The Role of Premature Evidence Accumulation in Making Difficult Perceptual Decisions under Temporal Uncertainty

Ciara A. Devine[1], Christine Gaffney[1], Gerard Loughnane[1], Simon P. Kelly[2] & Redmond G. O'Connell[1]

[1] Trinity College Institute of Neuroscience and School of Psychology, Trinity College Dublin, The University of Dublin, Dublin 2, Ireland.

[2] School of Electrical and Electronic Engineering, University College Dublin, Dublin 4, Ireland.

Abstract

The computations and neural processes underpinning decision making have primarily been investigated using highly simplified tasks in which stimulus onsets cue observers to start accumulating choice-relevant information. Yet, in daily life we are rarely afforded the luxury of knowing precisely when choice-relevant information will appear. Here, we examined neural indices of decision formation while subjects discriminated subtle stimulus feature changes whose timing relative to stimulus onset ('foreperiod') was uncertain. Joint analysis of behavioral error patterns and neural decision signal dynamics indicated that subjects systematically began the accumulation process before any informative evidence was presented, and further, that accumulation onset timing varied systematically as a function of the foreperiod of the preceding trial. These results suggest that the brain can adjust to temporal uncertainty by strategically modulating accumulation onset timing according to statistical regularities in the temporal structure of the sensory environment with particular emphasis on recent experience.

Introduction

To date, perceptual decision making, described as the process whereby noisy sensory information is accumulated over time towards a decision bound, has been studied predominantly using tasks in which there is little ambiguity as to when observers should begin accumulating sensory information (Gold & Shadlen, 2007; Ratcliff & Smith, 2004; Shadlen & Kiani, 2013). It has been documented extensively throughout the temporal attention literature, that there are significant advantages, across a wide-range of perceptual tasks, to knowing precisely when sensory evidence will be available (for reviews see Correa, Lupianez, Madrid & Tudela, 2006; Jepma, Wagenmakers & Niewenhuis, 2012; Niemi & Näätänen, 1981; Nobre, Correa & Coull, 2007; Nobre & van Ede, 2018). However, in the natural sensory world we are not always afforded the luxury of making decisions based on sensory events that have the same high degree of perceptual salience and/or predictability in terms of their onset timing as those typically used in experimental tasks. Such uncertainty presents a critical challenge because mistiming accumulation onset can have detrimental consequences: premature accumulation can lead to false alarms and fast errors while commencing accumulation late can lead to information loss and slow response times or misses (Teichert, Grinband & Ferrera, 2016). The aim of this study was to shed light on how evidence accumulation is initiated when sensory evidence onsets are temporally uncertain and hard to detect.

Recent research has highlighted a few mechanisms that the brain could feasibly deploy. For instance, several studies have pointed to the possibility that, under some circumstances, the brain relies on sensory-driven mechanisms that detect the appearance of goal-relevant sensory changes to trigger evidence accumulation. For instance, the Gated Accumulator Model, suggests that evidence accumulation onset is contingent on sensory input surpassing a threshold level (Purcell et al., 2010; Purcell, Schall, Logan & Palmeri, 2012; Schall, Purcell, Heitz, Logan & Palmeri, 2011). Similarly, evidence from human electrophysiology suggests that in a continuous monitoring context, where the timing of target sensory changes are unpredictable, early target selection mechanisms are involved in triggering evidence accumulation (Loughnane et al., 2016). However, the amplitude of target selection signals was found to be heavily dependent on the strength of the sensory evidence and therefore it is not clear how broadly such a stimulus driven strategy can be applied. An additional possibility suggested by Teichert and colleagues (2016) is that the onset of sensory

evidence accumulation may be subject to top-down control and adapted in accordance with the temporal context or demands of the sensory environment. Specifically, decision onset timing may be calibrated internally in order to optimise the trade-off between accumulating information too early, thereby risking performance errors, and accumulating too late, thereby losing valuable information. In the context of temporal uncertainty, behavioural modelling and psychophysics studies have indicated that prior knowledge of sensory evidence onset does modulate the timing of accumulation onset (Bausenhart, Rolke, Seibold & Ulrich, 2010; Jepma et al., 2012; Seibold, Bausenhart, Rolke & Ulrich, 2011). However, these studies have relied on indirect behavioural metrics for inferring accumulation onset time and direct neurophysiological evidence has been lacking.

In this study we aimed to examine the mechanisms of perceptual decision making when subjects have imperfect foreknowledge of the precise timing of a subtle, goal-relevant sensory change. To this end, we conducted two experiments, both of which required subjects to discriminate a subtle change in a single stimulus feature (relative contrast of two overlaid gratings in experiment 1 and coherent dot motion in experiment 2). Temporal uncertainty was created by pseudorandomly imposing one of three delays intervening between stimulus onset and this sensory change (from hereon referred to as the ‘foreperiod’) on each trial. We traced activity in two previously characterised decision variable signals that trace sensory evidence accumulation: the centro-parietal positivity (CPP; Kelly & O'Connell, 2013; 2015; O'Connell, Dockree & Kelly, 2012; Twomey, Murphy, Kelly & O’Connell, 2015) and effector-selective Mu/Beta band (10-30 Hz) activity (de Lange, Rahnev, Donner & Lau, 2013; Donner, Siegel, Fries & Engel, 2009; O'Connell et al., 2012; Siegel, Engel & Donner, 2011; Steinemann, O’Connell & Kelly, 2018). Each of these signals builds gradually during decision formation at a rate proportional to evidence strength, peaks around the time of a decision-reporting movement, predicts choice accuracy and RT across trials and exhibits amplitude modulations consistent with decision bound adjustments (Kelly & O’Connell, 2013; O’Connell et al., 2012; Steinemann et al., 2018; Loughnane et al., 2016; Murphy et al., 2015). Unlike the decision signals observed in premotor regions of the human and non-human brain, the CPP provides a domain-general or abstract representation of an emerging decision, tracing evidence accumulation even in conditions where no overt action is required and motor preparation is absent (O’Connell et al., 2012; Twomey, Kelly & O’Connell, 2016).

We found that in this situation where subjects are unable to time the onset of decision formation to coincide with the uncertain onset of choice-relevant sensory evidence (contrast

change in experiment 1 or coherent motion onset in experiment 2), they systematically initiated evidence accumulation in advance of the evidence. This was reflected in the build-up of the CPP during the foreperiod, which, in experiment 2, directly reflected random fluctuations in motion energy, and also in choice-predictive lateralisation of Mu/Beta band activity. The timing of accumulation onset relative to the sensory change accounted for the pronounced foreperiod effects that were observed on choice behaviour including increased miss rates on short foreperiod trials and elevated premature responses and fast errors on long foreperiod trials. Furthermore, we found that accumulation onset timing was not fixed across trials but varied systematically according to recent trial foreperiod duration, with accumulation commencing earlier if the previous foreperiod was shorter. This also gave rise to a congruency effect whereby choice accuracy was greater if the foreperiod duration on consecutive trials was congruent.

Results

Experiment 1: Contrast discrimination

We analysed 128-channel EEG data from 23 human subjects performing a two-alternative contrast discrimination task (see materials and methods; figure 1a). in which they compared the contrast of two overlaid grating stimuli presented at fixation. At the outset of each trail, the gratings were presented and held constant at 50% contrast for one of three foreperiod durations (800ms/1200ms/1600ms). Once the foreperiod had elapsed, the gratings underwent antithetical changes in contrast (one increased while the other decreased), and the participant was asked to report whether the left or right tilted grating was higher in contrast by clicking the corresponding mouse button with their corresponding thumb. The magnitude of the contrast change for each grating was determined separately for each subject during preliminary training using a staircase procedure (choice accuracy at 65-70%, mean contrast: 5.5% ±1.35, see materials and methods). Participants were informed clearly at the beginning of the study about the variability in the foreperiod durations. This was demonstrated to them during practise trials in which the contrast change was very large and evidence onsets clearly noticeable. Points were awarded on each trial in order to encourage participants to respond both as quickly and as accurately as possible. 40 points were awarded for correct responses with a speed bonus of 0-40 additional points. Meanwhile, incorrect, missed or premature responses were awarded 0 points. Performance feedback was presented on a trial-by-trial (correct/incorrect/clicked too soon/too late) and block-by-block basis (accuracy, mean reaction time and cumulative points earned).

Behaviour

Taken together our behavioural data suggest that subjects adopted a decision strategy that was tailored to the average onset time of the contrast change (i.e. the middle foreperiod) which resulted in responses that were more closely timed with respect to stimulus onset than to evidence onset. First, there were significant main effects of foreperiod on RT (figure 1b-c; one-way repeated-measures ANOVA, $F(1.13, 22.56)=1717$, $p=1.68\times10^{-39}$), accuracy (figure 1d-e; $F(1.43, 28.64)=60.08$, 8.91×10^{-13}), missed responses (figure 1f; $F(1.07, 21.38)=114.29$, $p=2.88\times10^{-17}$), premature responses (i.e. prior to sensory change onset, figure 1g; F(1.01 20.12)=39.55, $p=1.38\times10^{-5}$) and points earned (figure 1h; $F(1.17, 23.49)=26.55$, $p=4.60\times10^{-}$

[8]). Post-hoc tests (see figure 1b-h; asterisks indicate significance levels) indicate that a longer (1600ms) foreperiod resulted in faster but less accurate responses and an increased tendency to respond prematurely. The opposite was true of short (800ms) foreperiod trials where responses were slower and less error prone but misses more common. Overall performance peaked on intermediate (1200ms) foreperiod trials which yielded the highest levels of accuracy and greatest number of points earned per trial with relatively low levels of both misses and premature responses.

Next, we analysed accuracy as a function of RT separately for each foreperiod ('conditional accuracy functions'; figure 1d). This analysis revealed two key trends. Firstly, the diminished accuracy on longer foreperiod trials manifested specifically on trials with faster RTs, reaching chance levels for the subset of long foreperiod trials with RTs closest to the contrast-change onset. This observation suggests that responses made close to the target change onset were were based on little or no valid sensory evidence. Secondly, the conditional accuracy functions also highlight that beyond 2400ms, accuracy diminished monotonically as a function of RT across all three foreperiod durations.

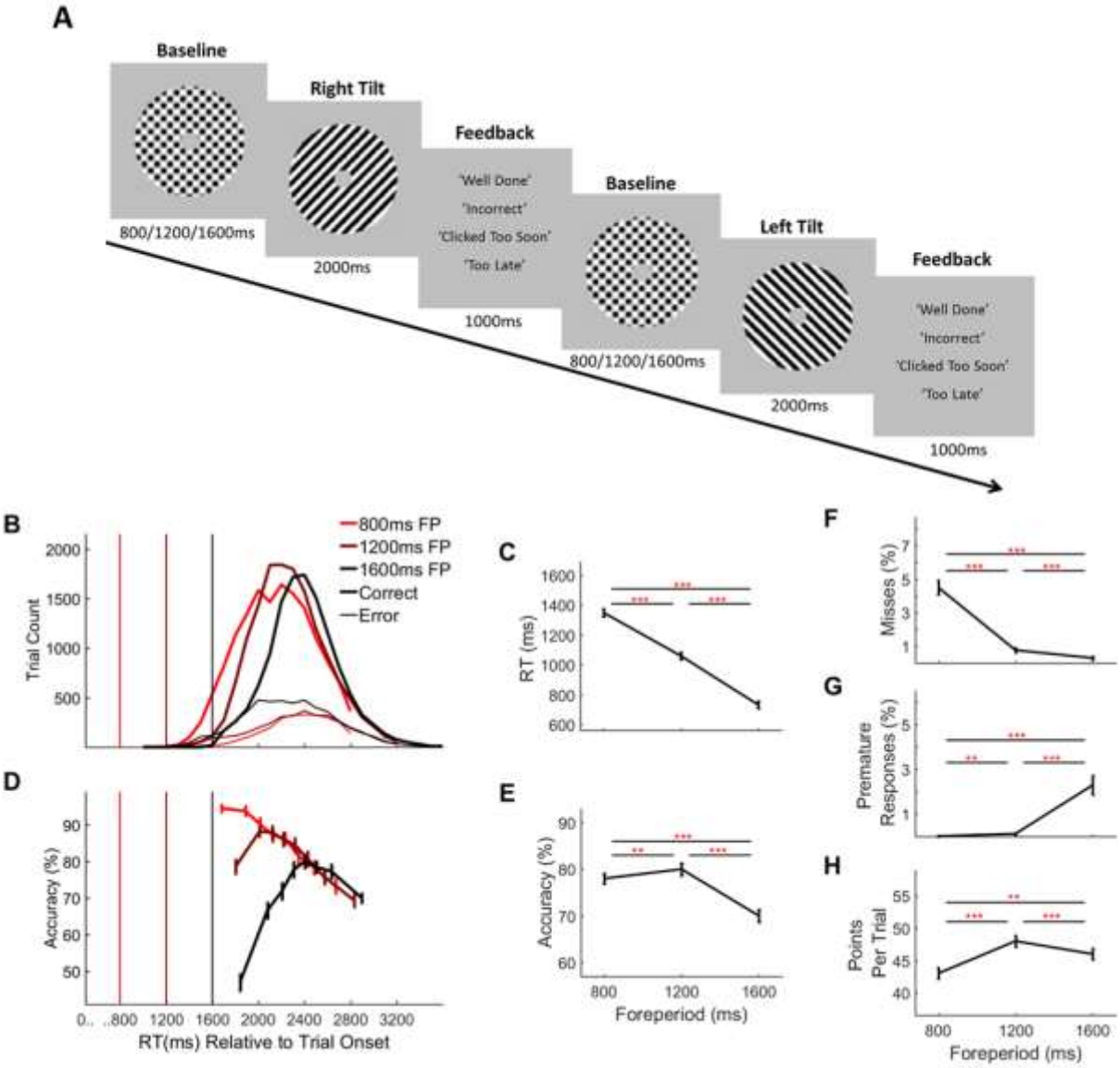


Figure 1 Two-Alternative Forced Choice Contrast Discrimination Task & Behaviour. A) Each trial commenced with the presentation of two overlaid left- and right-tilted gratings (45° relative to the vertical midline) within a circular aperture against a grey background. At baseline the gratings were each presented at 50% contrast. After an initial foreperiod, the duration of which varied unpredictably from trial-to-trial (800ms/1200ms/1600ms), one grating stepped up in contrast by an individually predetermined amount (*M*=5.5%, *SD*=1.35%, range=2-7%) while the other stepped down by a corresponding amount. Schematic depicts a right-tilted target followed by a left-tilted target with 50% contrast changes for illustration purpose only. Feedback was presented at the end of each trial. (B) Response time distributions separated by foreperiod. Vertical line markers indicate the times of the contrast change. (C) Average response times and D) accuracy, as a function of foreperiod. E) Conditional accuracy functions separated by foreperiod. Vertical line markers indicate the times of the contrast change. (F) Average missed response rate, (G) premature response rate and H) points earned

per trial, separated by foreperiod. Error bars represent standard error of the mean. Asterisks' indicate significance in post-hoc comparisons: **=p<.017, ***=p<.001; Bonferroni corrected critical p-value=.017.

Sensory Evidence Representation Unaffected by Foreperiod Duration

In order to determine whether sensory evidence representations were stable across the three foreperiod durations, we flickered each of the grating stimuli at a different frequency (20 Hz for left-tilted gratings, 25 Hz for right-tilted gratings) thereby eliciting separate SSVEPs in those corresponding frequency bands (20 Hz and 25 Hz respectively) over the occipital cortex (figure 2a). Key characteristics of these SSVEP signals indicate that they provide a read-out of the representation of the sensory evidence for contrast-based perceptual decisions: they are highly sensitive to modulations of stimulus contrast (Di Russo et al., 2007; Di Russo, Spinelli & Morrone, 2001; Di Russo, Teder-Sälejärvi & Hillyard, 2003; Kim, Grabowecky, Paller, Muthu & Suzuki, 2007; Norcia, Appelbaum, Ales, Cottereau & Rossion, 2015; O'Connell et al., 2012; Vialatte, Maurice, Dauwels & Cichocki, 2010) and, their amplitudes predict both response accuracy (Steinemann et al., 2018) and RT (Loughnane et al., 2018; O'Connell et al., 2012).

We analysed the SSVEPs as a function of foreperiod duration and stimulus (target vs non-target) using two-way repeated measures ANOVAs. As expected, the SNR of the target SSVEP (corresponding to the grating whose contrast increased) increased following the onset of the contrast change while that of the non-target (grating whose contrast decreased) decreased. Consistent with this we found a main effect of stimulus (target vs non-target) on SSVEP SNR in windows centered on 500ms post-contrast change (figure 2b; $F(1, 18)=13.93$, $p=.002$) and 200ms pre-response (figure 2c; $F(1, 18)=30.12$, $p=3.27\times10^{-5}$). These effects are also apparent in the difference SSVEP (d-SSVEP), which was calculated by subtracting the non-target from the target SSVEP (Figure 2d-e). Foreperiod on the other hand did not affect the SSVEP SNR (Post-contrast change: $F(2, 36)=.06$, $p=.94$; Pre-response: $F(1\ 18.40)=.01$ $p=.92$), irrespective of the stimulus (foreperiod x stimulus interactions: Post-contrast change: $F(2, 36)=.45$, $p=.64$; Pre-response: $F(1.05, 18.96)=.13$, $p=.73$). Together these findings indicate that the representation of sensory evidence was stable across the three foreperiods and could not account for the observed behavioural trends.

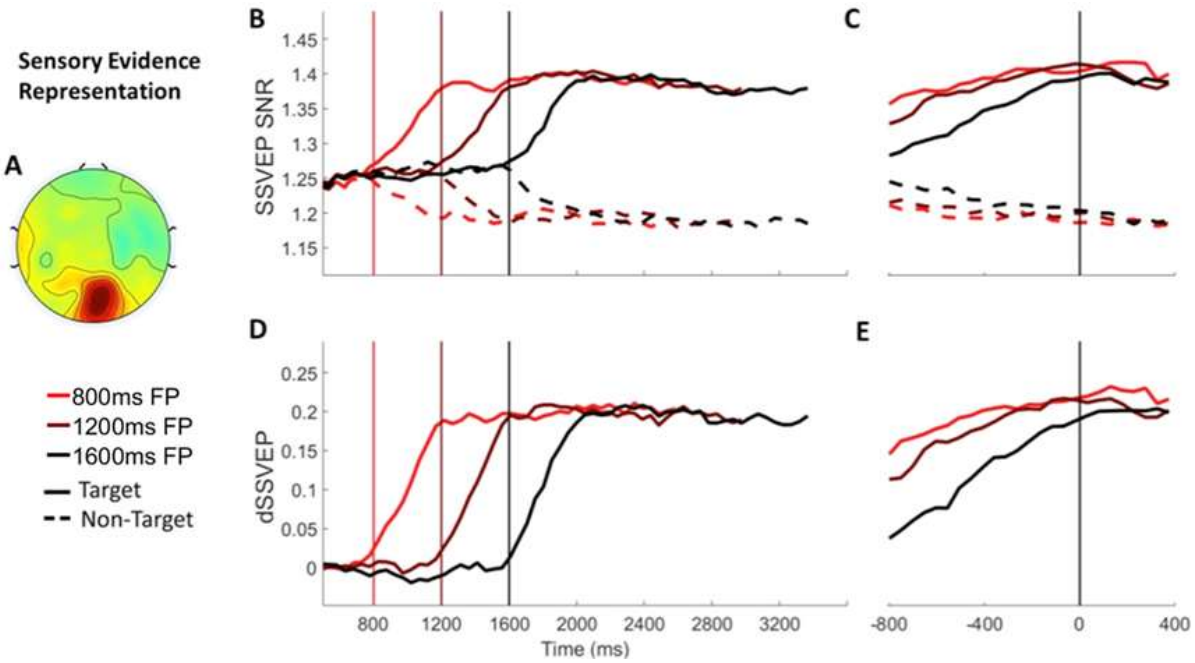


Figure 2 Sensory Evidence Representation (SSVEP) Signals Separated by Foreperiod. A) The topography of the d-SSVEP measured prior to response was maximal over the occipital cortex (Oz). B) Stimulus-aligned target and non-target SSVEP signals separated according to foreperiod duration, plotted relative to the onset of the stimulus. Vertical line markers at 800/1200/1600ms indicate the times of the contrast change. C) Response-aligned target and non-target SSVEP signals separated according to foreperiod duration. The vertical line marker at 0ms denotes the time of response. D) Stimulus-aligned d-SSVEP signal separated according to foreperiod duration, plotted relative to the onset of the stimulus. Vertical line markers at 800/1200/1600ms indicate the times of the contrast change. E) Response-aligned d-SSVEP signal separated according to foreperiod duration. The vertical line marker at 0ms denotes the time of response.

Premature Decision Formation on Trials with Longer Foreperiods

Figure 3b shows that, on longer foreperiod trials, significant CPP build-up was observed prior to the onset of the contrast change (figure 3b; one-sample t-tests: 800ms: $t(18)=1.79$, $p=.14$; 1200ms: $t(18)=3.71$ $p=.0016$; 1600ms $t(18)=5.52$, $p=2\times10^{-5}$). To explore this further, we analysed the CPP as a function of foreperiod duration and RT using two-way repeated measures ANOVAs. To this end we separated the data, within each foreperiod, into 6 equally sized bins, according to RT. Our analyses found that the CPP amplitude prior to evidence onset was larger on longer foreperiod trials ($F(1.27, 22.91=25.00$, $p=1.6\times10^{-5}$; linear contrast: $p=4.75\times10^{-5}$; quadratic contrast: $p=.17$). Moreover, there was an inverse relationship

between the amount of pre-evidence CPP build up and RT (figure 3d; $F(5, 90)=6.56$, $p=3x10-5$), irrespective of foreperiod duration (RT x foreperiod interaction: $F(5.33, 95.91)=1.07$, $p=.39$), indicating that greater pre-evidence CPP build-up was associated with faster responses. Further, the build-up rate of the CPP during the foreperiod (measured as the slope of a line fit to the CPP waveform in a 300ms window centered at 100ms prior to the contrast change) also predicted RT, irrespective of foreperiod duration (figure 3f; main effect of RT: $F(5, 90)=4.42$, $p=.001$; main effect of foreperiod: $F(2,36)=2.39$, $p=.11$; RT x foreperiod interaction: $F(4.81\ 86.54)=.63$, $p=.67$).

In light of the fact that the data from experiment 1 were collected over five separate experimental sessions, we sought to establish the degree to which the early build-up of the CPP during the foreperiod is a phenomenon that emerges only after extensive practice . As highlighted in supplementary figure 1, there was significant pre-evidence build-up of the CPP on longer foreperiod trials even on session 1 ($t(18)=4.12$, $p= 6.4x10^{-5}$) and the amount of pre-evidence CPP-build up did not vary systematically across training sessions (main effect of session: $F(4, 72)=1.18$, $p=.33$).

We also measured and analyzed CPP build-up rates in the response-aligned waveforms. There was no main effect of either RT ($F(5, 90)=.75$, $p=.59$) or foreperiod duration ($F(2, 36)=.56$, $p=.58$) but there was a significant cross-over interaction (figure 3g; $F(10, 180)=4.24$, $p=2.7x10^{-5}$), suggesting that the relationship between RT and pre-response CPP build-up rate was reversed for short and long foreperiod trials. Figure 3g illustrates that when the foreperiod was short, faster RTs were associated with steeper pre-response CPP slope. By comparison, the opposite was true of long foreperiod trials with faster responses having shallower pre-response CPP slopes. The shallower build-up rate of the pre-response CPP on faster RT trials can be explained in light of the close temporal proximity of these responses to the timing of the contrast change and hence the lesser contribution of valid sensory evidence to those decisions.

Alongside the early CPP build-up, we observed a progressive increase in motor preparation reflected in Mu/Beta desynchronisation over the hemispheres contralateral and ipsilateral to the response-executing hand throughout the pre-evidence foreperiod (figure 3i). To specifically isolate the choice-predictive component of motor preparation, we examined the differential motor preparation for the ultimately executed response relative to the withheld response, by computing the Mu/Beta lateralisation index (contralateral minus ipsilateral to the

response hand). This analysis indicated that there was significant pre-evidence Mu/Beta lateralisation on longer foreperiod trials (figure 3i inset; one-sample t-tests: 800ms: $t(18)=-1.27$, $p=.22$; 1200ms: $t(18)=-3.64$, $p=.002$; 1600ms: $t(18)=-3.91$ $p=.001$). To further examine this choice-predictive motor activity we analysed Mu/Beta lateralisation as a function of foreperiod and RT using two-way repeated measures ANOVAs. Longer foreperiods gave rise to greater pre-evidence Mu/Beta lateralisation ($F(2, 36)=3.64$, $p=.04$) with pre-evidence Mu/Beta lateralisation scaling inversely with RT irrespective of foreperiod duration (figure 3k; main effect of RT: $F(5, 90)=2.94$, $p=.017$; RT x foreperiod interaction: $F(5.22, 93.99)=1.68$, $p=.14$). This indicated that there was greater choice-selective motor preparation during the pre-evidence foreperiod predicted faster RTs. Further, the build-up rate of Mu/Beta lateralisation during the foreperiod (also measured as the slope of a line fit to the Mu/Beta lateralisation waveform in a 300ms window centered at 100ms prior to evidence onset) predicted RT irrespective of foreperiod duration (figure 3m; main effect of RT: $F(5, 90)=3.56$, $p=.006$; main effect of foreperiod: $F(2, 36)=.47$, $p=.63$; RT x foreperiod interaction: $F(4.82, 86.70)=1.45$, $p=.22$). As with the CPP, there was no main effect of foreperiod on the pre-response build-up rate (measured between -450 and -150ms) of Mu/Beta lateralisation ($F(2, 36)=.25$, $p=.78$) but there was a main effect of RT ($F(5, 90)=5.13$, $p=.01$) which was dependent on foreperiod duration (foreperiod x RT interaction: $F(10, 180)=2.08$, $p=.03$). This interaction indicates that the predicted pattern of faster pre-response Mu/Beta lateralisation build-up rates for earlier responses occurred only on short foreperiod trials but not when the foreperiod was longer (figure 3n).

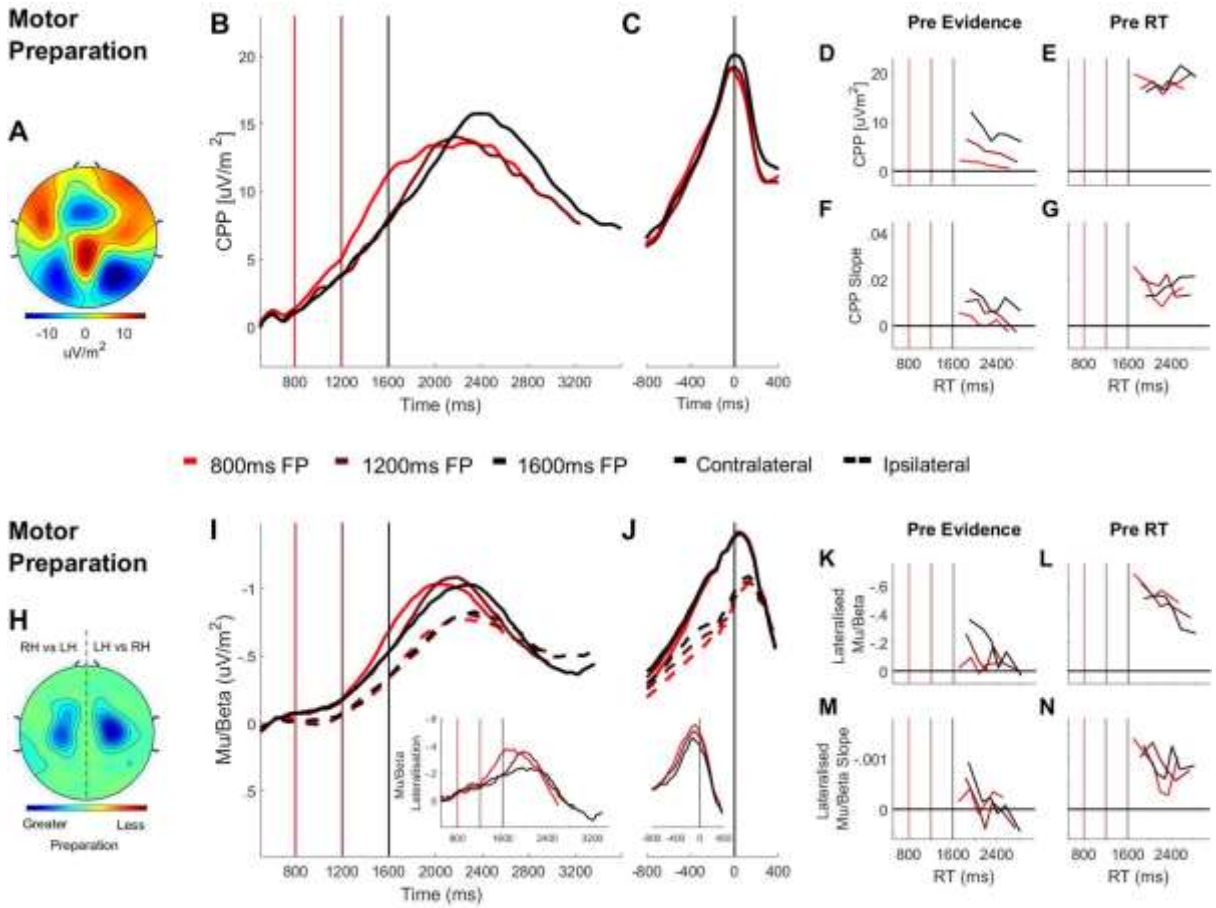
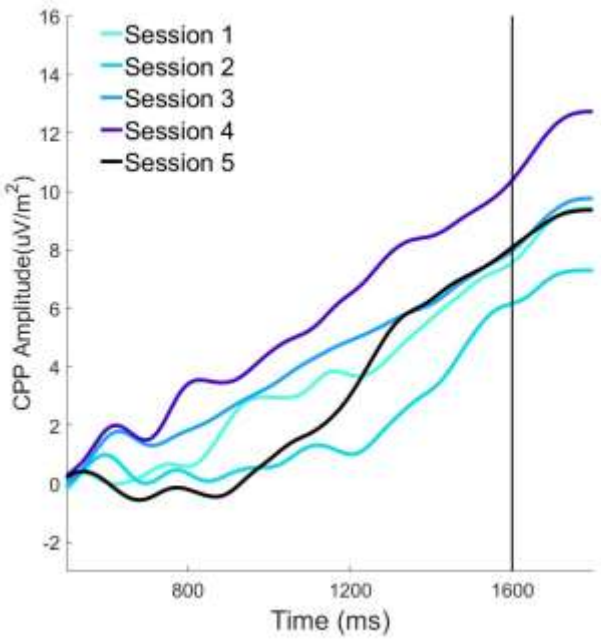


Figure 3 Domain-General (CPP) and Effector-Selective (Mu/Beta 10-30 Hz) Decision Signals Separated by Foreperiod (FP) in the Contrast Discrimination Task. A) Topography of the ERP signal measured prior to response (-150ms to -50ms) showing a positive going centroparietal component maximal over Pz. B) Stimulus-aligned CPP separated by foreperiod duration, plotted relative to the onset of the overlaid gratings stimulus. Vertical line markers at 800/1200/1600ms indicate the times of the contrast change across the three levels of foreperiod duration. C) Response-aligned CPP separated by foreperiod duration. The vertical line marker at 0ms denotes the time of response. D) CPP amplitude measured at contrast change (-50ms to 50ms) and E) at response (-150ms to -50ms) plotted as a function of RT separately for each foreperiod. F) Pre-evidence CPP build-up rate (-250ms to 50ms) and G) pre-response CPP build-up rate (-500ms to -200ms), plotted as a function of RT separately for each foreperiod. H) Topography of lateralised Mu/Beta band (10-30 Hz) activity measured prior to response (-150ms to -50ms) calculated separately for each hemisphere by subtracting ipsilateral from contralateral hand responses (LH=left hand; RH=right hand).The topography shows stronger lateralisation over each hemisphere when preparing a contralateral response. I) Stimulus-aligned contralateral and ipsilateral Mu/Beta waveforms, separated by foreperiod duration, plotted relative to the onset of the overlaid gratings. Vertical line markers at 800/1200/1600ms denote the times of the contrast change across the three levels of foreperiod

duration. Insert: stimulus-aligned Mu/Beta lateralisation (contralateral-ipsilateral) traces. J) Response-aligned contralateral and ipsilateral Mu/Beta waveforms, separated by foreperiod duration with a vertical line marker at 0ms denoting the time of response. Insert: response-aligned Mu/Beta lateralisation (contralateral-ipsilateral) traces. K) Mu/beta lateralisation at contrast change (-50 to 50ms) and L) response (-150ms to -50ms), plotted as a function of RT separately for each foreperiod. M) Pre-evidence Mu/Beta lateralisation slope (-250ms to 50ms) and N) pre-response Mu/Beta lateralisation slope (-500ms to -200ms) plotted as a function of RT separately for each foreperiod.

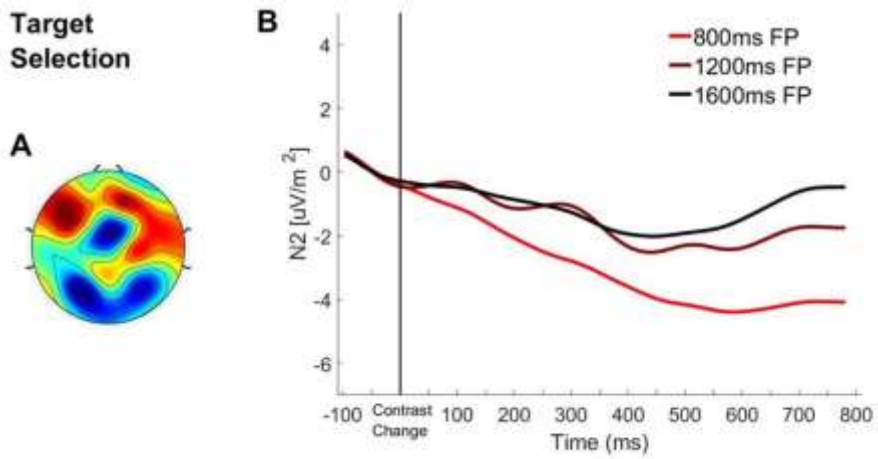


Supplementary Figure 1 Pre-Evidence CPP Build-Up as a Function of Session: The stimulus-aligned CPP, plotted for long foreperiod trials only, shows that the pre-evidence build-up occurs even in the very first testing session and that there is no systematic tendency for the degree of pre-evidence build-up to increase or decrease across sessions.

Subtle Sensory Evidence Onsets Do Not Elicit Early Target Selection Responses

Recent research and models of perceptual decision making have pointed to the importance of sensory-driven target selection mechanisms in triggering the onset of evidence accumulation (Purcell et al., 2010; Purcell et al., 2012; Schall et al., 2011). In the human brain, this target selection role has long been attributed to a negative potential over occipital scalp elicited approximately 200-300ms after the onset of a goal-relevant sensory stimulus

(Luck, 2012; Luck & Hilyard, 1994). Recently we showed that the target selection process is invoked bilaterally even for a single stimulus at fixation that has no ostensible need for "selection" and that in a continuous monitoring context, where target onsets are completely unpredictable, the signal temporally precedes and has a predictive influence on the onset and rate of evidence accumulation (Loughnane et al., 2016). This highlights the potential general role for this target selection signal in triggering the accumulation process under conditions of temporal uncertainty, but given its reliance on the presence of strong, detectable stimulus transitions (Loughnane et al., 2016), it is not clear whether it could play a major role in the current task. In fact, here, we did not observe any deflection in bilateral occipital ERP waveforms in the N2 time window although there was a slow, sustained negativity over the bilateral occipital cortex (supplementary figure 2a-b). Thus, in contrast to previous continuous tasks with larger evidence steps, low-level target-selection processes do not appear to play a role in dealing with temporal uncertainty in the present context.



Supplementary Figure 2 Bilateral Occipital ERP aligned to Contrast Change: A) Topography of the ERP signal measured 200-300ms post-contrast change. B) ERP waveform over bilateral occipital cortex aligned to the contrast change with waveforms separated according to foreperiod duration.

Foreperiod Duration Does Not Affect Decision Signal Amplitudes at Response

Recent neurophysiological studies have highlighted that when decisions must be made to a strict deadline, time-dependent urgency signals serve to expedite the decision process, progressively lowering the quantity of cumulative evidence required to trigger commitment

(Steinemann et al., 2018 Murphy et al., 2016; Hanks, Kiani & Shadlen, 2014; Thura & Cisek, 2016). In a recent study we showed that these urgency effects manifested in a time-dependent reduction in CPP amplitude (Steinemann et al., 2018). Here however, we did not find any variation in the CPP amplitude prior to response as a function of foreperiod duration ($F(2, 36)=.47, p=.63$; figure 3c) or response time ($F(5, 90)=.64, p=.67$; figure 3e). Similarly, the amplitude of contralateral Mu/Beta at the time of response did not vary as a function of foreperiod duration (figure 3j; $F(2, 36)=.49, p=.63$) or response time (figure 3n; $F(2.16, 38.95)=1.80, p=.18$), consistent with a time-invariant action-triggering threshold at the motor level (Murphy et al., 2016; Steinemann et al., 2018). By comparison, ipsilateral Mu/Beta prior to response, though not modulated significantly by foreperiod duration ($F(1.46, 26.30)=2.82, p=.09$), was more desynchronised prior to response on slower RT trials ($F(1.86, 33.40)=6.50, p=3x10^{-5}$) reflecting more preparation of the unchosen response hand on these trials and a closer 'race' between the alternative outcomes. Correspondingly, mu-beta lateralisation prior to response was stronger on faster response trials ($F(2.68, 48.31)=3.99, p=.01$), a finding that is consistent with previous studies (Murphy et al., 2016).

Summary: Experiment 1

The key results from experiment 1 show that in the absence of foreknowledge about the precise onset timing of a subtle goal relevant sensory change, subjects were unable to align the onset of decision formation to sensory evidence onset. On longer foreperiod trials, subjects consistently commenced decision formation before the target contrast change occurred, resulting in a higher incidence of false alarms and fast errors. The premature onset of the decision process was reflected in the build-up of the CPP and the emergence of choice predictive Mu/Beta lateralisation during the foreperiod. However, these data do not allow us to definitively establish whether the early build-up of decision-related activity reflected the accumulation of sensory noise, as opposed to some other process such as urgency (e.g. Cisek, Puskas & El-Murr, 2009; Thura, Beauregard-Racine, Fradet & Cisek, 2012; Thura & Cisek, 2016). To address this in experiment 2, we used the random dot motion task in which variability in physical dot motion energy from one frame to the next can be directly quantified by applying dot motion energy filtering. Random variations in motion energy during the foreperiod were leveraged in order to directly test for behavioural and electrophysiological indications of premature evidence accumulation.

Experiment 2: Dot Motion Discrimination

We analysed 128-channel EEG data from 23 human subjects who performed a two-alternative forced choice random dot motion discrimination task in which participants were required to judge the dominant direction of motion in a cloud of moving dots (see materials and methods; figure 4a). On each trial the dots initially moved randomly (0% coherence) for the duration of an initial foreperiod that varied unpredictably from one trial to the next across three levels (800ms/1200ms/1600ms). Once the foreperiod had elapsed, a portion of the dots began to move coherently in either a leftward or rightward direction. The portion of coherently moving dots, which was determined separately for each subject using a staircase procedure (see materials and methods) was, on average, 7.42% (*SD*=2.61%). Subjects were required to indicate the perceived direction of the coherent motion by clicking the corresponding (left/right) mouse button with their corresponding thumb (left/right). Participants were fully aware of the initial foreperiod and received feedback and points based on their performance following the same method implemented in experiment 1.

Behaviour

As in experiment 1, we first analysed behaviour separated according to foreperiod duration (summarised in Figure 4b-h) and observed highly similar trends including significant effects on RT (figure 4b-c; $F(1.07, 19.20)=150.60$, $p=1x10^{-10}$), accuracy (figure 4d-e; $F(1.47, 26.38)=7.69$, $p=.005$), miss rate (figure 4f; $F(1.08, 19.42)=63.76$, $p=9.7x10^{-8}$), premature response rate (figure 4g; $F(1.08, 19.40)=58.80$, $p=1.8x10^{-7}$) and points earned per trial (figure 4h; $F(142, 25.50)=10.10$, $p=.002$).

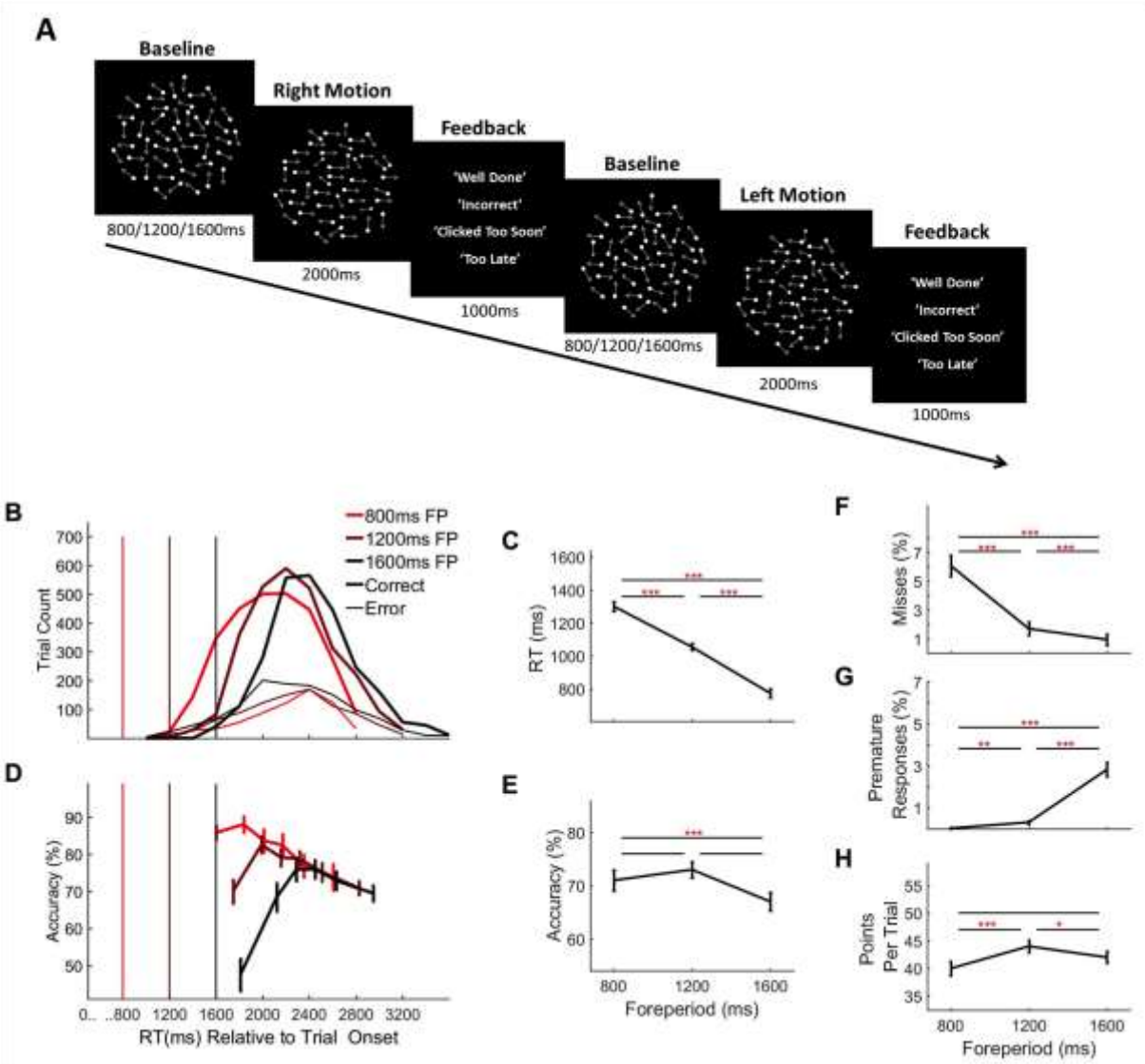


Figure 4 Random Dot Motion Discrimination Task & Behaviour. A) Each trial commenced with the presentation of a cloud of moving dots within a circular aperture against a black background. At baseline, the dots were displaced at random from one trial to the next. After an initial foreperiod, the duration of which varied unpredictably from trial-to-trial (800ms/1200ms/1600ms), a portion of the dots, determined separately for each individual subject (*M*=7.42%, *SD*=2.61%, range=3-12%), began to move coherently in either a leftward or rightward direction. The schematic depicts a rightward motion trial followed by a leftward motion trial (Note: Arrows are used here to illustrate the direction of motion but were not a feature of the actual stimulus). Feedback was presented at the end of each trial. (B) Response time distributions separated by foreperiod. Vertical line markers indicate the time of coherent motion onset. (C) Average response times and D) accuracy, as a function of foreperiod. E) Conditional accuracy functions separated by foreperiod. Vertical line markers indicate the time of coherent motion onset. (F) Average missed response rate, (G) premature response rate and H) points earned per trial, separated by foreperiod. Error bars represent standard error of the mean. Asterisks'

indicate significance in post-hoc comparisons: *=<.05, **=p<.017, ***=p<.001; Bonferroni corrected critical p-value=.017.

Decision Formation Commenced Prematurely on Longer Foreperiod Trials and Was Predictive of Behaviour

As in experiment 1 there was significant build-up of the CPP prior to the onset of the coherent motion (figure 5b) for longer foreperiod trials (one-sample t-tests: 800ms: t(17)=1.21 p=.24; 1200ms: t(17)=2.28, p=.03; 1600ms: t(17)=3.49, p=.003), with longer pre-evidence foreperiods resulting in larger CPP amplitudes at the time of coherent motion onset (F(2, 34) =5.45, p=.009). As in experiment 1, the build-up rate of the pre-evidence CPP did not vary as a function of foreperiod duration (F(1.45, 24.57)=1.96, p=.17). Likewise, pre-response CPP build-up rate was also invariant across foreperiod durations (ANOVA; main effect of foreperiod on pre-response CPP build-up rate: F(2, 34)=.04, p=.96).

Similar to the CPP, there was desynchronisation in Mu/Beta band activity over the hemispheres contralateral and ipsilateral to the response-executing hand throughout the pre-evidence foreperiod, with a progressive lateralisation toward the contralateral hemisphere on longer foreperiod trials (figure 5i; 800ms t(17)=-.40, p=.69; 1200ms: t(17)=-1.65, p=.12; 1600ms: t(17)=-3.23, p=.005) that increased as a function of foreperiod duration (F(3, 34)=3.49, p=.04). As reported in experiment 1 the rate of Mu/Beta lateralisation did not vary across foreperiod durations (F(3, 34)=.72, p=.49). Likewise, there was also no effect of foreperiod on the slope of pre-response Mu/Beta lateralisation (foreperiod main effect: (F(2, 34)=.09, p=.92).

Also in line with the results of experiment 1 the CPP amplitude at response did not vary significantly as a function of foreperiod duration (figure 5c; F(1.29, 21.92)=.04, p=.96) or response time (figure 5e; F(2, 34)=.52, p=.6). Similarly, the amplitude of contralateral Mu/Beta at the time of response did not vary as a function of foreperiod duration (figure 5j; main effect foreperiod: F(2, 34)=.13, p=.88) or response time (figure 5j; main effect response time: F(2, 34)=.74, p=.48).

Although the principle findings of experiment 1 were replicated in experiment 2, confirming the early onset of the decision formation process on longer foreperiod trials, some

trends from experiment 1 were notably absent in experiment 2. Specifically, the relationship between RT and pre-evidence CPP amplitude (F(2, 34)=.12, p=.89), pre-evidence CPP slope (F(1.45, 24.57)=1.96, p=.16) and pre-response CPP slope (F(2, 34)=2.08, p=.14) did not reach statistical significance. Likewise, the relationship between RT and the magnitude of Mu/Beta lateralisation in the pre-evidence (F(2, 34)=1.73, p=.19) or pre response time periods (F(2, 34)=1.80, p=.18), or the buildup rate in either the pre-evidence (F(2, 34)=.71 p=.50) or pre-response periods (F(2, 34)=.33, p=.72), did not reach statistical significance. A possible reason for these trends failing to replicate is the fact that there were far fewer trials in experiment 2, which consisted of a single experimental session (circa 500 trials per subject), than in experiment 1, which consisted of 5 separate experimental sessions (circa 2000 trials per subject). Consequently, when broken down by RT, our neural signal measurements may have been less robust in experiment 2 and therefore less sensitive to RT effects.

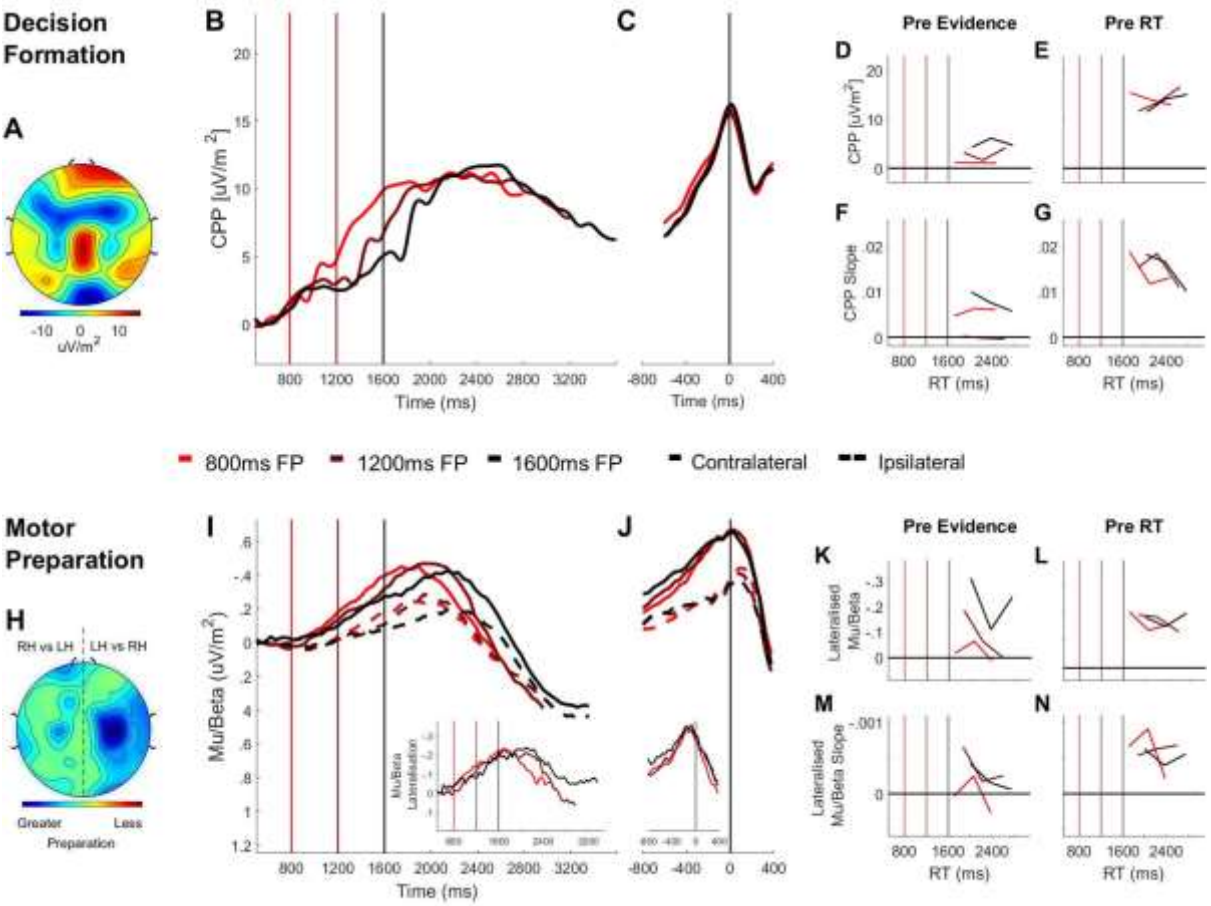


Figure 5 Domain-General (CPP) and Effector-Selective (Mu/Beta 10-30 Hz) Decision Signals Separated by Foreperiod (FP) in the Random Dot Motion Discrimination Task. A) Topography of the ERP signal measured prior to response (-150ms to -50ms) showing a positive going centroparietal component maximal over Pz. B) Stimulus-aligned CPP separated by foreperiod duration, plotted relative to the onset of the dot motion stimulus. Vertical line markers at 800/1200/1600ms indicate the times of coherent motion onset across the three levels of foreperiod duration. C) Response-aligned CPP separated according to foreperiod duration. The vertical line marker at 0ms denotes the time of response. D) CPP amplitude measured at coherent motion onset (-50ms to 50ms) and E) at response (-150ms to -50ms) plotted as a function of RT separately for each foreperiod. F) Pre-evidence CPP build-up rate (-250ms to 50ms) and G) pre-response CPP build-up rate (-500ms to -200ms), plotted as a function of RT separately for each foreperiod. H) Topography of lateralised Mu/Beta band (10-30 Hz) activity measured prior to response (-150ms to -50ms) calculated separately for each hemisphere by subtracting ipsilateral from contralateral hand responses (LH=left hand; RH=right hand). I) Stimulus-aligned contralateral and ipsilateral Mu/Beta waveforms separated by foreperiod duration, plotted relative to the onset of the dot motion stimulus. Vertical line markers at 800/1200/1600ms denote the times of coherent motion onset across the three level of foreperiod duration. Insert: stimulus-aligned Mu/Beta lateralisation (contralateral-ipsilateral) traces. J) Response-aligned contralateral and ipsilateral Mu/Beta waveforms, separated by foreperiod duration with a vertical line marker at 0ms denoting the time of response. Insert: response-aligned Mu/Beta lateralisation (contralateral-ipsilateral) traces. K) Mu/beta lateralisation at coherent motion onset (-50 to 50ms) and L) response (-150ms to -50ms), plotted as a function of RT separately for each foreperiod. M) Pre-evidence Mu/Beta lateralisation slope (-250ms to 50ms) and N) pre-response Mu/Beta lateralisation slope (-500ms to -200ms) plotted as a function of RT separately for each foreperiod.

Pre-Evidence Cumulative Motion Energy Predicts Fast Choices And Pre-evidence CPP Build-Up

In order to quantify the degree to which the dot motion stimulus by chance favoured one decision alternative or the other during the incoherent motion foreperiod, we performed dot motion energy filtering (see Methods). On the basis that the effects of random motion energy fluctuations should be most prominent for early choices, we separated long foreperiod trials into fast (<500ms) and slow responses (>500ms) and then calculated the integral of dot motion energy during the foreperiod (100ms-1600ms). One-sample t-tests confirmed that

there was significant sustained motion energy in the direction of the chosen response on fast RT trials (figure 6a; $t(17)=2.21$ $p=.04$) but not on slow RT trials (figure 6a; $t(17)=.03$, $p=.98$).

Next we sought to determine whether pre-evidence motion energy predicted the magnitude of pre-evidence CPP build-up. To this end we separated trials into two bins according to the slope of the pre-evidence dot motion integral (shallow vs. steep). This enabled us to to isolate trials where the cumulative motion energy during the foreperiod strongly and consistently favoured one choice or another by chance alone (figure 6b). For each subset of trials we measured the slope of the pre-evidence CPP over a broad window (600ms-1600ms) and the amplitude of the CPP in a 100ms window centered on the time of coherent motion onset (1600ms). We found that there was significant pre-evidence build-up of the CPP on trials where the cumulative motion energy consistently favoured a particular choice (figure 6c; one-sample t-tests: Amplitude: $t(17)=3.21$ $p=.005$; Slope: $t(17)=3.25$, $p=.005$) but not on those where the dot motion weakly and/or inconsistently favoured a specific choice (figure 6c; Amplitude vs 0: $t(17)=1.37$, $p=.19$; Slope vs 0: $t(17)=1.59$, $p=.13$). Paired samples t-tests comparing this pre-evidence CPP build-up between the two conditions (shallow vs steep pre-evidence motion integral) were significant (Amplitude: $t(17)=2.41$ $p=.03$) or close-to-significant (Slope: $t(17)=1.92$, $p=.07$) suggesting that the build-up of the CPP did reflect cumulative motion energy during the foreperiod.

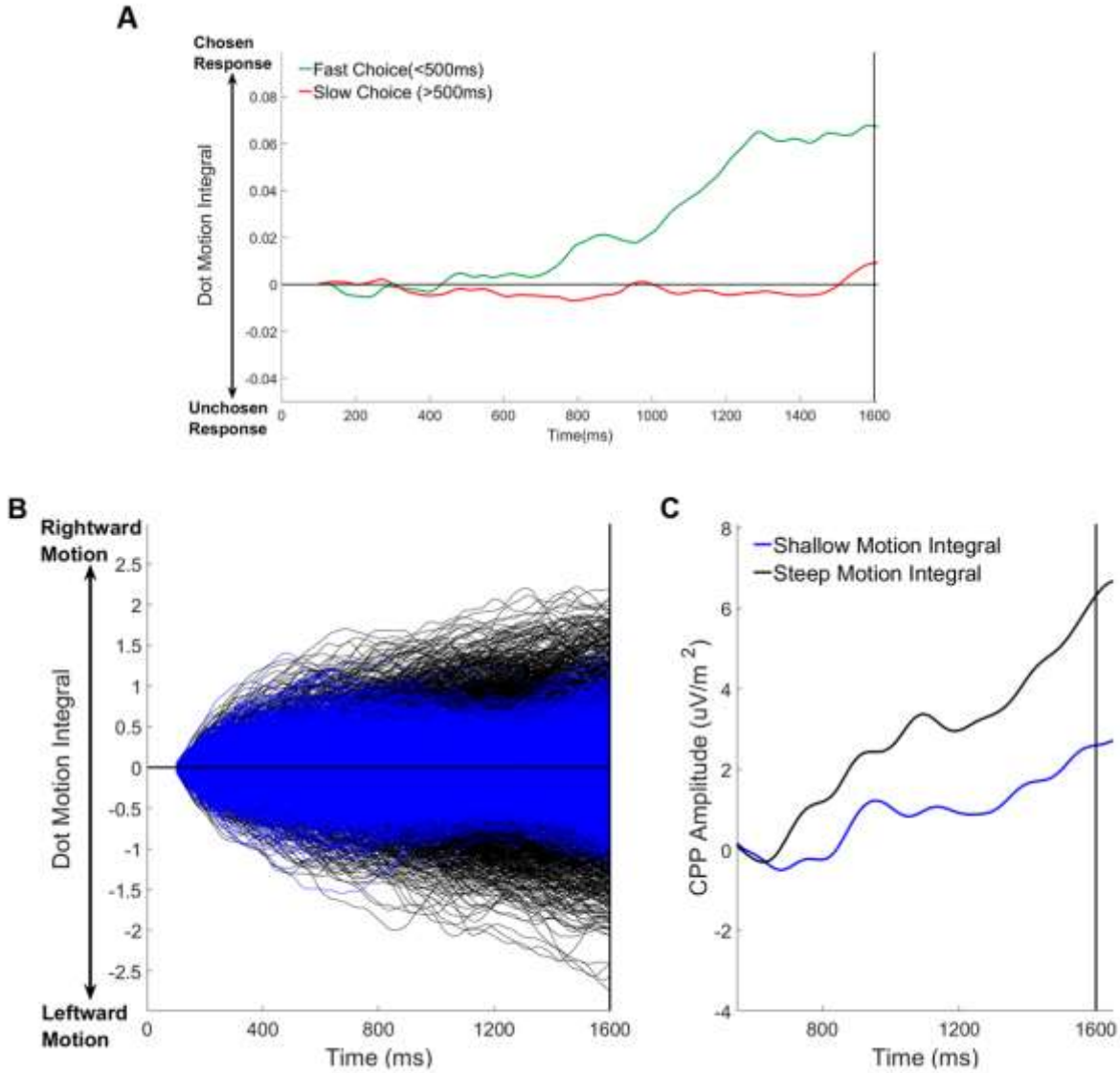
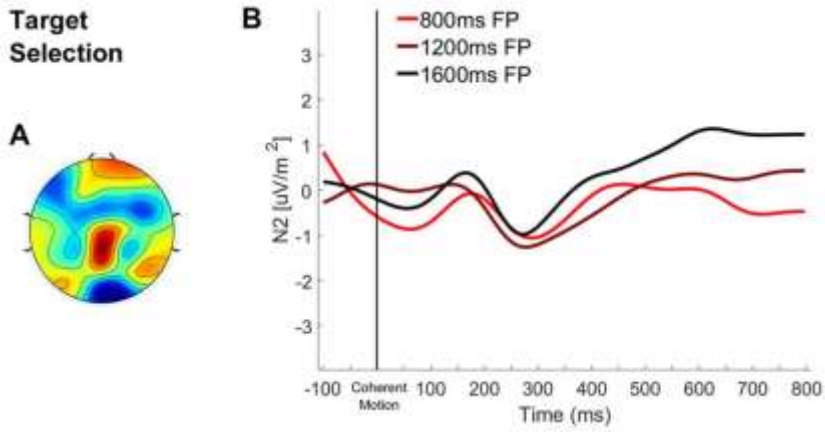


Figure 6 Relationship Between Pre-Evidence Cumulative Dot Motion Energy, Choice and Pre-Evidence CPP Build-Up (1600ms Foreperiod Trials Only). A) Average cumulative dot motion integral (measured between 100-1600ms) separated for fast (RT<500ms) and slow (RT>500ms) responses. B) Single trial cumulative dot motion integral traces separated according to the slope of the pre-evidence dot motion integral (100-1600ms). C) Grand average stimulus-aligned CPP separated according to the slope of the the pre-evidence dot motion integral.

Subtle Sensory Evidence Onsets Do Not Elicit Early Target Selection Responses

As in experiment 1, we were interested in establishing the extent to which target selection mechanisms played a role in decision formation in the context of discriminating a

subtle change in a sensory stimulus. Here, we observed a small, N2-like deflection in the time window 200-300ms relative to coherent motion onset over the bilateral occipital regions (supplementary figure 3a-b). However, relative to baseline, the magnitude of this component was not statistically significant for any of the foreperiods (all p>0.5 for t-tests against 0). Again, this suggests that low-level target-selection processes are not reliably elicited when coherent motion is weak.



Supplementary Figure 3 Bilateral Occipital ERP aligned to Coherent Motion: A) Topography of the ERP signal measured 200-300ms post-coherent motion onset. B) ERP waveform over bilateral occipital cortex aligned to coherent motion onset with waveforms separated according to foreperiod duration.

Decision Onset Timing is Modulated by Recent Temporal Structure

Thus far our data suggest that foreperiod effects on choice performance on these tasks can be attributed, in large part, to the premature onset of evidence accumulation occurring in this task. Next, we sought to examine the degree to which evidence accumulation onsets varied systematically as a function of trial history. Prior research in humans (Jazayeri & Shadlen, 2010) and in monkeys (Jazayeri & Shadlen, 2015) suggests that we exploit previous experiences and apply bayesian statistics in order to overcome uncertainty when estimating time or anticipating events. Specifically, these studies suggest that time estimates are affected by knowledge about the underlying distribution from which time intervals are drawn and are biased towards the mean of that distribution. Given that there was a uniform distribution of

the three foreperiods on our tasks, a bias towards the mean trial timing would lead to more favourable performance on intermediate foreperiod, which is what we observed.

To further explore this idea we conducted additional analyses of the data from experiment 1, hypothesising that, if decision onset timing is indeed affected by prior knowledge of temporal structure, then recent experiences may be particularly influential. For instance, in the temporal attention literature, the widely-documented sequential foreperiod effect suggests that the speed with which perceptual decisions are made is affected by previous trial timings (Capizzi et al., 2013; Los & Heslenfeld, 2005; Mento, 2013; 2017; Van der Lubbe, Los, Jaskowski & Verleger, 2004), though this phenomenon has not explicitly been linked or investigated with respect to the timing of evidence accumulation onset. Here, we separated trials according to the foreperiod on the current trial ($foreperiod_n$) and the foreperiod on the immediately preceding trial ($foreperiod_{n-1}$). We found that behavioural performance on $foreperiod_n$ was modulated significantly by $foreperiod_{n-1}$, such that responses were faster when $foreperiod_{n-1}$ was shorter (figure 7a; $F(2, 40)=140.17$, $p=2.49x10^{-19}$) and this effect was more pronounced when $foreperiod_n$ was long ($foreperiod_n$ x $foreperiod_{n-1}$ interaction: $F(4, 80)=7.37$, $p=4.17x10^{-5}$). Moreover, $foreperiod_{n-1}$ also affected accuracy ($F(2, 40)=5.23$, $p=.01$), but this effect was dependent on $foreperiod_n$ ($foreperiod_n$ x $foreperiod_{n-1}$ interaction: $F(4, 80)=14.90$, $p=3.85x10^{-9}$). The trends shown in figure 7b indicate that on short or long $foreperiod_n$ trials, foreperiod repetition led to a gain in accuracy whereas a change in foreperiod duration from one trial to the next was associated with a cost to accuracy. Furthermore, in line with the above effects, we found that the rate of missed responses on short $foreperiod_n$ trials was reduced if short foreperiods occurred consecutively (figure 7c; $F(2, 40)=36.51$, $p=9.52x10^{-10}$) while the rate of premature responses on long $foreperiod_n$ trials was reduced when long foreperiods occurred consecutively (figure 7d; $F(2, 40)=8.12$, $p=.001$).

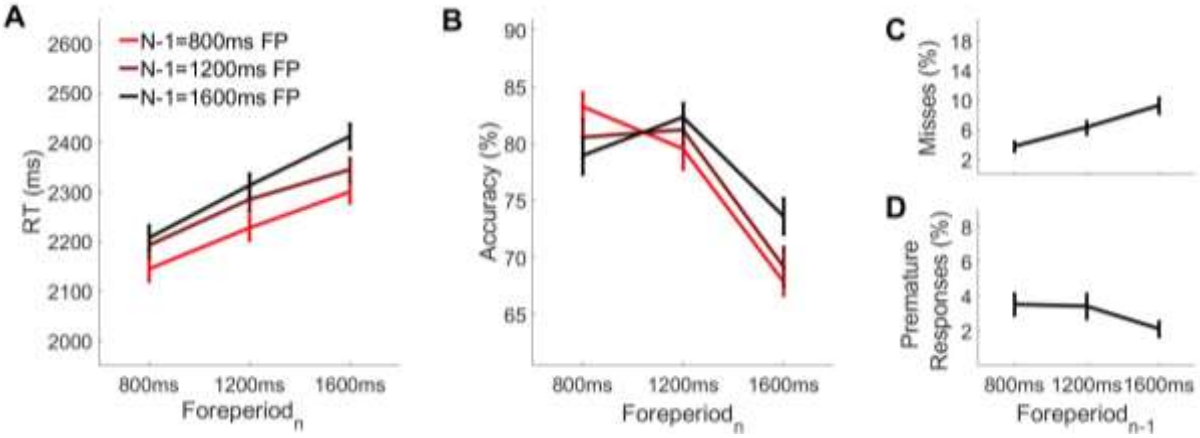


Figure 7: Behavioural performance as a function of foreperiod$_{n-1}$ and foreperiod$_n$.

A) On longer foreperiod$_n$trials, accuracy was higher if foreperiod$_{n-1}$was also long whereas on short foreperiod$_n$trials, accuracy was higher if foreperiod$_{n-1}$was short. **B)** Participants responded faster when foreperiod$_{n-1}$ was shorter. **C)** The rate of missed responses was higher when foreperiod$_{n-1}$was long. **D)**The rate of premature responses was higher when foreperiod$_{n-1}$was short.

Our analyses further show that the effects of foreperiod$_{n-1}$ on behaviour can be attributed, at least in part, to trial-by-trial modulation of evidence accumulation onset timing. As illustrated in figure 8a, the CPP reached a higher amplitude at the time of the contrast change when foreperiod$_{n-1}$ was shorter ($F(2, 36)=4.38$, $p=.02$). The absence of any effect of foreperiod$_{n-1}$ on pre-evidence CPP slope ($F(2, 36)=.54$, $p= .59$) suggests that the pre-evidence amplitude differences arose from differences in the onset, rather than the rate, of premature evidence accumulation. Together these findings indicate that decision onset timing is modulated by recent experience of temporal structure in the sensory environment.

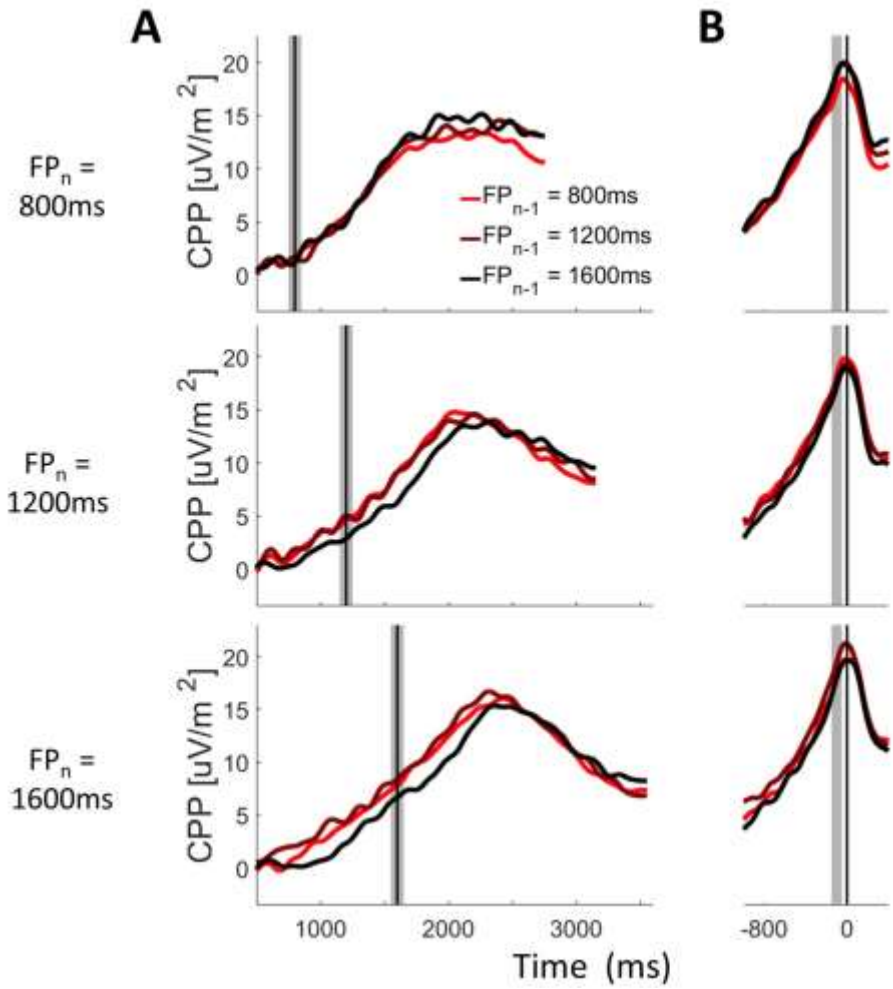


Figure 8: CPP as a function of $foreperiod_n$ and $foreperiod_{n-1}$

A) Stimulus-aligned and **B)** Response-aligned CPP waveforms separated according to $foreperiod_n$ and $foreperiod_{n-1}$. Shaded grey bars correspond to the pre-evidence (A) and pre-response (B) measurement windows. Solid black vertical lines mark the time of evidence onset (A) and response (B).

Discussion

Our investigation yielded three key findings. Firstly, in this situation where participants lacked exogenous cues marking the onset of temporally unpredictable sensory evidence, their choice behaviour reflected systematic variations in the onset of the neural evidence accumulation process, with longer foreperiods being associated with greater premature accumulation. Secondly, accumulation onset timing, which was not fixed across trials, appears to have been calibrated strategically and adjusted continuously based on experience of the underlying temporal structure of the task. Thirdly, our data suggests that participants could not rely on target selection mechanisms to initiate the onset of evidence accumulation. Recent work has highlighted that when sensory evidence onsets are unpredictable but relatively salient, early target selection signals are involved in triggering evidence accumulation onset (Loughnane et al., 2016; see also Purcell et al., 2010; 2012; Schall et al., 2011). Here however, our titration procedure ensured that the onset of the sensory changes were very weak and, by consequence, target selection signals were not reliably observed in our data forcing participants to rely on alternative strategies. Our data indicate that participants initiated evidence accumulation to coincide with the earliest possible evidence onset and, consequently, accumulated sensory noise during the foreperiods, resulting in faster, less accurate responses on trials with later evidence onsets.

In their recent review, Teichert et al. (2016) suggested that evidence accumulation onset timing can be adapted strategically in light of the particular challenges of a given scenario and is not merely triggered on the basis of sensory input. Two features of the data presented here suggest participants may have adapted the onset timing of accumulation based on their knowledge of the underlying temporal structure of the task. Firstly, overall performance peaked on intermediate foreperiod trials which yielded higher level of accuracy and points earned as well as low rates of both missed and premature responses. This pattern is interesting in light of research in the domain of time estimation, which suggests that under conditions of uncertainty the brain biases time estimates towards the centre of the distribution from which trial timings are drawn, a bayesian strategy known as regression to the mean (Jazayeri & Shadlen, 2010; 2015). Given that the foreperiod duration was drawn from a uniform distribution in the present study, the enhanced behavioural performance on intermediate foreperiod trials observed here, is consistent with a strategy of biasing the timing of evidence accumulation towards the centre of the foreperiod distribution. This can account

for the tendency to miss more targets on short foreperiod trials and to commit more premature responses on long foreperiod trials. This claim could be bolstered if, in future investigations, distinct behavioural patterns were observed when drawing the foreperiod from asymmetrical distributions. Secondly, we also observed a clear adjustment in evidence accumulation onset timing from one trial to the next, based on the duration of the foreperiod on the previous trial. According to the normative theories, intertrial dependencies such as this may reflect the inherent tendency for the brain to monitor and exploit local statistical patterns in the environment in order to reduce uncertainty (Jones, Curran, Mozer & Wilder, 2013; Wilder, Jones, Ahmed, Curran & Mozer, 2013; Wilder, Jones & Mozer, 2010; Yu & Cohen, 2008). Thus, these trial-by-trial adjustments in accumulation onset timing may reflect continuous updating of temporal expectations that inform accumulation onset timing. Together these two findings support the conclusion that adaptive, top-down factors impacted accumulation onset timing.

Our finding that evidence accumulation onset timing is adjusted from trial-to-trial also adds to a broader and growing body of research into the serial dependency in perceptual decision making. For instance, numerous studies have recently shed light on the fact that individuals exhibit a variety of idiosyncratic choice-history biases, which have variously been linked to biases in the starting point of evidence accumulation (Bode et al., 2012; de Lange et al., 2013; Yu & Cohen, 2008; Zhang, Huang & Yu, 2014) or biases in the weighting of sensory evidence (Braun, Urai & Donner, 2018; Kloosterman et al., 2018; Talluri, Urai, Tsetsos, Usher & Donner, 2018; Urai, de Gee, Tsetsos & Donner, 2019). Here we show that the timing of evidence accumulation onset is in part determined by the duration of the foreperiod on recent trials, with shorter foreperiods leading to earlier accumulation onset on subsequent trials and correspondingly faster RTs. Moreover, trial-by-trial adjustments in accumulation onset timing were also consequential for accuracy. When the same foreperiod duration was, by chance, repeated from one trial to the next, decision onset timing was more aligned to the physical sensory change than when the foreperiods were mismatched, and this incurred a corresponding benefit for accuracy. In turn, when the foreperiod mismatch on consecutive trials was maximal (i.e. short followed by long or vice-versa), decision onset timing was particularly misaligned with the physical sensory change and the greatest decrements in accuracy were incurred.

Our results also have the potential to shed light on the mechanisms underpinning a well-known phenomenon in the temporal attention literature, known as the sequential

foreperiod effect. This effect, which is another example of serial dependency, is characterised by the modulation of RT by previous trial foreperiod in the context of target detection tasks with variable foreperiods (Capizzi, Correa & Sanabria, 2013; Capizzi, Sanabria & Correa, 2012; Capizzi, Correa, Wojtowicz & Rafal, 2015; Los et al., 2017; Los & Van den Heuvel, 2001; Steinborn & Langner, 2012; Vallesi, McIntosh, & Stuss, 2009; Vallesi, Shallice, & Walsh, 2007). Our results replicate the sequential foreperiod effect using a task in which respondents performed a difficult perceptual discrimination task where the foreperiod varied pseudorandomly from trial to trial. In this context at least, the sequential foreperiod effect, the underlying mechanisms of which have not yet been fully explained, can be attributed to the adjustment in evidence accumulation onset timing. Interestingly, in the broader literature on temporal orienting, data from psychophysical and computational modelling studies have implicated decision onset timing in temporal orienting more generally (Bausenhart, Rolke, Seibold & Ulrich, 2010; Jepma, Wagenmakers & Nieuwenhuis, 2012; Rolke, 2008; Rolke & Hofmann, 2007; Seibold, Bausenhart, Rolke & Ulrich, 2011). Our findings provide the first neurophysiological evidence in support of this idea.

Finally, our study highlights how neural decision signals can be leveraged to gain insights into the timing of accumulation onset that are otherwise very difficult to discern through behaviour. Previous studies investigating decision onset timing have relied on indirect behavioural measures such as fast error rate, shifts in the leading edge of response time distributions, and estimates of non-decision time derived from sequential sampling models (Teichert et al., 2016). The use of non-decision time is problematic because the timing of evidence accumulation is just one of a number of factors, alongside afferent (e.g. sensory encoding, neural transmission) and efferent (e.g. motor execution) decision-related processes that contribute to this parameter. Our data indicate that it is possible to infer the timing of accumulation onset directly from the temporal dynamics of the CPP.

In summary, this study has shed light on the important role of accumulation onset timing in perceptual decision-making and has yielded two key insights in this regard. Firstly, we have shown that in the absence of foreknowledge about the precise timing of a subtle goal relevant sensory change, observers are unable to accurately time the onset of sensory evidence accumulation with respect to that event, resulting in instances where accumulation commences too early, leading to premature responses and fast errors, as well as instances where accumulation is delayed, leading to missed responses. Secondly, we have demonstrated that accumulation onset timing is not governed solely by sensory driven

mechanisms but instead can be controlled by top-down processes. To this end, we appear to rely on and continuously update our average representation of the temporal structure of a given task.

Materials and Methods

Subjects

Experiment 1: Contrast Discrimination Task

Twenty-three subjects aged 19-30 (12 males, 11 females) were recruited to take part in a two-alternative, forced choice, contrast discrimination task as part of a study on perceptual learning taking place over the course of 6 separate experimental sessions. This pre-planned sample size is consistent with previous electrophysiological studies conducted within our group (e.g. Murphy et al., 2015; O'Connell et al., 2012; Kelly & O'Connell, 2015; Steinemann et al., 2018; Twomey et al., 2015) and elsewhere (e.g. Diaz, Queirazza & Philiastides, 2017; McGovern, Webb & Peirce, 2012). All subjects were right-handed, had normal or corrected-to-normal vision and no history of personal or familial neurological or psychiatric illness. Subjects were requested to abstain from alcohol consumption and to maintain regular patterns of sleeping for the duration of the study so as to minimise potential confounds between sessions. One subject was excluded after reporting retrospectively that they had failed to comply with this prerequisite. One further subject dropped out after the second session and their data were excluded. Data from the final experimental session were excluded for one additional subject due to an error in setting the presentation monitor refresh rate on that day. This resulted in a final sample size of 21 (10 male, 11 female; Age: M=22.41, SD=2.96). For the purpose of electrophysiological analyses two further subjects were excluded, due to excessive blink and/or EEG artefacts (>40% trial loss), but their data were retained for behavioural analyses.

Prior to taking part, subjects were informed that they would receive a €10 reimbursement per testing session with the possibility of earning a performance bonus of up to €20 upon completing the study. In order to incentivise subjects, they were misled to believe that the exact value of their bonus was dependent on how many points (see design and procedure) they accumulated over the course of the study. In fact, all subject who completed the study were awarded the maximum additional bonus of €20 (total payment = €80) after being debriefed about the nature of the study and rationale for the deception. Written, informed consent was obtained from all subjects and all procedures were approved

by the Trinity College Dublin ethics committee and conducted in accordance with the Declaration of Helsinki.

Experiment 2: Dot Motion Discrimination Task

Nineteen subjects aged 18-37 (*M*=25.74, *SD*=5.92; 10 females, 9 males) were recruited to take part in a dot motion discrimination task comprising a single session of data collection. This pre-planned sample size is consistent with previous electrophysiological studies conducted within our group (e.g. Murphy et al., 2015; O'Connell et al., 2012; Kelly & O'Connell, 2015; Steinemann et al., 2018; Twomey et al., 2015). All subjects were right-handed, with normal or corrected-to-normal vision and had no history of personal or familial neurological or psychiatric illness. In the case of one subject, a file containing the data necessary for carrying out analyses of dot motion energy (see below) was missing. Their data were retained for all behavioural and electrophysiological analyses with the exception of those involving dot motion energy data.

Subjects were informed that they would earn €15 for taking part in the study with the possibility of earned an additional performance bonus of up to €10. In order to incentivise subjects, they were misled to believe that the exact value of the bonus was dependent on how many points they obtained over the course of the study. In fact all subjects were awarded the maximum bonus for completing the study after being debriefed about the nature of the study and rationale for the deception. Written, informed consent was obtained from all subjects and all procedures were approved by the Trinity College Dublin ethics committee and conducted in accordance with the Declaration of Helsinki.

Experiment 1: Two-Alternative Forced-Choice Contrast Discrimination Task

Subjects performed a difficult two-alternative contrast discrimination task, consisting of discrete trials, in which they were required to discriminate the direction (left or right) of a target (tilted grating stimulus) based on a change in the relative contrast between two overlaid grating stimuli (see figure 1a). The experiment consisted of 10 blocks, each containing 50 trials and was conducted in a dark, sound-attenuated room with subjects seated in front of a

51cm cathode ray tube (CRT) monitor (refresh rate: 100 Hz, 1024 x 768 resolution) at a distance of approximately 57cm. Stimuli used in this task were created using Psychtoolbox and the experiment was presented using Matlab. The stimuli consisted of two overlaid grating patterns (spatial frequency = 1 cycle per degree) presented in a circular aperture (inner radius = 11 outer radius = 6˚) against a dark grey background (luminance: 65.2 cd/m2). Each grating stimulus was tilted by 45˚ relative to the vertical midline (left tilt = -45˚, right tilt = +45˚). The gratings were "frequency tagged" in order to allow independent measurement of sensory evidence in favour of both possible choices, with the left-tilted grating flickering at 20 Hz and the right-tilted grating flickering at 25 Hz. On each trial the gratings were phase-reversed and these phase-reversals were pseudorandomly counterbalanced across trials.

Subjects commenced each trial by simultaneously clicking the left and right mouse button. A central fixation dot was presented, followed 200ms later by the presentation of the overlaid grating stimuli at 50% contrast. The stimuli were held constant for the duration of an initial foreperiod, which, as subjects were informed, varied pseudorandomly from trial to trial between 800ms, 1200ms and 1600ms. After the foreperiod, the gratings underwent opposite changes in contrast (the magnitude of which were determined individually using a staircase procedure - see *Staircase Procedure* below) whereby the target stepped up in contrast while the non-target stepped down in contrast by a corresponding amount. The gratings remained at the new contrast level for 2000ms and subjects were required to indicate the direction of the target stimulus by clicking the corresponding left or right mouse button with their left or right thumb. Feedback was presented at the end of each trial for 1000ms and at the end of each block (see *Feedback & Points* below).

Staircase Procedure:

The change in contrast of the grating stimuli was determined separately for each subject during an introductory phase of the study, which took place on the first of the six-day study. During this introduction phase, subjects were first trained to perform the task under relatively easy conditions (large changes in contrast). Once subjects performed the task under easy conditions at close to 100% accuracy, the magnitude of the contrast change was then adjusted downwards gradually over the course of 6-8 short blocks (20 trials each) until subjects were performing within the range of 65-70% accuracy. The average change in contrast resulting from this titration procedure was 5.5% (*SD*=1.35%, range: 2-7%).

Feedback & Points:

At the end of each trial (500ms post stimulus offset) feedback was presented on screen in text format, for 1000ms. This feedback indicated whether the subject responded correctly ("Correct"), incorrectly ("Error"), prematurely (responded before or within 150ms of the contrast change; "Clicked too Soon"), or failed to respond within the deadline ("Too Late"). At the end of each block, feedback was presented on screen informing subjects of their mean accuracy and response time during the block. This feedback remained on screen until either the subject or experimenter exited the screen. With this feedback, subjects were also given a score for each block reflecting their cumulative points earned during the block. Points were awarded on a trial by trial basis according to the accuracy and speed with which subjects responded. Every correct response was awarded 40 points plus a speed bonus, while incorrect, missed or premature responses were awarded 0 points. The maximum speed bonus was 40 points and this amount diminished linearly from 40 to 0 across the range of possible response times (150-2000ms). At the end of each block a bar graph was presented to subjects depicting their score for each completed block of the task.

Experiment 2: Random Dot Motion Discrimination Task

Subjects performed a difficult, discrete-trial version of the random dot motion discrimination task (figure 4a) in which they were required to judge the direction (left or right) of motion based on a cloud of moving dots. The experiment consisted of 8 blocks, each containing 66 trials and was conducted in a dark, sound-attenuated room with subjects seated in front of a 51cm CRT monitor (refresh rate: 100 Hz, 1024 x 768 resolution) at a distance of approximately 57cm. The stimuli used in this task, random dot kinematograms (RDK), were created using Psychtoolbox and the experiment was presented using Matlab. The RDKs consisted of a patch of 100 moving white dots (diameter = 4 pixels) presented within a circular aperture (outer radius = 4˚) against a black background. From frame to frame, the dots were displaced throughout the circular aperture creating a perception of motion.

At the onset of each trial, a central fixation dot was presented, followed 400ms later by the onset of the dot motion stimulus. Initially, the dots were displaced within the aperture

at random (0% coherence) giving rise to the perception of random, incoherent motion. The displacement of the dots remained at 0% coherence for the duration of a foreperiod, which as subjects were informed, varied pseudorandomly from trial-to-trial between 800ms, 1200ms and 1600ms. After this foreperiod a portion of the dots began to move coherently in the same direction (either leftward or rightward). This coherent motion was produced by displacing a portion of the dots on each frame, chosen at random, in a common direction relative to their position on the previous frame. All other dots on a given frame were displaced randomly to a new location within the aperture. The percentage of coherently moving dots was determined separately for each subject in the study during a pre-experimental introduction phase (see *"Staircase Procedure"* below). The duration of coherent motion was 2000ms and subjects were required to indicate its direction by clicking the corresponding left or right mouse button with their left or right thumb. As in experiment 1 feedback was presented at the end of each trial for 1000ms and at the end of each block (see *Feedback & Points* above).

Staircase Procedure:

The percentage of coherently moving dots was determined separately for each subject during an introductory phase prior to commencing the experiment. During this phase, subjects were first trained to perform the task under relatively easy conditions (high coherent motion). Once subjects were proficient at performing the task under easy conditions (were close to 100% accuracy), the difficulty of the task was adjusted using a 2-down 1-up staircase procedure over the course of an 80-trial block that titrated accuracy to 65-70% for each subject individually. Coherence levels ranged from 3-12% (M=7.42%, SD=2.61%).

***Motion Energy Filtering**:*

RDKs are suited to studying the dynamics of sensory evidence accumulation because the random displacement of dots from one trial to the next gives rise to objectively quantifiable variability in sensory evidence within a given trial around the nominal, pre-determined, motion coherence level. Momentary sensory evidence favouring leftward or rightward motion was calculated using dot motion energy filtering (Adelson & Bergen, 1985), using the implementation provided by Urai and Wimmer (2016; see also Urai et al., 2017). Momentary motion energy values within each trial were initially signed with respect to whether the movement of the dots on the given frame favoured the left or right choice. Positive values were attributed to momentary motion energy estimates that favoured rightward motion and negative values were assigned to those favouring leftward motion.

Behavioural Analysis

Across both experiments the data were segmented according to foreperiod duration (800/1200/1600ms). In experiment 1, the data were collapsed across experimental sessions. Response accuracy was defined as the percentage of trials where participants correctly reported the direction of the target. Response time was calculated in milliseconds (ms) relative to the time of the sensory change (i.e. the contrast change or onset of coherent motion). Missed responses were defined as those in which no response was made within the two-second response deadline (marked by the disappearance of the stimulus from the screen). Premature responses were defined as those made prior to or within 150ms of the onset of sensory evidence (i.e. the contrast change or onset of coherent motion). Points earned per trial were defined as the average number of points earned per trial including trials where no points were awarded due to missed, erroneous or premature responses. Each of these behavioural measures was analysed as a function of foreperiod duration using separate one-way repeated measures analyses of variance (ANOVA) followed by pairwise comparisons (paired-samples t-tests). The alpha level following Bonferroni corrections for each set of pairwise comparisons was .017 (.05/3). Accuracy was further examined as a function of response time by computing condition accuracy functions. To this end the data were broken down into equally spaced bins based on RT (8 bins in experiment 1 and 6 bins in experiment 2) and analysed using a two-way repeated measures ANOVAs.

In all of the analyses conducted here, prior investigation was carried out to determine whether the data were suitable for parametric analysis. Firstly, the data were screened for any major violations of normality. In the case of repeated measures ANOVAs, Mauchley's test of sphericity was also carried out to assess the assumption of equal variances of the differences across all combined levels of a single factor. In all instances where the assumption of sphericity was violated, greenhouse-geisser corrections were applied to the degrees of freedom and p-values reported.

The Relationship Between Pre-Evidence Cumulative Motion Energy & Choice

In light of evidence suggesting that the accumulation process may have commenced before the onset of the goal relevant sensory change, we sought to establish, in experiment 2, whether there was a relationship between random fluctuations in the sensory stimulus during

the pre-evidence foreperiod and choice. Specifically, we examined whether fast choices on long foreperiod trials could be attributed to dot motion energy favouring the chosen response by chance. To this end we expressed momentary motion energy values in terms of whether they favoured the chosen or unchosen response for a given trial. Positive values were attributed to motion energy values favouring the chosen response and negative values were assigned to those favouring the unchosen response. We then divided trials into fast (RT<500ms) and slow (RT>500ms) responses and calculated the integral of motion energy during the foreperiod by computing the cumulative sum of the momentary motion energy values expressed with respect to the chosen responses. The cut-off of 500ms between fast and slow responses was chosen in order to capture a sufficiently large subset of trials in which accuracy was very low, as evident in figure 4d. In the case of responses made after the coherent motion onset, the integral was calculated between 100-1600ms. For responses made prior to the onset of coherent motion, the integral was calculated between 100ms and the time of the response. The integral at 1600ms and the slope of the pre-evidence integral (measured as the slope of a line fit to the pre-evidence motion integral) for fast and slow responses was compared against 0 using single sample t-tests. Whereas the absolute value of the integral at 1600 provides a measure of how much evidence cumulatively favoured the chosen response, the slope of the integral is sensitive to the extent to which sensory evidence consistently favoured the chosen response.

EEG acquisition and Preprocessing (Experiment 1 & 2)

The following procedures for acquiring and preprocessing EEG data apply to data from experiment 1 and experiment 2. Continuous EEG data were acquired using an ActiveTwo system (BioSemi, The Netherlands) from 128 scalp electrodes and digitized at 512 Hz. Vertical eye movements were recorded using two vertical electrooculogram (EOG) electrodes placed above and below the left eye. Data were analysed using custom scripts in MATLAB (Mathworks, Natick, MA) and the EEGLAB toolbox (Delorme & Makeig, 2004). Continuous EEG data were low-pass filtered below 35 Hz, high-pass filtered above .05 Hz and detrended. EEG data were then re-referenced offline to the average reference. EEG data were segmented into stimulus- and response-aligned epochs. Stimulus-aligned epochs were extracted from stimulus onset to stimulus offset, which yielded different length windows for each foreperiod condition: 800ms (Epoch: 0-2,800ms), 1200ms (Epoch: 0-3,200ms) or

1600ms (Epoch: 0-3,600ms). Response-aligned epochs were measured from –1000ms pre-response to 600ms post response. The purpose of extending the epoch 600ms post-response was to allow four full fast fourier transform (FFT) windows to be calculated post-response (*see Time Frequency Decomposition below*). All epochs were baseline-corrected relative to the interval of 500-550ms post stimulus onset. This baseline window was chosen, with the aid of visual inspection of the grand-average event-related potential (ERP) waveform, to fall before the onset of the CPP but after the conclusion of evoked potentials elicited by the appearance of the stimulus on the screen and the mouse click that initiated the trial.

Trials were rejected if the bipolar vertical EOG signal (upper minus lower) exceeded an absolute value of 200 μV or if any scalp channel exceeded 100 μV at any time during the stimulus-aligned epoch. To avoid excessive trial loss, channels were interpolated if their individual artefact count exceeded 10% of the total number of trials for a given session. To avoid excessive channel interpolation, a maximum of 10% of the total number of channels were permitted to be interpolated for any given subject's data in a given session. Subjects were excluded from electrophysiological analyses entirely if, following these steps, more than 40% of trials were lost due to blinks and/or EEG artefacts. In order to mitigate the effects of volume conduction across the scalp, single-trial epochs underwent current source density (CSD) transformation (Kayser & Tenke, 2006), a procedure that has been used in order to minimise spatial overlap between functionally distinct EEG components (Kelly & O'Connell, 2013).

Stimulus- and response-aligned epochs were then decomposed into time frequency representations using the Short Time Fourier Transform (STFT) procedure. Each epoch was divided up into a series of overlapping 400ms time segments taken at 50-ms intervals across the epoch. The FFT was computed for each 400-ms segment. For the purpose of data analysis and plotting, each 400ms segment was identified by its median time point. The resulting time-frequency representations had a frequency resolution of 2.5 Hz and as such allowed for the direct measurement of SSVEPs at 20 Hz and 25 Hz, which corresponded to the flicker frequencies used in experiment 1 for the left- and right-tilted gratings respectively. Each 400ms window captured precisely 8 and 10 cycles of the 20 and 25 Hz SSVEPs respectively, thereby minimising spectral leakage. FFT power measurements were converted into amplitude measurements by dividing by half the length of the measurement window in samples.

Analysis of Electrophysiological Signals

The Effect of Foreperiod Duration on Sensory Evidence Representation

Sensory evidence representation was examined by tracing the time course of the SSVEPs elicited by the left- and right-tilted stimuli flickering at 20 Hz and 25 Hz respectively. The SSVEPs were normalised with respect to the average activity in the immediately adjacent frequencies by computing signal to noise ratios (SNR). In order to obtain a direct measurement of the sensory evidence upon which observers based their decisions (i.e. the relative contrast of the overlaid gratings) we then computed the difference SSVEP (d-SSVEP) by subtracting the non-target SSVEP from the target SSVEP. Based on visual inspection of the sample grand-average topography of the d-SSVEP (calculated at approximately 200ms prior to response) we identified a broad cluster of electrodes over the occipital cortex (centered around Oz) at which the d-SSVEP was maximal (figure 2a). At the single-subject level we then measured the SSVEPs by averaging data from the four electrodes within this cluster, at which the subject grand-average d-SSVEP was maximal. Single trial data were rejected from subsequent analyses if the d-SSVEP prior to response exceeded +/- 3 standard deviations from the within subject mean.

To determine whether foreperiod duration had any effect on sensory evidence representation the SSVEPs were examined as a function of foreperiod duration in the following windows: i) 500ms post evidence onset and ii) 200ms prior to response. Trials were excluded from this analysis if subjects responded quicker than 500ms in order to ensure that response-aligned SSVEP measurements did not include time windows prior to the contrast change and that post evidence measurements overlapped minimally with time points post response. This analysis was carried out using two-way repeated measures ANOVAs (factor 1: foreperiod duration; factor 2:stimulus). Effects of foreperiod duration on the d-SSVEP were inferred from the interaction between foreperiod duration and stimulus factors. We predicted that the target and non-target SSVEPs should undergo antithetical changes in SNR after the contrast change giving rise to a main effect of stimulus.

The Effect of Foreperiod Duration on Pre-Evidence Decision Formation

Domain-general decision formation was examined by measuring the dynamics of the CPP. Based on visual inspection of the sample grand-average topography of the pre-response CPP (response aligned window: -150ms : -50ms) we identified a broad cluster of electrodes over the centroparietal cortex (centered around Pz) at which the CPP was maximal (figure 3a; 6a). For each subject, the CPP was measured by averaging data from the four electrodes within this cluster, at which the subject grand-average was maximal. Effector selective decision formation was examined by measuring desynchronisation in the Mu/Beta frequency band (10 - 30 Hz) over the contralateral and ipsilateral response hemispheres (20 Hz and 25 Hz were excluded in order to avoid mixing motor and sensory activity). In order to measure its excursion and minimise the effects of general changes in amplitudes across testing sessions, Mu/Beta was baseline-corrected relative to 550ms post stimulus onset. This window was chosen with the aid of visual inspection of the grand average waveforms to ensure that the motor activity relating to the mouse button click at the beginning of each trial in experiment 1 was fully resolved. The Mu/Beta lateralisation index was computed by subtracting ipsilateral from contralateral Mu/Beta. Based on inspecting the grand average topography of pre response Mu/Beta lateralisation (response aligned window: -150 : -50ms), a cluster of electrodes was identified in each hemisphere over premotor regions of the scalp (C3 left hemisphere; C4 right hemisphere; figure 3h; 6h) that exhibited the largest lateralisation index. For each subject Mu/Beta was measured by averaging data from the four electrodes within this cluster, at which the subject grand-average was maximal. Single trial CPP and Mu/Beta data were rejected from subsequent analyses if the values estimated prior to response exceeded +/- 3 standard deviations from the within-subject mean.

Inspection of grand average CPP and lateralised Mu/Beta waveforms indicated that there was substantial build-up of both signals during the foreperiod on longer foreperiod trials suggesting that decision formation may have commenced prematurely on those trials. To investigate this phenomenon we initially pooled data from correct, incorrect, missed and premature response trials and separated the data according to foreperiod duration. We examined the amount of pre-evidence build-up in each signal by calculating the average amplitude of each signal within a 100ms window centered on the time of the sensory change. We also measured the rate of build-up during the foreperiod as the slope of a line fit to the waveforms between -250ms and 50ms relative to the sensory change. In order to determine whether there was significant pre-evidence decision related activity, the slope and amplitude

of CPP and Mu/Beta lateralisation at each level of foreperiod duration were compared against 0 using one-sample t-tests. We repeated this analysis for long foreperiod trials using only data from experimental session 1. Next, we divided the data into equally sized bins based on RT (6 bins in experiment 1; 3 bins in experiment 2) and examined the relationship between RT and pre-evidence decision related activity (amplitude and slope as defined above) using two-way repeated measures ANOVAs including foreperiod duration and RT bin as factors.

The Effect of Foreperiod Duration on Pre-Response Decision Related Activity

Decision formation dynamics were further investigated relative to response. The build-up rate of the CPP was measured by calculating the slope of a line fit to the waveform in the window -500ms to -200ms relative to response. The build-up rate of Mu/Beta lateralisation was examined by measuring the slope of a line fit to the lateralisation indices contained within the window -500ms to -200ms relative to response. We measured the amplitude of the CPP and magnitude of Mu/Beta desynchronisation and lateralisation prior to response in the window -150:-50ms. Contralateral and ipsilateral Mu/Beta were examined separately so as to evaluate preparation of both the chosen and unchosen responses. As before, the data were separated according to foreperiod duration and RT and analysed using two-way repeated measures ANOVAs.

Examining the Relationship Between Motion Energy & Premature Sensory Evidence Accumulation (Experiment 2)

In experiment 2 we sought to establish a relationship between cumulative motion energy during the foreperiod and pre-evidence CPP build-up. To this end, focusing exclusively on long foreperiod (1600ms) trials, we separated trials into two equally sized bins according to the steepness (slope) of the dot motion integral. This yielded a subset of trials in which the cumulative motion energy during the foreperiod strongly and consistently favoured either leftward or rightward motion by change and a subset of trials where the cumulative motion energy fluctuated closer to 0, favouring neither direction consistently. Single sample t-tests were first carried out, comparing the pre-evidence CPP amplitude and slope against 0, in order to determine whether there was any significant build-up of the CPP during the foreperiod on steep or shallow integral trials. The amplitude of the CPP was

measured in a 100ms window centred on coherent motion onset while the slope was measured in a broad window of 600-1600ms. To distinguish between the CPP on steep and shallow pre-evidence integral trials, additional paired samples t-tests were carried out comparing the amplitude and slope of the CPP between shallow and steep integral trials.

Examining The Effect of Previous Trial Foreperiod on Accumulation Onset Timing

In order to examine the sequential effect of foreperiod on evidence accumulation, the single trial data from experiment were subsequently further subdivided according to the duration of the foreperiod on the previous trial ($foreperiod_{n-1}$) and the duration of the foreperiod on the current trial ($foreperiod_n$). At the behavioural level accuracy and response times were examined using two-way repeated measures ANOVAs including $foreperiod_{n-1}$ and $foreperiod_n$ as separate independent factors. Missed response rate and premature response rate were examined using one-way repeated measures ANOVAs including only $foreperiod_{n-1}$ as an independent factor. The exclusion of $foreperiod_n$ as a factor in these analyses was necessary because missed responses were made almost exclusively on short $foreperiod_n$ trials while premature responses were made almost exclusively on long $foreperiod_{n-1}$ trials. At the neural level we also further analysed the CPP as a function of $foreperiod_{n-1}$ in order to determine whether the amount or rate of build-up, in either the pre-evidence onset or pre-response windows, was modulated by the duration of the previous trial foreperiod.

Examining Target Selection Signals Under Temporal Uncertainty

The target selection process was examined by measuring the bilateral N2 which is thought to play a role in triggering the onset of evidence accumulation (Loughnane et al., 2016). Based on visual inspection of the sample grand-average topography of the ERP in the window 200ms to 300ms post sensory change we identified a broad cluster of electrodes over the bilateral occipital cortex at which negative going activity was maximal (supplementary figure 2a; 3a). For each subject, the N2 was measured by averaging data from 3 electrodes centered on P7 and 3 electrodes centered on P8 at which the negativity was greatest. To establish whether target selection mechanisms were reliably elicited by the sensory change in this task we compared the amplitude of the N2, measured in the window 200ms-300ms post sensory change, against 0 using separate single sample t-tests for each level of foreperiod.

References

Andelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America, 2*(2), 284-299.

Bausenhart, K. M., Rolke, B., Seibold, V. C., & Ulrich, R. (2010). Temporal preparation influences the dynamics of information processing: Evidence for early onset of information accumulation. *Vision Research, 50*(11), 1025-1034.

Bode, S., Sewell, D.K., Lilburn, S.D., Forte, J.D., Smith, P.L., & Stahl, J. (2012). Predicting perceptual decision biases from early brain activity. *The Journal of Neuroscience, 32*(36), 12488-98.

Braun, A., Urai, A. E., & Donner, T. H. (2018). Adaptive history biases result from confidence-weighted accumulation of past choices. *Journal of Neuroscience, 38*(10), 2418-2429

Capizzi, M., Correa, A., & Sanabria, D. (2013). Temporal orienting of attention is interfered by concurrent working memory updating. *Neuropsychologia, 51*(2), 326-339.

Capizzi, M., Correa, Á., Wojtowicz, A., & Rafal, R.D. (2015). Foreperiod priming in temporal preparation: Testing current models of sequential effects. *Cognition, 134*, 39-49.

Capizzi, M., Sanabria, D., & Correa, Á. (2012). Dissociating controlled from automatic processing in temporal preparation. *Cognition, 123*(2), 293-302.

Cisek, P., Puskas, G. A., & El-Murr, S. (2009). Decisions in changing conditions: the urgency-gating model. *Journal of Neuroscience, 29*(37), 11560-11571.

Correa, Á., Lupiáñez, J., Madrid, E., & Tudela, P. (2006). Temporal attention enhances early visual processing: A review and new evidence from event-related potentials. *Brain Research, 1076*(1),116-128.

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

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9-21.

Diaz, J. A., Queirazza, F., & Philiastides, M. G. (2017). Perceptual learning alters post-sensory processing in human decision making. *Nature Human Behaviour, 1* (35): 1-9.

Di Russo, F., Pitzalis, S., Aprile, T., Spitoni, G., Patria, F., Stella, A., Spinelli, D., & Hillyard, S. A. (2007). Spatiotemporal analysis of the cortical sources of the steady-state visual evoked potential. *Human Brain Mapping, 28,* 323-34.

Di. Russo, F. D., Spinelli, D., & Morrone, M. (2001). Automatic gain control contrast mechanisms are modulated by attention in humans: evidence from visual evoked potentials. *Vision Research, 41, 2*435-2447.

Di Russo, F., Teder-Sälejärvi, Wolfgang, A., & Hillyard, S. A. (2003). Steady-State VEP and Attentional Visual Processing. In A., Zani, A. M. Proverbio., & M. I. Posner (Eds.), *The Cognitive Electrophysiology of Mind and Brain* (259-274). Academic Press

Donner, T. H., Siegel, M., Fries, P., & Engel, A. K. (2009). Build-up of choice-predictive activity in human motor cortex during perceptual decision making. *Current Biology, 19*(18), 1581-1585.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience, 30,* 535-74.

Hanks, T., Kiani, R., & Shadlen, M. N. (2014). A neural mechanism of speed-accuracy tradeoff in macaque area LIP. *eLife, 3,* e02260.

Jazayeri, M., & Shadlen, M. (2010).Temporal context calibrates interval timing. *Nature Neuroscience, 13*(8), 1020-1026.

Jazayeri, M., & Shadlen, M. N. (2015). A Neural Mechanism for Sensing and Reproducing a Time Interval. *Current Biology, 25*(20), 2599–2609.

Jepma, M., Wagenmakers, E.-J., & Nieuwenhuis, S. (2012). Temporal expectation and information processing: A model-based analysis. *Cognition, 122*(3), 426-41.

Jones, M., Curran, T., Mozer, M. C., & Wilder, M. H. (2013). Sequential effects in response time reveal learning mechanisms and event representations. *Psychological Review, 120*(3), 628-666.

Kayser, J., & Tenke, C. E. (2006). Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: I. Evaluation with auditory oddball tasks. *Clinical Neurophysiology, 117*(2), 348-368.

Kelly, S. P., & O'Connell, R. G. (2015). The neural processes underlying perceptual decision making in humans: recent progress and future directions. *Journal of Physiology-Paris, 109*(1-3), 27-37.

Kelly, S. P., & O'Connell, R. G. (2013). Internal and external influences on the rate of sensory evidence accumulation in the human brain. *Journal of Neuroscience, 33*(50), 19434-19441.

Kim, Y. J., Grabowecky, M., Paller, K. A., Muthu, K., & Suzuki, S. (2007). Attention induces synchronization-based response gain in steady-state visual evoked potentials. *Nature Neuroscience, 10*(1), 117-125.

Kloosterman, N. A., de Gee, J. W., Werkle-Bergner, M., Lindenberger, U., Garrett, D. D., & Fahrenfort, J. J. (2018). Humans strategically shift decision bias by flexibly adjusting sensory evidence accumulation. *bioRxiv*, 229989.

Los, S. A., & Heslenfeld, D. J. (2005). Intentional and unintentional contributions to nonspecific preparation: electrophysiological evidence. *Journal of Experimental Psychology: General, 134*(1), 52-72.

Los, S. A., Kruijne, W., & Meeter, M. (2017). Hazard versus history: Temporal preparation is driven by past experience. *Journal of Experimental Psychology: Human Perception and Performance, 43*(1), 78-88.

Los, S. A., & Van Den Heuvel, C. E. (2001). Intentional and unintentional contributions to nonspecific preparation during reaction time foreperiods. *Journal of Experimental Psychology: Human Perception and Performance, 27*(2), 370-386.

Luck, S. J. (2012). Electrophysiological correlates of the focusing of attention within complex visual scenes: N2pc and related ERP component. In S. J. Luck and E. S. Kappenman (Eds.), *The Oxford Handbook of Event-related Potential Components* (pp. 329-360). New York: Oxford University Press.

Loughnane, G. M., Newman, D. P., Bellgrove, M. A., Lalor, E. C., Kelly, S. P., & O'Connell, R. G. (2016). Target selection signals influence perceptual decisions by modulating the onset and rate of evidence accumulation. *Current Biology, 26*(4), 496-502.

Loughnane, G. M., Newman, D. P., Tamang, S., Kelly, S. P., & O'Connell, R. G. (2018). Antagonistic interactions between microsaccades and evidence accumulation processes during decision formation. *Journal of Neuroscience, 38*(9), 2163-2176.

McGovern, D. P., Hayes, A., Kelly, S. P., & O'Connell, R. G. (2018). Reconciling age-related changes in behavioural and neural indices of human perceptual decision-making. *Nature Human Behaviour*, 1.

McGovern, D.P., Webb, B.S., & Peirce, J.W. (2012). Transfer of perceptual learning between different visual tasks. *Journal of Vision, 12 11*, 4.

Mento, G. (2013). The passive CNV: carving out the contribution of task-related processes to expectancy. *Frontiers in Human Neuroscience, 7,* 827.

Mento, G. (2017). The role of the P3 and CNV components in voluntary and automatic temporal orienting: A high spatial-resolution ERP study. *Neuropsychologia, 107,* 31-40.

Murphy, P. R., Boonstra, E., & Nieuwenhuis, S. (2016). Global gain modulation generates time-dependent urgency during perceptual choice in humans. *Nature Communications, 7,* 13526.

Murphy, P. R., Robertson, I. H., Harty, S., & O'Connell R. G. (2015). Neural evidence accumulation persists after choice to inform metacognitive judgments. *eLife,* e11946.

Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychological Bulletin, 89*(1),133-162.

Nobre, A. C., Correa, A., & Coull, J. T. (2007). The hazards of time. *Current Opinion in Neurobiology, 17*(4), 465-470.

Nobre, A. C., & van Ede, F. (2018). Anticipated moments: temporal structure in attention. *Nature Reviews Neuroscience, 19*(1), 34-48.

Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottereau, B. R., & Rossion, B. (2015). The steady-state visual evoked potential in vision research: A review. *Journal of Vision, 15*(6), 4.

O'Connell, R. G., Dockree, P. M., & Kelly, S. P. (2012). A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nature Neuroscience, 15*(12), 1729-1735.

O'Connell, R. G., Shadlen, M. N., Wong-Lin, K. F., & Kelly, S. K. (2018). Bridging neural and computational viewpoints on perceptual decision-making. *Trends in Neurosciences, 41*(11), 838-852.

Pouget, P., Logan, G.D., Palmeri, T.J., Boucher, L., Paré, M., & Schall, J. D. (2011). Neural basis of adaptive response time adjustment during saccade countermanding. *Journal of Neuroscience, 31*(35), 12604-12612.

Purcell, B. A., Heitz, R. P., Cohen, J. Y., Schall, J. D., Logan, G. D., & Palmeri, T. J. (2010). Neurally constrained modeling of perceptual decision making. *Psychological Review, 117*(4), 1113-1148.

Purcell, B. A., Schall, J. D., Logan, .G. D., & Palmeri, T. J. (2012). From Salience to Saccades: Multiple-Alternative Gated Stochastic Accumulator Model of Visual Search. *Journal of Neuroscience, 32*(10), 3433-3446.

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review, 111*(2), 333-367.

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: current issues and history. *Trends in Cognitive Sciences, 20*(4), 260-281.

Rolke, B. (2008). Temporal preparation facilitates perceptual identification of letters. *Perception & Psychophysics, 70,*1305-1313.

Rolke, B., & Hofmann, P. (2007). Temporal uncertainty degrades perceptual processing. *Psychonomic Bulletin & Review, 14*(3), 522-526.

Schall, J. D., Purcell, B. A., Heitz, R. P., Logan, G. D, & Palmeri, T. J. (2011) Neural mechanisms of saccade target selection: gated accumulator model of the visual–motor cascade. *European Journal of Neuroscience 33*(11), 1991-2002.

Seibold, V. C., Bausenhart, K. M., Rolke, B., & Ulrich, R. (2011). Does temporal preparation increase the rate of sensory information accumulation?. *Acta Psychologica, 137*(1), 56-64.

Shadlen, M. N., & Kiani, R. (2013). Decision making as a window on cognition. *Neuron, 80*(3), 791-806.

Shadlen, M. N., & Newsome, W. T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *Journal of Neuroscience, 18*(10), 3870-3896.

Siegel, M., Engel, A. K., & Donner, T. H. (2011). Cortical network dynamics of perceptual decision-making in the human brain. *Frontiers in Human Neuroscience, 5*, 21.

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences, 27*(3), 161-168.

Steinborn, M. B., & Langner, R. (2012). Arousal modulates temporal preparation under increased time uncertainty: Evidence from higher-order sequential foreperiod effects. *Acta Psychologica, 139*(1), 65-76.

Steinemann, N. A., O'Connell, R. G., & Kelly, S. P. (2018). Decisions are expedited through multiple neural adjustments spanning the sensorimotor hierarchy. *Nature Communications,9*(1), 3627.

Talluri, B. C., Urai, A. E., Tsetsos, K., Usher, M., & Donner, T. H. (2018). Confirmation bias through selective over-weighting of choice-consistent evidence. *Current Biology, 28*(19), 3128-3135.

Teichert, T., Grinband, J., & Ferrera, V. (2016). The importance of decision onset. *Journal of Neurophysiology, 115*(2), 643-661.

Thura, D., Beauregard-Racine, J., Fradet, C. W., & Cisek, P. (2012). Decision making by urgency gating: theory and experimental support. *Journal of Neurophysiology, 108*(11), 2912-2930.

Thura, D., & Cisek, P. (2016). Modulation of premotor and primary motor cortical activity during volitional adjustments of speed-accuracy trade-offs. *Journal of Neuroscience, 36*(3), 938-956.

Twomey, D. M., Kelly, S. P., & O'Connell, R. G. (2016). Abstract and effector-selective decision signals exhibit qualitatively distinct dynamics before delayed perceptual reports. *Journal of Neuroscience, 36*(28), 7346-7352.

Twomey, D. M., Murphy, P. R., Kelly, S. P., & O'Connell, R. G. (2015). The classic P300 encodes a build-to-threshold decision variable. *European Journal of Neuroscience, 42*(1), 1636-1643.

Urai, A. E., Braun, A. & Donner, T. H. (2017). Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. *Nature Communications, 8,*14637.

Urai, A., E., de Gee, J. W., Tsetsos, K. & Donner, T. B. (2019). Choice history biases subsequent evidence accumulation. *BioRxiv*, 251595. https://doi.org/10.1101/251595

Urai, A., E., & Wimmer, K. (2016). Spatiotemporal motion energy filtering: a Matlab implementation. *Zenodo,* http://doi.org/10.5281/zenodo.594505.

Vallesi, A., McIntosh, A. R., & Stuss, D. T. (2009). Temporal preparation in aging: A functional MRI study. *Neuropsychologia, 47*(13), 2876-2881.

Vallesi, A., Shallice, T., & Walsh, V. (2007). Role of the prefrontal cortex in the foreperiod effect: TMS evidence for dual mechanisms in temporal preparation. *Cerebral Cortex, 17*(2), 466-474.

van der Lubbe, R. H., Los, S. A., Jaskowski, P., & Verleger, R. (2004). Being prepared on time: on the importance of the previous foreperiod to current preparation, as reflected in speed, force and preparation-related brain potentials. *Acta psychologica, 116*(3), 245-62.

Vialatte, F. B., Maurice, M., Dauwels, J., & Cichocki, A. (2010). Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. *Progress in Neurobiology, 90*(4), 418-438.

Wilder, M. H., Jones, M., Ahmed, A. A., Curran, T., & Mozer, M. C. (2013). The persistent impact of incidental experience. *Psychonomic Bulletin & Review, 20*(6), 1221-1231.

Wilder, M., Jones, M., & Mozer, M. C. (2010). Sequential effects reflect parallel learning of multiple environmental regularities. In Y. Bengio, D. Schuurmans, J. Lafferty, C.K.I.

Williams, & A. Culotta (Eds.), *Advances in Neural Information Processing Systems 22* (pp. 2053-2061). La Jolla, CA: NIPS Foundation.

Yu, A. & D. Cohen, J. (2008). Sequential effects: Superstition or rational behavior? *Advances in Neural Information Processing Systems, 21,* 1873-1880.

Zhang, S., Huang, H. C, & Yu, A. J. (2014). Sequential effects: A Bayesian analysis of prior bias on reaction time and behavioral choice. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 36.

Zohary, E., Shadlen, M. N., & Newsome, W. T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature, 370*(6485), 140-143.