecdotally, in pilot testing on authors and lab members, that these parameters seemed to elicit a robust TOE. After test Gabor presentations, participants indicated which side of the test Gabor they perceived to have the higher spatial frequency, by pressing the left or right mouse button. The difference between split-half test spatial frequencies was manipulated according to 1-up 2-down staircase procedures. During these, an incorrect judgment results in an increased difference, and two successive correct judgments in a decreased difference (0.2 c/dva increments initially, then 0.02 c/dva after two incorrect judgements). Such staircase procedures converge on ~71% correct task performance. Each staircase was instigated at a difference of 1 c/dva. Feedback

was provided after each response by way of an ascending tone for correct responses and a descending tone for incorrect responses.

During each block of trials, a single staircase procedure was conducted for each trial type. Each type of trial was presented 60 times, for a total of 180 individual trials, all completed in a random order. Each block of trials provided three distributions of correct task performance as a function of the magnitude of split-half spatial frequency differences (one for each trial type). Logistic functions, based on maximum likelihood analyses, and spanning 50 to 100%, were fit to distributions of correct task performance, and physical spatial frequency differences coinciding with 75% points of fitted functions were taken as estimates of the Just Noticeable Difference threshold (JND) for split-half test spatial frequency differences.

Results

Exclusion criteria. Participants whose estimated JND for a specific trial type were > 3 SDs from the mean were excluded from analysis. Four participants were excluded on this basis.

Sensitivity thresholds. The basis for our analyses were the JND threshold scores for each individual for each test condition (see Table 2 for mean of JND thresholds).

Table 2	Mean spatial frequency JND thresholds scores for each test condition	
Condition	All (400ms)	
	M	*SD*
Repeat	0.086	0.066
Random	0.075	0.044
Oddball	0.053	0.035

Difference scores. As per our analysis plan, individual difference scores were calculated by subtracting JND threshold estimates for repeat and oddball tests, from JND estimates for random tests. Positive difference scores indicate that participants were *less sensitive* to spatial frequency differences relative to random baseline tests, whereas negative difference scores indicate greater sensitivity.

Difference scores for the repeat condition (M = 0.01 c/dva, SD = 0.07) indicated that participants were equally sensitive to spatial frequency differences in repeat and random baseline tests (t_{29} = 0.86, p = .396; see Figure 4). By contrast, difference scores involving the oddball condition (M = -0.02, SD = 0.05) indicated that participants were *more* sensitive to spatial frequency differences within oddballs, relative to random baseline tests (t_{29} = 2.58, p = .015). These two sets of difference scores could also be statistically distinguished (t_{29} = 2.68, p = .011).

Figure 4. Spatial frequency JND threshold difference scores. Positive difference scores indicate decreased sensitivity relative to random baseline tests, whereas negative difference scores indicate

increased sensitivity. Grey bars depict average difference scores. Black error bars depict +/- 1 standard error. Coloured circles depict individual difference scores.

Consistent with frequentist test results, moderate evidence for the alternative, relative to the null, hypothesis was found for difference scores involving oddball tests (BF_{10} = 3.167). Moderate evidence for the null, relative to the alternative, hypothesis was also found for differences scores involving repeat tests (BF_{10} = 0.273). Moderate evidence for the alternative was also found in relation to the difference between these two sets of difference scores (BF_{10} = 3.855).

Discussion

Experiments 1 and 2 provide a consistent pattern of results. Repeat identical tests, and tests following a train of pseudo-randomly oriented inputs, are matched in terms of perceived duration and sensitivity to content. Oddball tests, by contrast, elicit longer time reproductions and heightened sensitivity relative to random tests. Interpretation of these results depends on the status of random tests.

The random condition was intended as a baseline, against which repeat and oddball tests could be referenced. The different orientations of successive inputs should reduce low-level orientation adaptation, preventing this from contributing to the TOE or sensitivity changes. Consistent with a previous result (Schindel et al., 2011), we found that preventing low-level orientation adaptation failed to equate the perceived durations of (or sensitivity to content within) oddball and random tests. Does this mean reduced responding to successive inputs plays no role in generating the TOE? Is the TOE generated entirely by an *exaggerated* response to unexpected inputs?

The answers to these questions hinge on how the brain responds to a sequence of differently oriented inputs. It is possible these are treated as members of a single class of *predictable* input, as each event within these 'random' sequences has an identical spatial frequency, contrast, matching predictable timings, and they each have a reliable difference in orientation from the last input. The brain might therefore become less responsive to such inputs, including the final test, as these are highly *predictable* given the context. We felt this dictated that reduced responding to inputs based on predictability was still a viable interpretation of our results.

To explore these issues further, in Experiment 3 we avoid using a repetitive sequence of events as a baseline. Instead, we compare all three successive presentation protocols against a common baseline – a single isolated test Gabor. We also have participants provide two responses per trial, allowing us to explore how modulations of predictability in repetitive sequences shape both perceived duration and sensitivity in a common set of trials.

Experiment 3

Methods

Details were as for Experiment 2, with the following exceptions (https://aspredicted.org/ix878.pdf).

Participants. We *pre-registered* a target *N* of 20 participants. This number is similar to or greater than samples in comparable studies (see: Cai, Eagleman, & Ma, 2015; Schindel et al., 2011), but represents an increase relative to Experiment 1. Guided by the results of Experiment 1, we had revised our expectations of the temporal oddball effect size (to ~0.7), and so estimated that a larger *N* of 20 would be necessary to promote sufficient power (0.84). A total of 26 participants were recruited for testing via a research participation scheme at the

University of Queensland. None had participated in either of the earlier experiments. All reported having normal or corrected-to-normal visual acuity. Six were removed from analysis according to pre-registered exclusion criteria (see results section), leaving a total of 20 participants (the pre-registered target *N*; 6 males). Ages for the final sample ranged from 18 to 29 (*M* = 20.45, *SD* = 2.91).

Procedure. There were two experimental blocks, for which the order of completion was counterbalanced across participants. In the test block, each trial involved six to ten 650ms standard Gabor presentations, separated by 300ms ISIs. These were followed by a test Gabor presentation of either 500 (short) or 800ms (long). In this block, each type of trial (random, repeat, & oddball) was presented 30 times for each test duration, for a total of 180 individual trials, all completed in a random order. In the baseline block, a single test Gabor was presented on each trial (with no preceding standards). In this block, each test duration was presented 45 times, for a total of 90 individual trials, all completed in a random order.

All orientations were possible in this experiment, whereas in previous experiments orientations were restricted to four candidates. Initial orientations were selected at random on a trial-by-trial basis. Repeat and oddball trials used the same rules as previous experiments. On random trials each successive Gabor was rotated between 30 to 60° (clockwise or counter-clockwise) relative to the previous Gabor, with rotation direction and magnitude randomised on a presentation-by-presentation basis.

After test Gabor presentations, participants first performed a duration reproduction task (as per Experiment 1), followed by a spatial frequency judgement (as per Experiment 2). For duration reproductions, a white horizontal bar would appear, and grow from left to right against a black

background while the participant held down the left mouse button (to reproduce the test duration). The reproduced duration (in seconds) was visible in text below the finished bar. After each time reproduction, participants had the option to accept their estimate, by pressing the middle mouse button, or to trigger a repeat of the time reproduction sequence (by pressing the right mouse button). We introduced this feedback to give participants a more accurate sense of their' duration reproductions, and to allow them to re-record any that they felt were erroneous. For sensitivity judgements, all details were the same as for Experiment 2, except that the choice options (left or right side) were displayed on screen following the duration judgement as a prompt for the participant.

Results

Exclusion criteria. Participants whose average reproduced durations for long (800ms) and short (500ms) trials (regardless of trial) were not significantly different ($p < .05$) were excluded from further analysis. Six participants were excluded on this basis.

Reproduced durations. The basis for our analyses were the median durations reproduced by each individual for each test condition (see Table 3 for mean of medians).

Table 3   Mean of median durations (ms) reproduced by each participant, for each test condition for each test duration, along with standard deviations between individual medians.

Condition	Short (500ms)		Long (800ms)	
	M	*SD*	*M*	*SD*
Baseline	386	115	492	149
Repeat	374	123	465	154
Random	379	117	482	155
Oddball	379	120	507	166

Difference scores. As per our pre-registered analysis plan, individual duration difference scores were calculated for long repeat, oddball and random tests, all relative to long Baseline tests, by subtracting median reproduced test durations from median reproduced durations for baseline tests (see Table 2 for median reproduced durations for each condition).

Difference scores for repeat (M = -28ms, SD = 58ms; t_{19} = 2.10, p = .049) oddball (M = 15ms, SD = 66ms; t_{19} = 0.98, p = .336) and random (M = -10ms, SD = 61ms; t_{19} = 0.75, p = .460) tests suggested that each of these conditions was associated with a *similar* perceived duration, relative to single isolated tests (see Figure 5a). Note that these results were indexed against an appropriately conservative alpha (0.017), and that no differences were found when short test data were used in analyses (p > .410). We had anticipated this last set of results and had pre-registered that these data would be excluded from analyses.
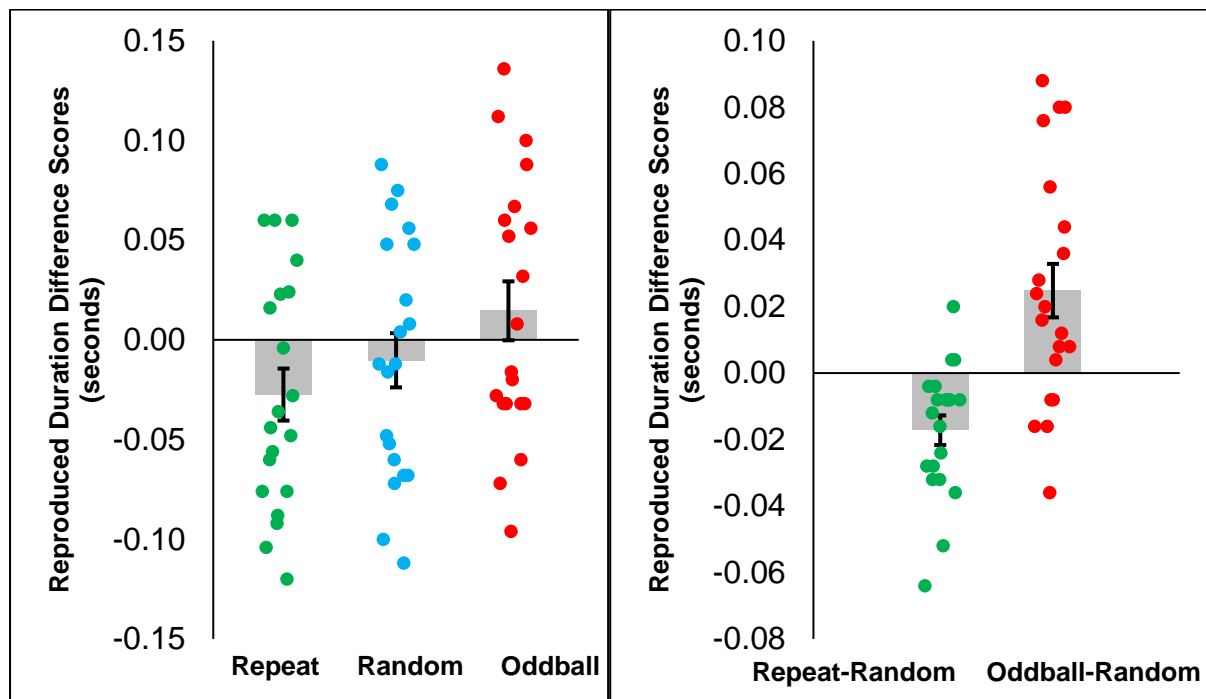

Figure 5 **a)** Reproduced duration difference scores for the long (800ms) trials. Positive difference scores indicate longer reproduced durations relative to single baseline, whereas negative difference scores indicate shorter reproduced durations. Grey bars depict average difference scores. Black error bars depict +/- 1 standard error. Coloured circles depict individual difference scores. **b)** Difference

scores for oddball and repeat tests, relative to random tests (as per Experiment 1). Other details are as for Figure 5a.

Results described for this experiment up to this point were pre-registered. Hereafter, we describe additional follow-up analyses concerning reproduced durations.

Bayes factor analyses suggest that reproduced durations for oddball ($BF_{10} = 0.356$) and random ($BF_{10} = 0.299$) tests had *matched* reproduced durations for single tests (the null hypothesis). For repeat tests, a Bayes factor analysis provided anecdotal evidence for the alternate hypothesis ($BF_{10} = 1.411$), that repeat test reproductions would be *shorter* than single test reproductions.

To aid comparison across experiments, results were re-analysed in the same way as Experiments 1 and 2, with difference scores calculated for oddball and repeat tests relative to random tests. Note, however, that these experiments had different task demands (a single response per trial in Experiments 1 & 2, different sequential responses in Experiment 3) and different procedures (visual feedback re time reproductions in Experiment 3, no feedback Experiment 1). We thus caution against a direct comparison of absolute time reproductions across experiments, but consideration of the patterns of conditional differences *within* each experiment should be informative.

Difference scores calculated on the basis described above suggest that participants reproduced *shorter* durations for repeat, relative to random tests ($M = -17$ms, $SD = 20$ms; $t_{19} = 3.86$, $p = .001$) and *longer* durations for oddballs relative to random tests ($M = 25$ms, $SD = 36$ms; $t_{19} = 3.08$, $p = .006$). There was also a difference between these two sets of difference scores ($t_{19} = 4.84$, $p < .001$; see Figure 5b).

As in Experiment 1, if interval timing error had increased in proportion to Weber's law (Treisman, 1963), variability for long (800ms) tests should have been greater than that for short (500ms) tests (by a factor of 1.6). The actual increase was markedly less (1.04), and the variance between reproductions for the two test durations was statistically equivalent (500ms condition SDs: $M = 106$ms, $SD = 34$ms; 800ms condition SDs: $M = 110$ms, $SD = 41$ms; $t_{19} = 0.46$, $p = .651$). This validates our omission of short tests from statistical appraisal of these time reproductions.

Sensitivity thresholds. The basis for our analyses were the JND threshold scores for each individual, for each test condition (see Table 4 for mean of JND thresholds).

Table 4	Mean spatial frequency JND threshold scores for each condition	
Condition	All (500 & 800ms)	
	M	*SD*
Baseline	0.037	0.016
Repeat	0.076	0.038
Random	0.064	0.026
Oddball	0.049	0.025

Difference scores. As per our pre-registered analysis plan, sensitivity difference scores were calculated as above (for data from both test durations), using JND threshold estimates (as per Experiment 2). Difference scores showed that participants were *less sensitive* to spatial frequency differences in repeat ($M = 0.04$, $SD = 0.03$; $t_{19} = 5.10$, $p < .001$) and random ($M = 0.03$, $SD = 0.03$; $t_{19} = 3.95$, $p < .001$) tests relative to single isolated baselines, whereas sensitivity for oddballs was *similar* to isolated baselines ($M = 0.01$, $SD = 0.03$; $t_{19} = 2.04$, $p = .055$, see Figure 6a)

Results described for this experiment up to this point were pre-registered. Hereafter, we describe additional follow-up analyses concerning sensitivity thresholds.

Bayes factor analyses suggested that sensitivity had been different for *all* test conditions, relative to single baselines (the alternate hypothesis). In each case evidence was in favour of *worse* test sensitivity. This evidence was very strong for repeat tests ($BF_{10} = 415$), strong for random tests ($BF_{10} = 42$), and anecdotal for oddball tests ($BF_{10} = 1.287$).

To aid comparison across experiments, sensitivity results were re-analysed as per Experiment 2. These difference scores suggested sensitivity to content in repeat and random tests was equivalent ($M = 0.01$, $SD = 0.04$; $t_{19} = 1.54$, $p = .140$), whereas sensitivity for oddball tests had been *greater* relative to random tests ($M = -0.02$, $SD = 0.02$; $t_{19} = 2.96$, $p = .008$). There was also a difference between these two sets of difference scores ($t_{19} = 3.28$, $p = .004$; see Figure 6b).
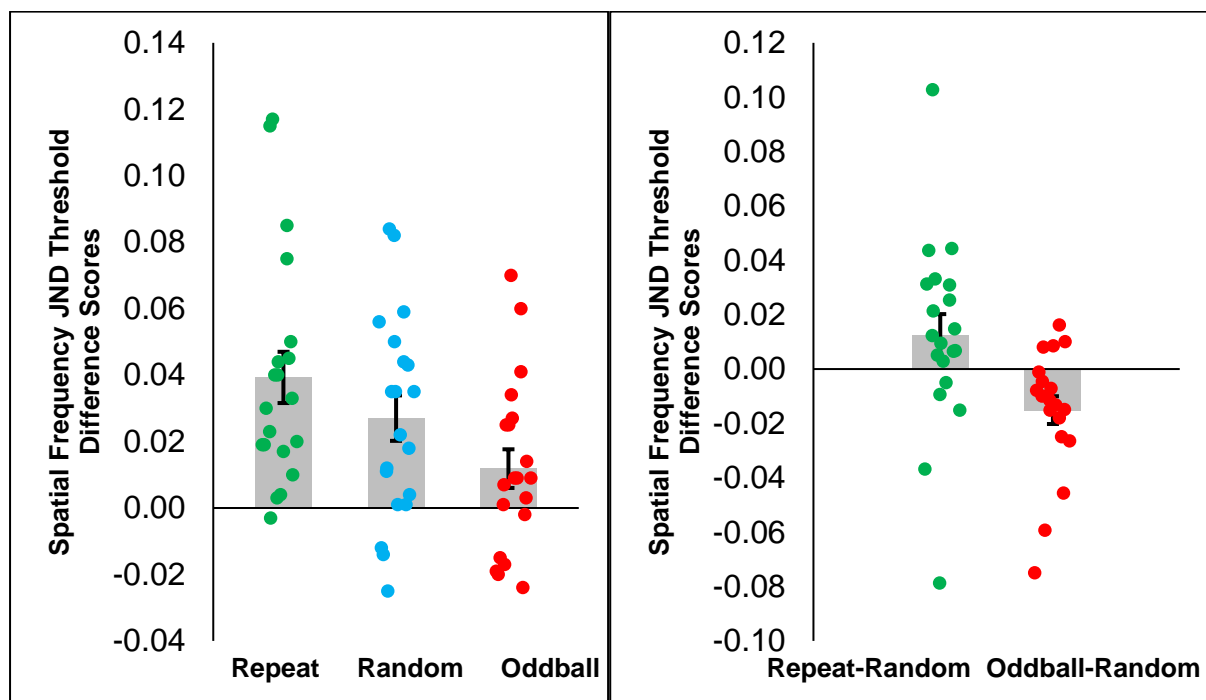

Figure 6. **a)** Spatial frequency JND threshold difference scores. Positive difference scores indicate decreased sensitivity relative to single baseline, while negative difference scores indicate increased

sensitivity. Grey bars depict average difference scores. Black error bars depict +/- 1 standard error. Coloured circles depict individual difference scores. **b)** Difference scores for oddball and repeat tests, relative to random tests (as per Experiment 2). Other details are as for Figure 6a.

Discussion

The pattern of results across Experiments 1 to 3 are clear for sensitivity, but less clear for time perception. In both Experiments 2 and 3 people were more sensitive to oddball content relative to random and repeat tests (see Figures 4 & 6). Moreover, sensitivity to oddball content was approximately *equal* to sensitivity for single tests (Experiment 3), whereas sensitivity for repeat and random tests was markedly *reduced* relative to this same 'baseline' (see Figure 6a). Time perception results were more variable.

In Experiments 1 and 3 evidence suggested people had reproduced *longer* times for oddballs relative to random tests (at least for longer tests – 650ms in Experiment 1, 800ms in Experiment 2), but had reproduced similar times for repeats and random tests (see Figures 2 and 5b). However, in Experiment 3, when these three test conditions were indexed relative to single isolated tests (as per our pre-registered analysis plan), there were no robust differences for *any* test condition (see Figure 5a). Overall, we felt time perception results to this point had been underwhelming, so we decided to attempt a further experiment, addressing shortcomings of Experiments 1 and 3.

The most glaring fault of Experiments 1 and 3 was that we effectively eliminated *half* our data from critical analyses, as motor planning noise had seemed to disproportionately impact reproductions of *shorter* tests. We were determined not to repeat this mistake. It was not clear, however, precisely what would constitute a 'too short' test, or if this would be the same for all participants. So, instead of trying to pre-determine what duration might be sufficiently long, and testing just 2 durations above that limit, in Experiment 4 we decided instead to test a much

larger number of test durations, ranging from potentially too short (250ms) to sufficiently long (550ms, suggested by pilot testing). This approach allowed us to extract information from intermediate test durations, and to investigate reproduction results via regression analyses that encompass *all* data.

Experiment 4

Methods

Details were as for Experiment 3, with the following exceptions. (https://aspredicted.org/at2bp.pdf)

Participants. Following Experiment 3, we further downgraded our expectations of the temporal oddball effect size (to 0.55), so we estimated that a larger *N* of 30 would be necessary to provide sufficient power (0.83). Thirty-four participants were recruited for testing via a research participation scheme at the University of Queensland. None had participated in any of the earlier experiments. All reported having normal or corrected-to-normal visual acuity. Four were removed from analysis according to pre-registered exclusion criteria (see results section), leaving a total of 30 participants (the pre-registered target *N*; 16 males). Ages for the final sample ranged from 18 to 24 (*M* = 19.63, *SD* = 1.61).

Stimuli and apparatus. As no spatial frequency judgement was included, stimuli were reverted back to the single Gabor patches used in Experiment 1.

Procedure. There were two experimental blocks, for which order of completion was counterbalanced across participants. In the test block, each trial involved four to eight 400ms standard Gabor presentations, separated by 200ms ISIs. We adopted the shorter ISI to speed up

the trial sequence, in order to gather more responses while shortening the overall experiment. Pilot testing had suggested this shorter ISI would be sufficient. ISIs were followed by a test Gabor, presented for between 250 and 550ms (10ms intervals; 31 test durations). In the test block each type of trial (random, repeat, & oddball) was presented twice for each test duration, for a total of 186 individual trials, all completed in a random order. In the baseline block, a single test Gabor was presented (no standards). In this block each test duration was presented twice, for a total of 62 individual trials, all completed in a random order. After test Gabor presentations participants performed a duration reproduction task (as per Experiment 3).

Results

Exclusion criteria. Participants whose average reproduced durations for shorter-half tests ($<$ 400ms) were not significantly ($p < .05$) shorter relative to longer-half tests (> 400ms), regardless of trial type, were excluded from further analysis. Four participants were excluded on this basis.

Reproduced durations. The mean of the two presentations for each test duration per condition was used in subsequent analyses.

Table 5   Mean durations (ms) reproduced by each participant for each test condition, pooled across test durations, along with standard deviations between individuals.

Condition	All (250 to 550ms)	
	M	*SD*
Baseline	214	65
Repeat	195	56
Random	201	59
Oddball	211	61

Difference scores. As per our pre-registered analysis plan, difference scores were calculated for repeat, oddball, and random tests, all relative to isolated baselines, by subtracting baseline reproductions from tests of the same physical duration.

Participants reproduced *shorter* durations for repeat (M = -19ms, SD = 44ms; t_{30} = 8.76, p < .001; see Figure 7a) and random tests (M = -13ms, SD = 45ms; t_{30} = 5.69, p < .001) relative to single baselines. While participants reproduced numerically shorter durations for oddballs, relative to single baselines (M = -3ms, SD = 44ms), this difference was not robust (t_{30} = 2.15, p = .040, indexed relative to a Bonferroni corrected alpha of 0.017).
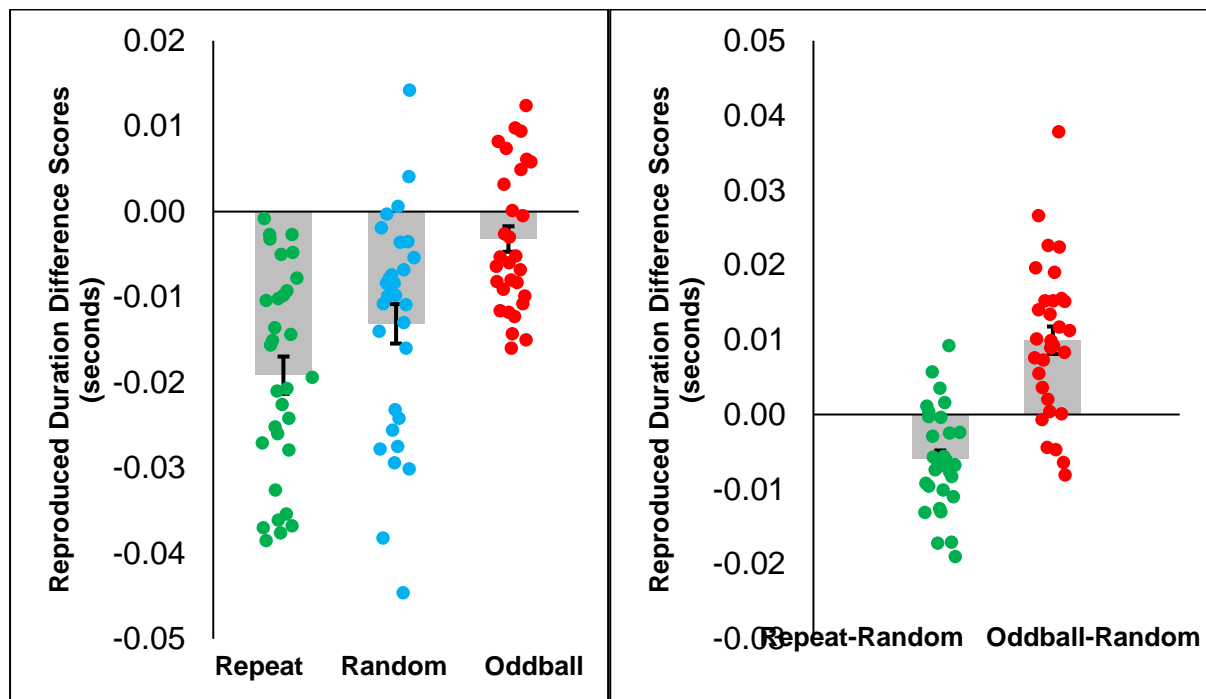

Figure 7. **a)** Reproduced duration difference scores for each test condition, relative to isolated single tests. Positive scores indicate longer reproduced test durations relative to single baselines, while negative difference scores indicate shorter reproduced test durations. Grey bars depict average difference scores. Black error bars depict +/- 1 standard error. Coloured circles depict individual difference scores. **b)** Difference scores for oddball and repeat tests, relative to random tests (as per Experiment 1). Other details are as for Figure 7a.

Regression. We conducted linear regression analyses on mean duration difference scores for each type of test, as a function of physical test durations (see Figure 8). These difference scores are indexed relative to single isolated tests. They reveal a negative relationship for both repeat

$(R^2 = 0.49, F_{1,29} = 27.30, p < .001)$ and random $(R^2 = 0.31, F_{1,29} = 13.00, p = .001)$ tests, with

people reproducing progressively *shorter* test durations relative to single isolated events as

physical test durations increase. There was no similarly robust relationship for oddball tests $(R^2$

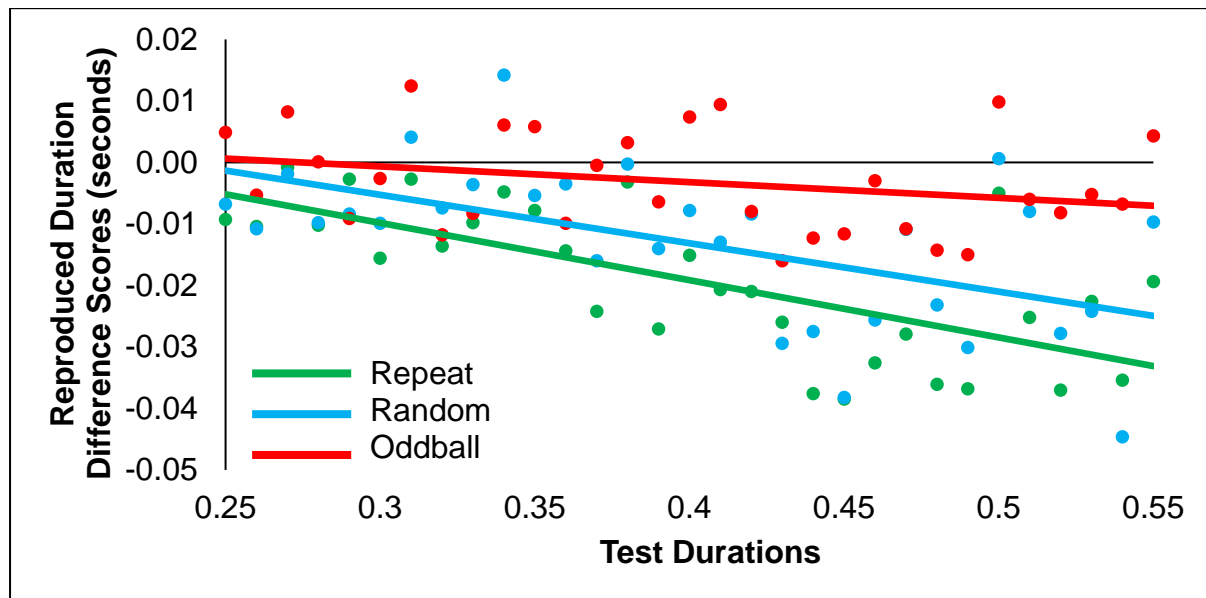$= 0.08, F_{1,29} = 2.47, p = .127)$.

Figure 8. Linear regressions overlayed on reproduced test duration difference scores, as a function of physical test durations.

Follow-up analyses. Results described for this experiment up to this point were pre-registered.

Hereafter, we describe additional follow-up analyses.

Bayes factor analyses showed that data were more likely under the alternate hypothesis (that

reproduced durations had been different relative to isolated baselines) for *all* test conditions. In

each case evidence was in favour of *shorter* reproduced test durations. This evidence was

extremely strong for repeat $(BF_{10} = 1.266e+07)$, and random tests $(BF_{10} = 5739)$, but anecdotal

for oddball tests $(BF_{10} = 1.428)$.

To aid comparison across experiments, we re-analysed results in the same way as Experiment 1. While we caution against a *direct* comparison of absolute reproduction times across experiments (due to different task demands and experimental protocols), a comparison of *patterns* of difference between the conditions in each experiment is meaningful. Difference scores conducted on this basis suggested participants reproduced *shorter* durations for repeat relative to random tests ($M = $ -6ms, $SD = $ 9ms; $t_{29} = 3.47$, $p = .002$), and *longer* durations for oddballs relative to random tests ($M = 10ms$, $SD = $ 14ms; $t_{29} = 3.87$, $p < .001$). There was also a difference between these two sets of difference scores ($t_{29} = 6.20$, $p < .001$, see Figure 7b).

General Discussion

Returning to the two core questions that motivated this project, we can answer one with confidence – presentation protocols that promote the TOE also modulate visual sensitivity. This is apparent from results of Experiments 2 and 3, where sensitivity to oddball content was enhanced relative to random tests (see Figures 4 and 6b). It is also important to note that sensitivity to repeat and random tests was equivalent (Figures 4 and 6b), that sensitivity for these tests was reduced relative to single isolated tests (Figure 6a), and that sensitivity to oddball and single isolated tests was closely matched (Figure 6a).

The other core question was whether the TOE results from a contraction of perceived duration for repetitive inputs, or from a dilation of perceived time for surprising events. We believe the overall pattern of results, across all experiments, argues in favour of a *contraction* of time for repetitive inputs.

The core assumption of the duration contraction account of the TOE (Pariyadath & Eagleman, 2007) is that surprising oddballs engender equal visual responses relative to the first event in a

train of presentations, whereas repetitive inputs elicit reduced responses. It is further assumed that perceived time is a metric of neural response magnitudes (Pariyadath & Eagleman, 2007). In Experiment 1 oddballs had an exaggerated duration relative to both repeat and random tests (at least for longer test durations, see Figure 2). The duration contraction account can explain these data, if repetition suppression magnitudes are similar for identical repeats and for trains of repetitive inputs. We think this is a reasonable proposition (as the 'random' events were physically identical, except for a *reliable* difference in orientation), but acknowledge that the duration contraction account would be more parsimonious with results if identical repeat durations had been *more* contracted than random tests – which brings us to Experiments 3 and 4.

In Experiments 3 and 4, repeat tests seemed shorter than random tests, and random tests seemed shorter than oddballs – precisely the ordering predicted by duration contraction (see Figures 5b, 7b and 8). All time perception results of Experiment 3 were equivocal when tests were indexed relative to single isolated events (see Figure 5a), but in Experiment 4 oddball test durations were equivalent to isolated tests, *largely* predictable 'random' tests had a reduced duration, and *entirely* predictable repeats had an even greater reduction (see Figures 7a and 8). Not only is this ordering predicted by the duration contraction account, so too is the equivalence of oddball and single isolated tests – as unanticipated inputs are thought to evoke a release from repetition suppression (Pariyadath & Eagleman, 2007).

How well does the duration dilation account for surprising inputs explain the same sequence of results? This can explain Experiment 1 results, provided oddballs are more surprising than events in a train of differently oriented inputs. This account is also broadly consistent with the ordering of apparent test durations in Experiments 3 and 4. This account is not, however,

consistent with comparisons involving single isolated tests. These were at least as predictable as random tests – arguably more so. Participants knew when a single isolated test stimulus, of an unknown orientation, would be presented and that they would have to reproduce its duration. Random tests had a similarly unknown orientation, but participants could not be sure if they were witnessing a test or a preliminary input, and the successive changes in orientation could reasonably be expected to engender some level of surprise. Random tests did not, however, seem equivalent to isolated tests in duration – they seemed shorter. This result is not at all consistent with a duration dilation for surprising inputs.

Predictive coding and the TOE. We believe that the TOE can be explained by extending the duration contraction account of the TOE into a predictive coding framework (Rao & Ballard, 1999; Rauss, Schwartz, & Pourtois, 2011). As repeated inputs are encountered, a prior expectation would be formed regarding subsequent events (Knill & Pouget, 2004), allowing the brain to become less responsive to predicted inputs (Feldman & Friston, 2010; Friston, 2010; Koch & Poggio, 1999). This would entail reduced sensitivity to slight changes in content from prior expectations, as input signals would be summated with predictions (Rao & Ballard, 1999; Rauss et al., 2011). Equal sensitivity to isolated tests and oddballs could also be expected, as in both cases one would not predict a strong prior regarding the precise spatial features of the test, so more emphasis should be placed on sampling of input (Rao & Ballard, 1999; Rauss et al., 2011). The predictive coding framework can therefore explain both the patterns of perceived test duration across conditions we have observed, and our findings regarding TOE protocols modulating sensitivity.

Are comparisons to an isolated baseline appropriate? It could be suggested that comparing isolated events to events at the end of a repetitive train is inappropriate, as the two classes of

event differ in too many ways. For instance, it could be asserted that the two classes of event differ in terms of entrained attention, temporal uncertainty, and or the levels of boredom that are elicited. All of these differences could variably impact perceived duration. There is no clear empirical evidence that these factors have impacted perceived duration in our experiments, and we note that in the absence of evidence one could equally make these assertions of any contrast of two experimental conditions – including comparisons of our different repetitive protocols. Further, we believe it is important to bear in mind the core prediction of the duration contraction account of the TOE – that repetitive events should have a reduced apparent duration due to repetition suppression (see Pariyadath & Eagleman, 2007). This prediction cannot be tested without a condition that is immune to the effects of repetition.

Of our repetitive protocols, results for our repeat and random conditions were most similar (in some cases indistinguishable), whereas random and oddball conditions were never equivalent. When we compared repetitive conditions to an isolated baseline (immune to the effects of repetition suppression), we found that the oddball condition was equivalent, whereas the repeat and random conditions were associated with *shorter* perceived durations and *reduced* visual sensitivity. We regard the overall qualitative pattern of results across all four experiments as highly consistent, and note that these are well-predicted by the duration contraction/predictive coding account of the TOE. Explaining the overall pattern of results via the duration dilation account of the TOE (Tse et al., 2004) would seem to require reliance on additional assumptions, for example, that differences between the isolated baseline and our repeat and random repetitive conditions (but not our oddball condition) interact in a manner yielding a relative expansion of the isolated stimulus, with a magnitude that is very similar to traditional the oddball effect. As there is currently no substantive evidence for this assumption, we prefer the

more parsimonious account, but recognise that the issue of an appropriate baseline in oddball tasks is contentious.

Our findings in context. How do our findings relate to previous similar studies? Ulrich et al. (2006) also examined the impact of expectation on perceived duration. They presented a pair of sequential stimuli on each trial, a standard then a frequent (70% of trials) or infrequent (30%) comparison. Infrequent (unexpected) comparisons were perceived as lasting longer, consistent with a negative relationship between expectation and perceived duration.

Cai et al. (2015) also investigated the impact of repetition and expectation on perceived duration, and found that while repetition produced a reliable duration contraction, a higher-order expectation violation (an inconsistent number following a sequence) had no strong impact (although this may have modulated a stronger repetition effect – see Figure 2 in that paper). These data are broadly consistent with the predictive coding framework, as identical repeated inputs should prompt strong prior expectations. The failure of high-level expectations might speak to the locus of these effects, as primarily sensory rather than cognitive

We are aware of one piece of apparently contradictory evidence, regarding the impact of expectation on perceived duration. Birngruber et al., (2018) had people predict which of two equally probable tests would be presented – *before* the test was presented. They found that predicted inputs had a relatively *exaggerated* duration, which was associated with heightened visual sensitivity to duration. While these data accord with a positive relationship between subjective duration and visual sensitivity, the impact of expectation on time perception seems *opposite* relative to the TOE and our data. These data may speak to a difference between intrinsically generated predictions, and ones that are guided by environmental statistics.

Conclusion. Overall our data are consistent with a predictive coding framework, in which repetitive events become anticipated, resulting in a reduction in apparent duration and sensitivity. Our data are less consistent with surprising inputs triggering a greater sampling of input. These possibilities should be further examined using a combination of psychophysical manipulations and recordings of brain activity. Regardless of the outcome of these future investigations, based on our investigation it would seem that when life delivers a succession of predictable events, time really does seem to fly.

Open Practices Statement: Experiments 1, 3, and 4 were pre-registered (links are provided at the beginning of each pre-registered experiment section). Contact corresponding author for data or analysis scripts.

References

Arnold, D.H., Nancarrow, K. & Yarrow, K. (2012). The critical events for motor-sensory temporal recalibration. *Frontiers in Human Neuroscience, 6*, 235 doi:10.3389/fnhum.2012.00235

Astle, A. T., Webb, B. S., & McGraw, P. V. (2010). Spatial frequency discrimination learning in normal and developmentally impaired human vision. *Vision Research, 50*(23), 2445-2454. doi:10.1016/j.visres.2010.09.004

Augustine, St. (1960). *The confessions of St Augustine*. New York, NY: Image Books.

Barlow, H. (1958). Temporal and spatial summation in human vision at different background intensities. *Journal of Physiology, 141,* 337 – 350.

Birngruber, T., Schröter, H., Schütt, E., & Ulrich, R. (2018). Stimulus expectation prolongs rather than shortens perceived duration: Evidence from self-generated expectations. *Journal of Experimental Psychology: Human Perception and Performance, 44*(1), 117-127. doi:10.1037/xhp0000433

Birngruber, T., Schröter, H., & Ulrich, R. (2015). The influence of stimulus repetition on duration judgments with simple stimuli. *Frontiers in Psychology, 6*, 1213. doi:10 .3389/fpsyg.2015.01213

Block, R. A., & Grondin, S. (2014). Timing and time perception: A selective review and commentary on recent reviews. *Frontiers in Psychology, 5*, 648. doi:10.3389/fpsyg.2014.00648

Cai, M. B., Eagleman, D. M., & Ma, W. J. (2015). Perceived duration is reduced by repetition but not by high-level expectation. *Journal of Vision, 15*(13), 1–17. doi: 10.1167/15.13.19

Clarke, F. J. J., & Belcher, S. J. (1962). On the localisation of Troxler's effect in the visual pathway. *Vision Research, 2*(1-4), 53-68. doi:10.1016/0042-6989(62)90063-9

Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience, 4*(215). doi:10.3389/fnhum.2010.00215

Friston, K. J. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience, 11*, 127-138. doi:10.1038/nrn2787

Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics, 72*(3), 561–582. doi:10.3758/APP.72.3.561

Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993). Focal visual attention produces illusory temporal order and motion sensation. *Vision Research, 33*(9), 1219-1240. doi:10.1016/0042-6989(93)90210-N

Kanai, R., & Watanabe, M. (2006). Visual onset expands subjective time. *Perception & Psychophysics, 68*(7), 1113-1123. doi:10.3758/BF03193714

Keane, B., Spence, M., Yarrow, K. & Arnold, D.H. (2015). Perceptual confidence demonstrates trial-by-trial insight into the precision of audio-visual timing perception. *Consciousness & Cognition 38,* 107 – 117.

Knill, D., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences, 27*(12), 712-719. doi:10.1016/j.tins.2004.10.007

Koch, C., & Poggio, T. (1999). Predicting the visual world: Silence is golden. *Nature, 2*(1), 9-10. doi:10.1038/4511

Marinovic, W. & Arnold, D.H. (2012). Separable temporal metrics for time perception and anticipatory actions. *Proceedings of the Royal Society of London, Series B: Biological Sciences 279,* 854 - 859.

Matthews, W. J. (2011). Stimulus repetition and the perception of time: The effects of prior exposure on temporal discrimination, judgment, and production. *PLoS ONE, 6*(5), e19815. doi:10.1371/journal.pone.0019815

Matthews, W. J., & Meck, W. H. (2016). Temporal cognition: Connecting subjective time to perception, attention, and memory. *Psychological Bulletin, 142*(8), 865–907. doi:10.1037/bul0000045

Mella, N., Conty, L., & Pouthas, V. (2011). The role of physiological arousal in time perception: Psychophysiological evidence from an emotion regulation paradigm. *Brain and Cognition, 75*(2), 182-187. doi:10.1016/j.bandc.2010.11.012

Mioni, G., Stablum, F., McClintock, S. M., & Grondin, S. (2014). Different methods for reproducing time, different results. *Attention, Perception, & Psychophysics, 76*(3), 675-681. doi:10.3758/s13414-014-0625-3

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research, 29*(11), 1631-1647. doi:10.1016/0042-6989(89)90144-2

Pariyadath, V., & Eagleman, D. (2007). The effect of predictability of subjective duration. *PLoS ONE, 2*(11), e1264. doi:10.1371/journal.pone.0001264

Patel, A., Maurer, D., & Lewis, T. L. (2010). The development of spatial frequency discrimination. *Journal of Vision, 10*(14):41, 1-10. doi:10.1167/10.14.41

Peters, J. C., van den Boomen, C., & Kemner, C. (2017). Spatial Frequency Training Modulates Neural Face Processing: Learning Transfers from Low- to High-Level Visual Features. *Frontiers in Human Neuroscience, 11*(1), 1-9. doi:10.3389/fnhum.2017.00001

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature, 2*(1), 79-87. doi:10.1038/4580

Rauss, K., Schwartz, S., & Pourtois, G. (2011). Top-down effects on early visual processing in humans: A predictive coding framework. *Neuroscience and Biobehavioural Reviews, 35*(5), 1237-1253. doi:10.1016/j.neubiorev.2010.12.011

Schindel, R., Rowlands, J., & Arnold, D. (2011). The oddball effect: Perceived duration and predictive coding. *Journal of Vision, 11*(2), 1-9. doi:10.1167/11.2.17

Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock." *Psychological Monographs, 77*(13), 1-31. doi:10.1037/h0093864

Tse, P., Intriligator, J., Rivest, J., & Cavanagh, P. (2004). Attention and the subjective expansion of time. *Perception & Psychophysics, 66*(7), 1171-1189. doi:10.3758/BF03196844

Ulrich, R., Nitschke, J., & Rammsayer, T. (2006). Perceived duration of expected and unexpected stimuli. *Psychological Research, 70*(2), 77-87. doi:10.1007/s00426-004-0195-4

Vierordt, K. (1868). *Der Zeitsinn nach Versuchen*. Tübingen, Germany: Laupp.

Wittmann, M. (2013). The inner sense of time: How the brain creates a representation of duration. *Nature Reviews Neuroscience, 14*(3), 217–223. doi:10.1038/nrn3452

Yarrow, K., & Rothwell, J. C. (2001). Manual chronostasis: Tactile perception precedes physical contact. *Current Biology*, *13*(13), 1134–1139. doi:10.1016/S0960-9822(03)00413-5