



Prior expectations modulate unconscious evidence accumulation



Leonardo S. Barbosa^{a,b,*,1}, Alexandra Vlassova^{a,1}, Sid Kouider^a

^a Brain and Consciousness Group (ENS, EHESS, CNRS), École Normale Supérieure – PSL Research University, Paris, France

^b École Doctorale Cerveau Cognition Comportement, Université Pierre et Marie Curie, Paris, France

A B S T R A C T

Unconscious processes have been shown to affect both perception and behaviour. However, the flexibility of such processes remains unknown. Here we investigate whether unconscious decisional processes can adapt to the utility of sensory information. To this end, we had participants gradually accumulate information from noisy motion stimuli, until a decision was reached. We titrated conscious awareness of these stimuli by simultaneously presenting a dynamic dichoptic mask. Crucially, we manipulated the likelihood that the suppressed portion of each presentation would contain useful information. Our results show that the statistics of the environment can be used to modulate unconscious evidence accumulation, resulting in faster choices. Furthermore, computational modelling revealed that this modulation is due to a change in the quality of unconscious evidence accumulation, rather than a conscious change in strategy. Together, these results indicate that unconscious decisional mechanisms are capable of optimising performance by flexibly adapting to the statistical environment.

1. Introduction

While for many years the existence of unconscious processing was the subject of intense debate (Eriksen, 1960; Holender, 1986; Merikle & Daneman, 1998), today it is generally accepted that unconscious processes can affect both our perception and behaviour (Kouider & Dehaene, 2007; Van den Bussche, Van den Noortgate, & Reynvoet, 2009). Furthermore, recent evidence suggests that information can be accumulated in the absence of conscious awareness and be used to improve decision-making (Vlassova, Donkin, & Pearson, 2014). However, the nature of these unconscious decisional mechanisms remains unclear. In particular, the extent to which these mechanisms can engender flexible adaptive behaviour in the absence of awareness remains unknown.

Our ability to flexibly adapt to the current state of the world allows us to modify our behaviour to optimise performance. Previous studies have demonstrated that prior experiences and expectations can influence subsequently made choices (e.g. de Lange, Rahnev, Donner, & Lau, 2013; Mulder, Wagenmakers, Ratcliff, Boekel, & Forstmann, 2012; for a review see Summerfield & de Lange, 2014). Such context-driven effects have also been shown to occur in the absence of awareness. For example, the magnitude of unconscious perceptual priming effects has been found to depend on the likelihood that stimuli masked from visual awareness contain useful information (Bodner & Masson, 2001, 2003; Bodner, Masson, & Richard, 2006; Jaśkowski, Skalska, & Verleger, 2003; Kinoshita, Mozer, & Forster, 2011). However, it remains unclear whether the statistics of the environment can exert a similar influence on unconscious decisional processing. If the environmental statistics dictate that the unconscious content is likely to contain valuable decision-relevant information, are we more likely to utilise this information? Likewise, when the unconscious content has no utility, are we able to ignore or discount it?

* Corresponding author at: Brain and Consciousness Group (ENS, EHESS, CNRS), École Normale Supérieure – PSL Research University, 123 rue d'Ulm, 75005, Paris, France.

E-mail address: leonardo.barbosa@usp.br (L.S. Barbosa).

¹ Authors contributed equally to this paper.

Here, we address these questions by investigating whether the statistics of the environment in which decisions are made can bias the unconscious accumulation of evidence. Using the dichoptic suppression paradigm developed by [Vlassova et al. \(2014\)](#), we were able to titrate perceptual awareness of random-dot kinematograms. This paradigm allows us to investigate how the accumulation of evidence in the absence of visual awareness can affect our perceptual decision-making. Subjects were asked to make decisions about the direction of motion in these partially suppressed stimuli: the first 400 ms of each trial was suppressed from conscious awareness and was immediately followed by an additional 300 ms of unmasked visible motion. We induced decisional expectations by presenting the stimuli in two contexts: in one, the motion direction in the suppressed portion of the trial was more likely to be congruent with the motion direction of the visible portion. In a second context, the suppressed motion stimulus was more likely to contain purely random motion.

We modelled participants' performance using a drift diffusion model ([Ratcliff, 1978](#)), according to which decisions are made by the gradual accumulation of evidence over time, until such a point when the amount of information in favour of a particular choice has reached a decision threshold and a response is obtained. This allowed us to disentangle the different mechanisms driving the decisional process, such as differentiating between changes in performance driven by changes in the rate of evidence accumulation or changes driven by strategic shifts of the decision threshold. Our results indicate that the statistical context in which decisions are made can modulate the rate at which unconscious information is accumulated. Our results therefore highlight that our perceptual decision-making processes can flexibly adapt the rate of unconscious accumulation in line with the potential utility of incoming sensory information.

2. Materials and methods

2.1. Participants

19 participants (7 male; 18–33 years of age) were prescreened for inclusion in this experiment. 5 participants were excluded because they could not identify the catch trials above 90% correct, while an additional 3 participants were excluded because they reported seeing the suppressed grey dots (suppression breaks) on more than 20% of the trials (see Section 2.4). A further 3 participants were removed as they were responding before the stimulus presentation had finished and a technical issue prevented us from recording these responses accurately. Finally, one participant was excluded for being unable to perform the task (performance on the motion discrimination task remained at chance throughout 7 days of training). This left us with 7 participants (2 male; 18–33 years of age), all of whom were able to accurately report the catch trials ($M = 96.65\%$, $SD = 1.54$) and who had minimal suppression breaks ($M = 5.78\%$, $SD = 5.40$). All participants had normal or corrected-to-normal vision, and provided informed written consent. Participants were compensated for their time at a rate of 10€ per hour, for 7 total hours completed over 7 days of testing.

2.2. Apparatus

Participants were seated in a dark room on a height-adjustable chair at a distance of 43 cm from a BenQ XL2420-B LCD screen, with a resolution of 1657×932 and a refresh rate of 75 Hz. Participants' heads were stabilised using a chin and headrest housing a mirror stereoscope apparatus adjusted for each subject. This stereoscope uses circular mirrors to display images presented on the screen separately to each eye, which perceptually overlap one another to form a single image when viewed binocularly. Stimuli were presented using MATLAB (R2012b) Psychophysics Toolbox 3 ([Brainard, 1997](#); [Pelli, 1997](#)), on a HP 2400 Workstation machine running Windows 7 Professional (SP1).

2.3. Stimuli

The motion stimuli used in this study were random dot kinematogram (RDK) displays, which are commonly used in research in perceptual decision-making (for a review see: [Smith & Ratcliff, 2004](#)). The RDK stimulus parameters were set to those previously reported in [Vlassova et al. \(2014\)](#), as follows: 100 grey dots (21 cd/m^2) were displayed within an invisible 8.2° circular aperture, with a central 0.7° fixation point. Participants were instructed to fixate on this point throughout the experimental sessions in order to facilitate fusion. Each dot was a 1×1 pixel square and moved at a speed of 6.1° per second. Dots were presented on a black background at a density of 1.9 dots/deg^2 . In order to conserve dot density, dots that moved along a trajectory that would place them outside of the circular aperture were wrapped around to the opposite side and maintained their original motion direction. Three uncorrelated random sequences of dot movement were generated, and frames were interleaved so that each coherent dot was only correlated with dots three frames forwards or backwards and not the subsequent frame ([Roitman & Shadlen, 2002](#); [Shadlen & Newsome, 2001](#)). That is, coherent dots were first displayed for one frame, and then three frames later they were displaced by 0.25° in the direction of the overall coherent motion, while remaining dots were replotted randomly. In this way, coherently moving dots could not be individually tracked over time.

During suppressed trials, RDK dots were suppressed from conscious awareness using a dynamic dichoptic mask configuration that has been previously shown to effectively suppress dot-motion stimuli from conscious awareness for durations up to 500 ms ([Vlassova et al., 2014](#)). The mask consisted of 350 green dots (58.5 cd/m^2), each a 1×1 pixel. Mask dots were displayed within an invisible 9.8° circular aperture, around a central 0.7° fixation point, with an average dot density of 3.3 dots/deg^2 . During the visible trials, RDK dots were rendered visible by changing the mask to 25 green dots, resulting in an average dot density of 0.24 dots/deg^2 . The dots moved concentrically around the central fixation point (clockwise) at a rate of 1.67 revolutions per second, and were presented on every third frame only. The ocular dominance of each participant was assessed using a simple pointing test, and in order to facilitate suppression of the RDK stimulus, the mask stimulus was presented to each participant's dominant eye.

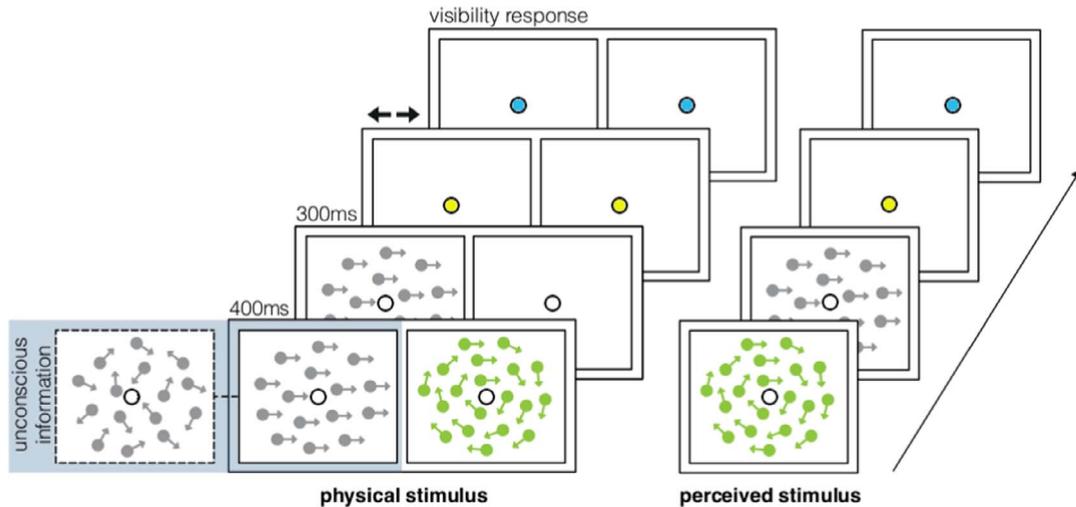


Fig. 1. Task design. The left panel shows the timeline of the stimulus presentation, while the right panel illustrates a subject's perception of the presented stimuli. The mask consisted of green dots spinning concentrically around the fixation point (not drawn to scale), which was presented to the participant's dominant eye. A Random Dot Kinematogram (RDK) was concurrently presented to the other eye, and consisted of 100 moving grey dots. The blue-shaded region highlights the first 400 ms of the trial. During this initial 400 ms, on half the trials the RDK was rendered unconscious by the mask (suppressed trials). In the other half, mask effectiveness was weakened by reducing the number of green dots, rendering the grey dots visible (visible trials). Additionally, during this first 400 ms, on half the trials (random trials) each grey dot moved in a random direction, while in the other half (coherent trials) a fraction (40%) of the grey dots moved coherently in one of two directions, left or right. During the following 300 ms, the mask disappeared and only the RDK was presented (always fully visible). In the coherent condition, the coherent fraction of the grey dots continued to move in the same direction as in the first 400 ms of the trial, while in the random condition a fraction (also 40%) started to move coherently in one of the two directions. At stimulus offset, the fixation point changed color and prompted participants to report the direction of motion as quickly and accurately as possible. Participants had to choose the direction of the moving grey dots. After their response, the fixation point changed color again. Participants were then asked to indicate whether they saw any grey dots during the first half of the stimulus presentation (i.e. concurrently with any green dots). Importantly, during half of the experiment (congruent context), 80% of the trials were coherent, while in the other half (incongruent context), only 20% of the trials were coherent. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.4. Procedure

The RDK stimulus was presented for 700 ms, with the mask stimulus presented concurrently for the first 400 ms only (see Fig. 1A for a timeline). In a congruent context, on 80% of the trials the RDK stimulus consisted of 40% coherent motion. On the remaining 20% of trials, all of the RDK dots moved randomly while the mask stimulus was concurrently presented (first 400 ms), and then 40% of the dots moved coherently following mask offset (for the last 300 ms of visible presentation). This was reversed in the incongruent context condition, with 80% of trials having random RDK motion while the mask stimulus was presented (first 400 ms), and the remaining 20% of trials having coherent motion for the entire stimulus presentation duration. To allow for the learning of the context and the build-up of expectations, in both the conditions half of the trials were rendered visible by eliminating mask effectiveness: half the trials contained the normal number of green dots in the mask (350 green dots, suppressed trials), while the other half contained a heavily reduced number of mask dots in the first half of the stimulus presentation (25 green dots, visible trials).

On each trial, participants were asked to decide the overall motion direction (left or right) of the RDK stimulus, and report their decision following RDK offset as quickly and accurately as possible using the arrow keys on a standard AZERTY keyboard. To help cue the participants to respond, the previously empty fixation point was filled with yellow following stimulus offset. Once participants executed their motion discrimination response, the fixation point was filled with blue and participants were then asked to press the 'A' key if they saw any grey dots while the green mask dots were also on the screen, and the 'Z' key if they did not. All subsequently reported analyses for the suppressed condition include only trials in which participants indicated that the RDK stimulus was successfully suppressed (95% of all suppressed trials). Conversely, for the visible condition we include only trials where participants reported seeing the grey dots (98% of all visible trials). To obtain an objective measure of participants' visibility reports we included two types of catch trials: during the first 400 ms, either the grey dots were presented to both eyes and the number of mask dots was reduced by 50% (simulating a break in suppression), or no grey dots were presented and only the mask dots were presented (simulating successful suppression). Both types of catch trials were presented 6 times in each block.

Before starting the main experiment, participants completed a short training session to familiarise them with the paradigm. Participants first completed one block of 24 trials (16 normal trials and 8 catch trials) where they performed a simple left/right motion discrimination task. During this training block participants received auditory feedback for correct and incorrect responses on a trial by trial basis. Participants repeated similar blocks until their performance identifying the left/right motion was above 55% correct (average 1 block per participant). Next they completed one block of 32 trials (16 normal trials and 16 catch trials) where they were asked to perform the full task and therefore also report stimulus visibility after making the left/right motion direction discrimination. No more feedback was given on a trial by trial basis until the end of the session. Participants repeated similar blocks until they were correctly identifying both catch trials above 90% correct

(average 4 ± 2 blocks per participant). Participants that could not correctly identify the catch trials did not return for the main experiment. Following the training session, participants completed 6 sessions of the main experiment. On each subsequent day, participants completed a single training block followed by 6 experimental blocks consisting of 112 trials (100 normal trials and 12 catch trials), for a total of 42 experimental blocks. Participants were encouraged to take short breaks in between blocks to prevent fatigue.

2.5. Drift diffusion model

We next wanted to determine whether changes in task performance were being driven by changes in unconscious decisional-information processing, or whether changes in RDK performance simply reflected a change in decision criteria. To this end, we fitted a drift diffusion model (DDM) to the responses of each participant. The drift diffusion model assumes that noisy information is sampled continuously from a stimulus, until the amount of evidence accumulated reaches a response boundary and a decision is made (Ratcliff & Rouder, 1998; Ratcliff & McKoon, 2008). We used the fast-dm (Voss & Voss, 2007) implementation of the DDM to fit a Wiener distribution (Ratcliff, 1978) to the behavioural data. This implementation has four parameters: the drift rate represents the quality of evidence accumulation, the response boundary represents how much evidence is accumulated before a response is given, the bias represents a shift in the starting point of the accumulator towards one of the two response boundaries, and the non-decision time incorporates factors unrelated to the decision process itself, such as the encoding of evidence and response execution.

Here for each subject, we fitted one drift and one boundary value for each context and coherence condition, as well as a single constant non-decision component. Additionally, for each subject we fitted the variance across trials of the non-decision time and the variance of the drift rate for each context. Since we used correct and incorrect responses as response boundaries, and since participants had no information about the upcoming direction prior to the stimulus presentation (half the trials contained leftward motion and the other half rightward motion), in this analysis we fixed the bias parameter in the model to 0.5 (no bias). Since the model was very sensitive to outliers in the reaction time distribution, we used a strict criterion and removed all the trials faster than 200 ms and slower than 1.5 s (9% of all trials). The model of each subject was fitted using a gradient descent algorithm (Voss & Voss, 2007) combined with a Chi Square test to decide the convergence (Ratcliff & McKoon, 2008). This fitting produced a good fit for each subject's data, as revealed by the lack of significant statistical difference between surrogated simulation data and the real data (Kolmogorov-Smirnov p -value > 0.05 ; see Fig. 3A). Conventional ANOVA analyses were performed in the drift and boundary parameters across subjects.

3. Results

When discriminating the direction of the random-dot kinematograms (RDKs), participants were above chance for both suppressed (mean accuracy = 87%, $p < 0.001$) and visible (mean accuracy = 85%, $p < 0.001$) trials. This shows that participants correctly learned how to perform the task, and the information present during the second part of the trial was sufficient for them to identify the RDK direction, even when the first part of the trial was suppressed. To investigate the effect of congruence and coherence of the first part of the trial over performance (see Section 2), we next conducted an ANOVA with congruence and coherence as main factors for both RDK accuracy and reaction time (RT). The analyses of RDK accuracy failed to show an interaction between congruence and coherence for both suppressed and visible trials (all $p > 0.1$). Importantly, the analyses of RTs showed a significant interaction between congruence and coherence for suppressed trials [$F(1,6) = 18.34$, $p < 0.01$]. Post-hoc analyses revealed that this effect was driven by the trials in the congruent context: while there was no significant difference between random and coherent trials during the incongruent context ($p > 0.1$), participants were slower to respond to random trials in comparison to coherent trials ($\Delta RT = 22$ ms, $p < 0.05$) during the congruent context (see Fig. 2A).

This result indicates that unconscious perceptual decision-making can be influenced by expectations induced by the different contexts. Here, such expectations were created by the statistics of the environment. That is, when trials were expected to start more often with relevant (coherent) information, participants were slower to identify the direction of the RDK when the trial started with irrelevant (random) motion. Importantly, this occurred despite the first part of the stimulus presentation being suppressed from perceptual awareness. While the same analyses for the visible trials did not reveal a significant interaction between congruence and coherence ($p > 0.1$), a significant main effect of coherence was found [$F(1,6) = 18.34$, $p < 0.01$; see Fig. 2B]. This result shows that participants used the coherent information to improve the conscious perceptual decision-making, and the lack of interaction might indicate that the task was too easy for participants (mean accuracy across all conditions was 86%). This could be a consequence of the high coherence chosen in the experiment in order to allow participants to easily differentiate random from coherent visible trials, and therefore learn the different contexts. It is possible that the effect of the congruence may have its strongest impact on performance when the task is difficult and the sensory information is ambiguous. While this might have prevented an effect of congruence over RT for visible trials, the significant interaction between congruence and coherence for invisible trials argues for the fact that participants indeed learned the different contexts.

3.1. Drift diffusion model

While our results show that unconscious decisional processing can be modulated by context-induced expectations, it is unclear which aspect of the decisional process is driving such changes. For example, our results may reflect a change in the quality of unconscious evidence accumulation, or they could also stem from changes in conscious strategies. During the congruent context, participants may simply adapt their response strategy and thereby accumulate less evidence before making a decision. This would result in faster reaction times for coherent trials, since during the random trials no evidence was accumulated during the first part of the trial. Another possibility is that the quality of unconscious information used in the decision process is modified by expectations. In this case, the change in RT would result from an improved quality in the unconscious accumulation of evidence provided by the suppressed RDK.

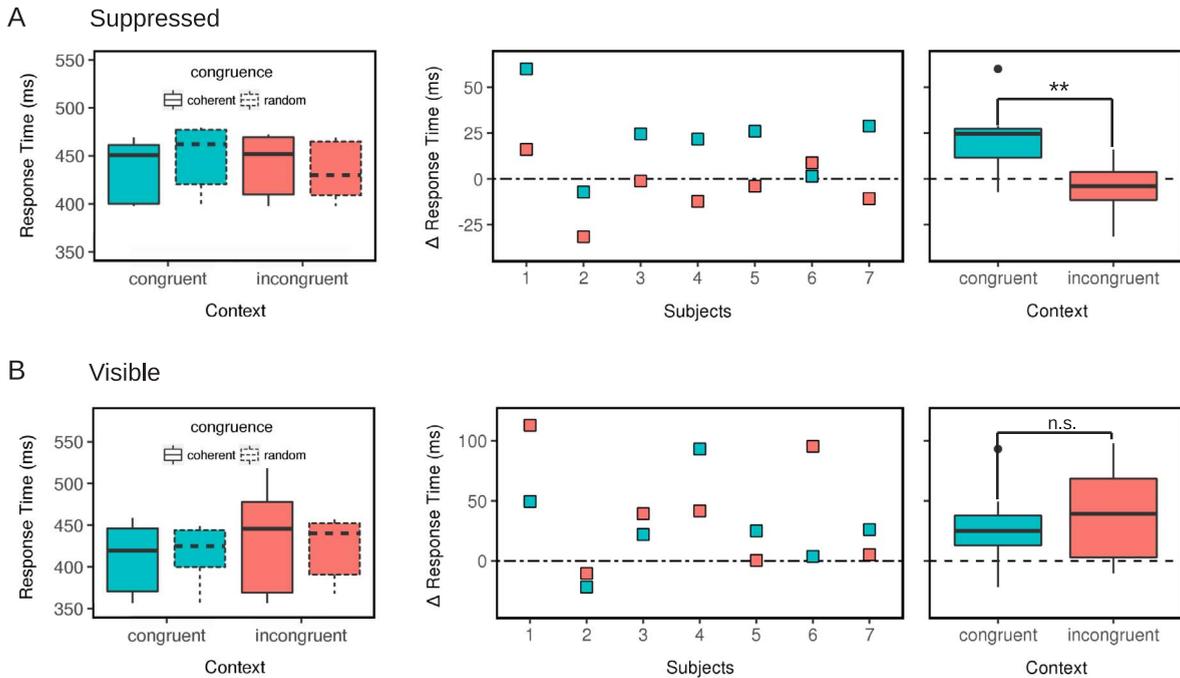


Fig. 2. Reaction time results. (A) Shows the results for suppressed trials, while (B) shows the results for visible trials. The leftmost panel shows the mean response times for all conditions. The coherent condition has a solid outline and the random condition has a dotted outline; blue represents the congruent context, while red represents the incongruent context. Note that subject 1 had outlier mean response times (> 600 ms) that are outside of the scale plotted here. In the center panels, each square shows the difference in reaction time between random and coherent trials for each subject. Right panels show the average of the differences across subjects (p-values: n.s. > 0.05 , * < 0.05 , ** < 0.01 , *** < 0.001). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

To test these two hypotheses, we modelled our data using a Drift Diffusion Model (DDM; Ratcliff, 1978; Ratcliff & McKoon, 2008). This model allows us to capture changes to the rate at which evidence is accumulated (drift rate) separately from changes to the decision criteria (boundary separation), as well as obtaining estimates for the effects of other non-decisional processes, such as stimulus encoding and response execution. After fitting the model for each subject (see Section 2; Fig. 3A), we obtained estimates for our two parameters of interest: the drift rate and the boundary separation. A larger drift rate indicates better quality of evidence accumulation, while a higher response boundary indicates that participants accumulate more evidence before responding.

We next performed an ANOVA on the drift rate in suppressed trials, which revealed a significant interaction between congruence and coherence [$F(1,6) = 6.26$, $p < 0.01$; see Fig. 3B]. Post-hoc analyses confirmed that this effect was due to an increased drift rate for coherent with respect to random trials ($\Delta\text{Drift} = 0.26$, $p < 0.01$) during the congruent context, while no difference was found for trials during the incongruent context ($p > 0.1$). Similar analyses failed to reveal any effect for the response boundary ($p > 0.1$; see Fig. 3C).

One possibility is that the non-decision time was driving the observed effect in the drift. Since in the model presented here the non-decision time was constrained to not vary between conditions, we further fit an additional model to explore this possibility. We removed the across-trial variability from this model to reduce the number of parameters and increase the number of data points per cell (Lerche & Voss, 2016; Lerche, Voss, & Nagler, 2016). We found that as before, there was no effect on boundary and a significant interaction on the drift rate for suppressed trials [$F(1,6) = 8.17$, $p < 0.05$]. Crucially, there was also no effect for non-decision time ($p > 0.1$). These results suggest that the difference in perceptual decision-making between coherent and random trials in the suppressed condition during the congruent context is the result of a qualitative change in the unconscious accumulation of evidence, rather than a change in response strategies.

4. Discussion

Our results show that unconscious perceptual decision-making can be modulated by expectations induced by the statistical context in which the choices are made. We found that when the unconscious portion of a trial was likely to contain coherent information, participants were able to discriminate the overall motion direction faster. Conversely, when the suppressed portion was highly likely to contain only noise, participants did not utilise suppressed information when making their decision. Furthermore, computational modelling revealed that the decisional context modulated the rate at which unconscious information was accumulated. Together, these results indicate that unconscious decisional mechanisms are capable of flexibly adapting to the statistical environment in order to optimise performance.

Such flexible behaviour can be accounted for by a theoretical framework which assumes that the brain's primary purpose is to detect violations of expectations (Friston, 2005; Lee & Mumford, 2003; Mumford, 1992; Rao & Ballard, 1999). According to the predictive coding framework, perception (and in recent versions, action; Friston, 2010) emerges as a result of the continuous comparison between top-down activity (reflecting the expected percept) and bottom-up activity (reflecting the difference between sensory evidence and the expected

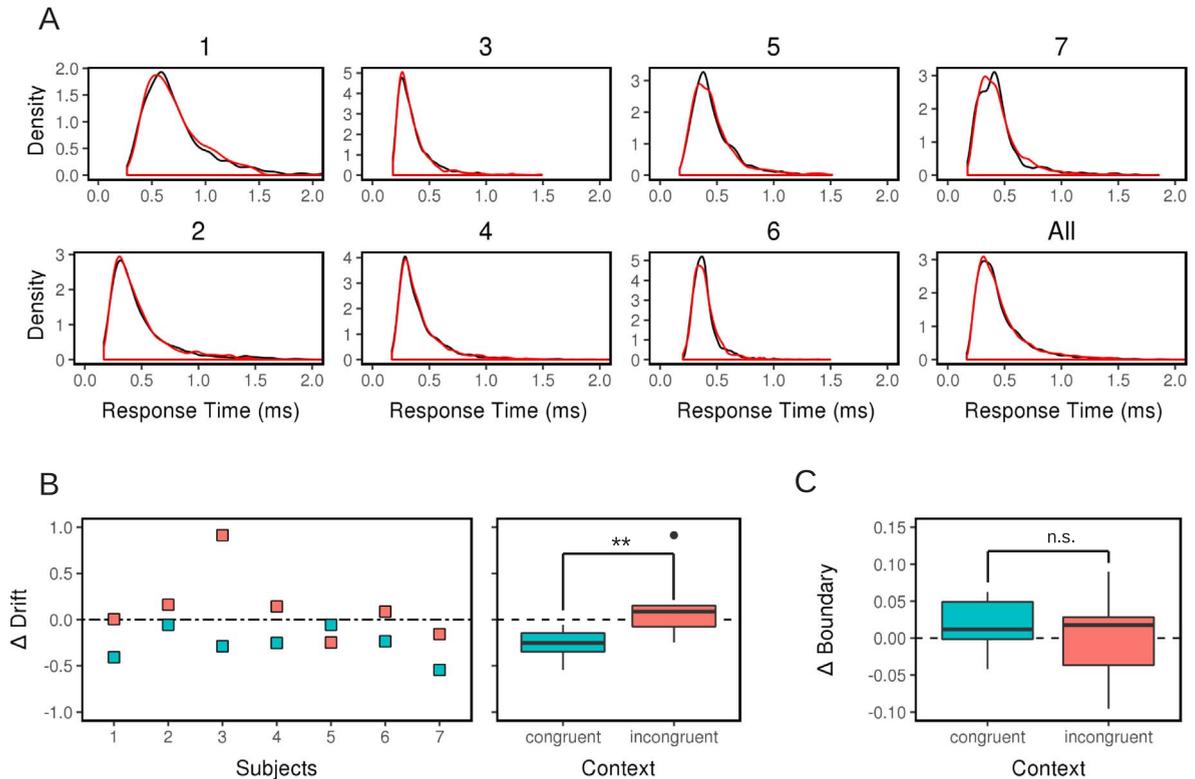


Fig. 3. Drift diffusion model results for suppressed trials. (A) Shows the density of reaction times for original (black) and simulated (red) data for each subject. A Kolmogorov-Smirnov test failed to show any difference between the real and simulated distributions (p -value > 0.1). (B) In the left panel, each square shows the difference in drift between random and coherent trials for each subject. Blue represents the congruent context, while red represents the incongruent context. The right panel shows the average of the differences across subjects. (C) Shows the differences between boundary for each context averaged across subjects. (p -values: n.s. > 0.05, * < 0.05, ** < 0.01, *** < 0.001). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

percept). In this scenario, the difference between the expected and actual stimulus is fed-forward to higher-level regions and used to update current predictions. The new predictions are then propagated backwards to lower regions; top-down activity would therefore play a key role during perceptual decision-making (for a review see Summerfield & de Lange, 2014).

While most studies have investigated the effects of expectations on the starting bias during perceptual decisions (Rao et al., 2012; de Lange et al., 2013; Summerfield & de Lange, 2014), here we focused on the impact of predictions on unconscious accumulation of evidence. By design, participants could not predict the direction of accumulation before the start of the trial, since the directions were not cued and were equally likely to be left or right. Therefore, performance could only be optimised by taking advantage of the statistics of the environment either by increasing the rate of evidence accumulation or by changing the amount of evidence required to trigger a choice. We showed that predictions updated through the statistics of the environment change the rate of unconscious evidence accumulation.

Our results further showcase that unconscious decisional processing can actually be more optimal than conscious processing, as participants were able to utilise the statistics of the environment to improve their performance in the congruent context only for invisible trials. Conversely, there were no differences between the two statistical contexts for visible trials, indicating that participants were unable to discount the first half of the visible trials when the content present in these periods had no utility. In general, the predictive coding framework predicts that expectations should have more influence on perception when sensory evidence is noisy (Hohwy, 2013; Hohwy, Roepstorff, & Friston, 2008). In our case, when the stimuli were suppressed from conscious awareness, the incoming sensory information was more ambiguous, and participants were therefore able to integrate additional sources of information by utilising the block statistics to improve their decisional processing. This finding dovetails nicely with previous work showing that the influence of strong priors in the visual system is maximal when information is noisy but still strong enough to be used (de Gardelle, Kouider, & Sackur, 2010). One of the main ideas under the predictive coding framework is that predictions lead to optimality (Friston, 2010). Our results can help us to understand the role of perceptual awareness in such optimization processes, since participants relied on prior information to maximize performance only when stimulus was invisible and ignored the prior information during visible trials, arguably because maximum performance could be achieved relying exclusively on sensory information.

In summary, we have shown here that predictions can bias performance during unconscious perceptual decision-making. We further demonstrated that these biases are the result of an increased rate of unconscious evidence accumulation in a context where unconscious evidence is often useful, resulting in faster discriminations. Together these results indicate that unconscious decisional mechanisms are capable of flexibly adapting to the statistical environment in order to optimise performance, in line with the predictive coding framework.

Acknowledgements

We would like to thank Jérôme Sackur for helpful discussion and comments. This research was supported by funding from the European Research Council (to S.K.).

References

- Bodner, G. E., & Masson, M. E. J. (2001). Prime validity affects masked repetition priming: Evidence for an episodic resource account of priming. *Journal of Memory and Language*, *45*, 616–647.
- Bodner, G. E., & Masson, M. E. (2003). Beyond spreading activation: An influence of relatedness proportion on masked semantic priming. *Psychonomic Bulletin & Review*, *10*, 645–652.
- Bodner, G. E., Masson, M. E. J., & Richard, N. T. (2006). Repetition proportion biases masked priming of lexical decisions. *Memory & Cognition*, *34*, 1298–1311.
- Brainard, D. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436.
- de Gardelle, V., Kouider, S., & Sackur, J. (2010). An oblique illusion modulated by visibility: Non-monotonic sensory integration in orientation processing. *Journal of Vision*, *10*, 6.
- de Lange, F. P., Rahnev, D. A., Donner, T. H., & Lau, H. (2013). Prestimulus oscillatory activity over motor cortex reflects. *The Journal of Neuroscience*, *33*, 1–11.
- Eriksen, C. W. (1960). Discrimination and learning without awareness: A methodological survey and evaluation. *Psychological Review*, *67*, 279–300.
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*, *360*, 815–836.
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*, 127–138.
- Hohwy, J. (2013). *The predictive mind*. Oxford University Press.
- Hohwy, J., Roepstorff, A., & Friston, K. (2008). Predictive coding explains binocular rivalry: An epistemological review. *Cognition*, *108*, 687–701.
- Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behavioral and Brain Sciences*, *9*, 1–23.
- Jaśkowski, P., Skalska, B., & Verleger, R. (2003). How the self controls its “automatic pilot” when processing subliminal information. *Journal of Cognitive Neuroscience*, *15*, 911–920.
- Kinoshita, S., Mozer, M. C., & Forster, K. I. (2011). Dynamic adaptation to history of trial difficulty explains the effect of congruency proportion on masked priming. *Journal of Experimental Psychology General*, *140*, 622–636.
- Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: A critical review of visual masking. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*, *362*, 857–875.
- Lee, T. S., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, *20*, 1434–1448.
- Lerche, V., & Voss, A. (2016). Model complexity in diffusion modeling: Benefits of making the model more parsimonious. *Frontiers in Psychology*, *7*.
- Lerche, V., Voss, A., & Nagler, M. (2016). How many trials are required for parameter estimation in diffusion modeling? A comparison of different optimization criteria. *Behavior Research Methods*.
- Merikle, P. M., & Daneman, M. (1998). Psychological investigations of unconscious perception. *Journal of Consciousness Studies*, *5*, 5–18.
- Mulder, M. J., Wagenmakers, E.-J., Ratcliff, R., Boekel, W., & Forstmann, B. U. (2012). Bias in the brain: A diffusion model analysis of prior probability and potential payoff. *Journal of Neuroscience*, *32*, 2335–2343.
- Mumford, D. (1992). On the computational architecture of the neocortex II: The role of cortico-cortical loops. *Biological Cybernetics*, *66*, 241–251.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.
- Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*, 79–87.
- Rao, V., DeAngelis, G. C., & Snyder, L. H. (2012). Neural correlates of prior expectations of motion in the lateral intraparietal and middle temporal areas. *Journal of Neuroscience*, *32*(29), 10063–10074.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873–922.
- Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*(5), 347–356.
- Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience*, *22*(21), 9475–9489.
- Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology*, *86*(4), 1916–1936.
- Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, *27*(3), 161–168.
- Summerfield, C., & de Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. *Nature Reviews Neuroscience*, *15*, 745–756.
- Van den Bussche, E., Van den Noordgate, W., & Reynvoet, B. (2009). Mechanisms of masked priming: A meta-analysis. *Psychological Bulletin*, *135*, 452–477.
- Vlassova, A., Donkin, C., & Pearson, J. (2014). Unconscious information changes decision accuracy but not confidence. *Proceedings of the National Academy of Sciences*, *2014*, 1–8.
- Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods*, *39*(4), 767–775.