Effect of Post-decisional Evidence Accumulation on Type-I and Type-II Decisions

A thesis submitted to the degree of

Master of Cognitive Science

Erva Nihan KANDEMIR

Université Paris Cité

École Normale Supérieure

Laboratoire des systèmes perceptifs (LSP)

Supervisors: Pascal MAMASSIAN & Tarryn BALSDON

June 2022

ervanihank@gmail.com

Table of Contents

Declaration of Originality

Integration to the boundary is a decision process that gathers information until it reaches a threshold, resulting in the fastest choices for a given accuracy. This mechanism is optimal since it allows observers to decide which decision to commit to and how long or how much evidence to gather before committing to their decision.

It was recently shown that when observers are not allowed to control the amount of evidence they are presented with, they commit their decision prematurely without accumulating all the evidence they were exposed to, while continuing to monitor additional evidence for evaluating confidence (Balsdon et al., 2020, Balsdon et al., 2021).

Here, we aimed to push the limits and examine what happens if observers encounter a lot of supportive or disconfirming sensory evidence after they already committed to a decision. Based on the previously mentioned findings, that suggest separate mechanisms for confidence and perceptual processes, we specifically investigated the effect of the disconfirming post-decisional evidence separately on 1) the perceptual decision itself and on 2) confidence judgments.

We are not the first researchers to account for the post-decisional information in evidence accumulation frameworks; there are already a number of studies establishing the continued accumulation of post-decisional evidence and how it relates to changes of mind (e.g., Resulaj et al., 2009) and confidence (e.g., van den Berg et al., 2016; Moran et al., 2015; Pleskac and Busemeyer, 2010; Yu et al., 2015). Nonetheless, in these studies, the quality of post-decisional evidence was not systematically manipulated; there was either no further information after the initial decision, or evidence that supported the information presented during the pre-decisional stage to make the post-decisional evidence helpful. (Limitations of these studies are discussed in detail in the introduction section). Thus, what the present study adds to the literature is to ask about the role of the post-decisional evidence that contradicts the evidence preceding the initial decision.

Previously the post-decisional evidence accumulation was generally tested by using the classical Random-Dot Kinematogram stimulus (e.g., Fleming et al., 2018) or sequences of stimuli (as in the Balsdon et al.'s (2020) study). However, using this classic stimulus, it would have been difficult to answer our research question because of the difficulty of presenting enough contradictory evidence after the decision commitment without an abrupt change. Thus, to address our research question, we adapted a stimulus first used to study the pooling of visual motion signals across space (Amano et al., 2009).

This new stimulus enabled the presentation of pro-signal, anti-signal, and noise elements simultaneously. Furthermore, a fraction of elements in the trial, controlled throughout the stimulus duration, changed their status of being signal, anti-signal, or noise. This simultaneous presence of three elements and their gradual status change enabled us to manipulate the stimulus within the trial. In this way, we were able to present a lot of supportive or contradictory evidence after the initial decision but still within the same trial in a smooth and nonabrupt way. In the end, unlike the abovementioned studies, this gradual stimulus transformation within the trial allowed us to better test our research question.

The experimental design was also original in a way that the free-response trials (observers terminate the evidence themselves) were implemented in the interrogation session (the experimenter controlled the termination of evidence). This implementation, which does not exist in the already published studies on the topic, enabled us to control the potential confound effect that can be caused by the different accumulation strategies in the free and interrogation sessions.

In summary, the present study is original in a way that allows us to go beyond the already published studies by 1) asking an unanswered research question (the effect of the disconfirming post-decisional evidence on perceptual decisions and confidence, 2) choosing a novel stimulus, and 3) using an experimental design in the ambition of answering our research question.

Declaration of Contribution

The project was created in collaboration with Pascal Mamassian (PM) and Tarryn Balsdon (TB).

- **Identifying and clarifying the research question:** Although I desired to work on the visual confidence and propose my ideas, PM also suggested potential research questions. We discussed the limitations and originality of these questions together. Both PM and TB helped me in the relevant literature to clarify my research question.
- **Literature Review:** I reviewed the literature on the topic. Both PM and TB aided me in finding the relevant literature and sent me papers that they believed might be helpful in the project.
- **Choice of methodology:** Relevant methodology was recommended both by PM and TB. I researched the suggested methods myself and the methodology to best answer our research question, and necessary updates were decided together.
- **Programming the experiment:** The experiment was written in Matlab by myself and TB. TB always checked, reviewed, and made the necessary improvements in my code. PM wrote the part of the program where the fraction of status-changing Gabors was controlled by a Markov chain.
- **Subject recruitment and testing:** I collected the data.
- **Data analysis:** I carried out the analysis with the help of TB.
- **Interpretation of the results:** We discussed the theoretical interpretation of the results together.
- **Writing the thesis:** I wrote the thesis.
- **Proof-reading the thesis:** Both TB and PM did the proof-reading and helped me with the presentation of my project

I am very grateful for all their contributions.

Abstract

Perceptual decisions are accompanied by confidence, which refers to the observers' subjective estimation of their performance accuracy. Therefore, the whole process can be divided into two parts: Type-I judgments, which the decision-making process leads to, and Type-II judgment, which estimates the validity of the Type-I decision. There is already a rich literature proposing integration to boundary models as an optimal Type-I decision algorithm in which the sensory information is gathered until a decision boundary. These models were recently extended also to explain the construction of Type-II decisions. Still, it remains unclear what happens if observers receive further contradictory evidence after committing to a decision and how this post-decisional evidence affects perceptual and confidence judgments. Here, to clarify the effect of post-decisional evidence on the Type-I and Type-II decisions separately, we exposed 20 human observers to a lot of supportive or contradictory evidence after their estimated decision-time in a motion direction discrimination task. After presenting this post-decisional evidence, observers were asked for a final response and confidence rating. Results show that if observers receive enough contradictory evidence after their initial decision, they become more likely to change their already committed decision compared to the condition in which they received supportive post-decisional evidence. Additionally, this contradictory evidence reduces the sensitivity of the confidence ratings to the total accumulated evidence. Together, these findings suggest that the additional evidence received after the initial response is not entirely ignored: the evidence accumulation mechanism continued to some extent in the post-decisional period for both Type-I and Type-II decisions.

Keywords: decision-making, visual confidence, metacognition, integration to boundary, post-decisional evidence

Pre-registration Document

1.1.Study Information

Hypotheses

Our main hypothesis is that in a visual decision-making task, post-decision evidence will affect confidence but not alter the initial decision. When the post-evidence is in favor of the initial decision, confidence will increase, and when evidence is against the initial decision, confidence will decrease. So, our null hypothesis implies that post-decisional evidence has no effect on confidence and no effect on performance. A second alternative hypothesis is that this post-decisional evidence has an effect on both confidence and performance (the observer will change their initial decision when provided with sufficient counter-evidence).

1.2.Design Plan

Study type

Experiment - A researcher randomly assigns treatments to study subjects, this includes field or lab experiments. This is also known as an intervention experiment and includes randomized controlled trials.

Blinding

For studies that involve human subjects, they will not know the treatment group to which they have been assigned.

Is there any additional blinding in this study?

The participant will not know what condition they are performing.

Study design

We will use a within-subject design. The stimulus will be Global Drifted Gabor Stimulus based on Amano et al.'s (2009) original study. The task will be the motion direction discrimination task. On each trial, the participant must decide if the global motion is to the left or right. There will be two versions of this task: In the Free Task observers can enter their response as soon as they feel ready; in the Replay Task, ¾ of the trials will have a fixed duration and the participant will be cued to enter their final response after that duration. Free Task (Figure-1): We will run the first block to estimate each observers' default decision time for a particular pre-generated trial. Each unique pre-generated trial will be presented multiple times, and the median reaction time of the responses given to the same trial will be calculated for each participant. Then the decision time will be estimated by subtracting the estimated non-decision time (~200 ms) from each of the unique trials' median reaction times. So we will end up with each participants' default decision time for each unique pre-generated trial. These unique decision times will determine the stimulus duration for each condition in the Second Part of the experiment, where we will actually test our hypothesis. Replay Task (Figure-2): Task will be the same as the Free Task, but the stimulus duration will differ depending on the default decision time for each trial calculated based on the data coming from the Free Task. In Replay Task, five conditions (Less, Free in Replay, More Contrary, More Neutre, and More Supportive Conditions) will be presented in a mixed form and in random order. In the Replay Task, all answers will be followed by a confidence rating concerning the last responses given. Confidence rating will be on a scale of 1-4 (1 indicating the minimum confidence level and 4 indicating the maximum confidence level).

Randomization

The trial order in both the Free and Replay tasks will be randomized. All participants will be presented with the same pre-generated stimulus set, but the order of presentation of these pre-generated trials will be randomized between subjects.

1.3. Sampling Plan

Existing Data

Registration prior to creation of data

Explanation of existing data

We collected pilot data from 2 participants and analyzed it to see the approximate time required for the task completion and set the parameters appropriately. We will run 1 final pilot and analyze it to see the parameters change we did after the two previous pilots run smoothly.

Data collection procedures

20 participants who have normal or corrected to normal vision will be recruited by mailing lists and word of mouth. They will be required to speak English to understand the full explanation of the task. Written consent will be requested after explaining the task and before beginning the experiment.

Sample size

We will use a similar procedure that is used in Balsdon et al.'s (2020) study. Therefore, we will include a total of 20 participants by referencing the moderate effect size aimed in this study (d= 0.68) with a power of 0.8 (alpha = 0.05). Participants who have below chance level (50% correct) performance will be excluded and replaced by another participant to preserve the total number of data included in the full analysis.

Sample size rationale

The sample size was determined by power analysis by aiming for moderate effect size (d= 0.68) with a power of 0.8 (alpha = 0.05).

Stopping rule

No more than 20 participants' data will be included in the final analysis.

1.4. Variables

Manipulated variables

In the Replay Task, we will manipulate the strength of the presented evidence, by predicting that increasing the proportion of Gabors moving in a coherent direction will increase performance and decrease reaction times. We will also manipulate the duration of stimulus presentation in the Replay task: Less, Free response, and More Conditions. In the More conditions of the Replay Task, we will manipulate the post-decision evidence which will either confirm, contradict, or have no change on the information presented up to the decision point. This manipulation will allow us to test our main hypothesis, where this post-decision evidence is expected to have an influence on confidence but not on the final perceptual decision.

Measured variables

The outcome variable will be the choice on each trial, the reaction time of each choice, and the confidence rating in the Replay Task.

1.5. Analysis Plan

Statistical models

The main analysis will be to test our hypothesis to examine the effect of extra post-decisional information that contradicts the pre-decision information on the performance and confidence ratings. This effect will be examined by non-parametric within-subject statistics on the d’ and confidence ratings between five conditions. Specifically, to test our hypothesis, we will investigate whether the observers are more likely to change their initial response in the More Contrary Condition than in the More Supportive Condition. We will further examine the data by using Drift Diffusion Models to explain the decision-making quantitatively and to see whether the Type-I and Type-II decisions have the same or different boundaries. We will first fit parameters of different versions of models to describe behavior in the Free trials, and then compare how well this model predicts responses in the More conditions across accumulating all evidence and accumulating evidence only to the bound.

Inference criteria

The null hypothesis will be rejected at the alpha level of 0.05 ($p<0.05$ to reject the null hypothesis). Our hypothesis also includes the performance, so the difference in the performance level across conditions will also be assessed against an alpha level of 0.05.

Data exclusion

Participants who have below chance level performance (50% correct) on the first block and participants who wait more than enough time to respond both in the Free Task (more than 1.5 seconds for 80% of trials) and Replay task (waiting in the majority of trials until the end of the stimulus presentation in the Free in Replay condition) will be excluded. Additionally, participants who will show biases consistently (80% of trials) choosing a certain confidence level in the confidence task may also be excluded. All the excluded participants' data will be replaced by another participant to preserve the total number of data included in the full analysis. Even if participants only give a very late response (response given after 3 sec) in a small number of trials, that is not enough to exclude the whole data (not correspond to 80% of the trials); these trials with very late responses will not be included in the analysis

Missing data

Only complete data sets will be included in the analysis

Exploratory analysis

We have no current exploratory analysis planned.

Effect of Post-decisional Evidence Accumulation on the Type-I and Type-II Decisions

1. INTRODUCTION

Every second, we make countless decisions, from deciding an object's color or motion direction, to what we will eat for dinner. While making these decisions, we rely on our perceptual system, which gathers sensory evidence provided by the physical environment. But our perceptual system is itself imperfect and relies on probabilistic inference to resolve uncertain sensory information (Helmholtz, 1856). Thus, sensory decision-making, fed by these probabilistic inferences, is also naturally imperfect. Given this suboptimality in perceptual decision-making, rather than blindly trusting the validity of the output of those noisy inferences, a proper perceptual decision requires an accompanying mechanism to evaluate the percept quality. This concomitant mechanism to make a judgment over another judgment is known as metacognition and can be quantified in vision research by asking the observer to rate their confidence level on the correctness of their perceptual decision (e.g., "how confident are you that your decision on the color of this object is correct?"; see the review on visual confidence by Mamassian, 2016).

A wide range of recent studies demonstrates that this subjective sense of decision accuracy is useful in monitoring the decision-making process (Boldt et al., 2019; Carlebach et al., 2020), modulating the subsequent cognition (Desender et al., 2019; van den Berg et al. 2016) and even in being a predictor for learning (Veenman et al., 2004). Considering this benefit of confidence and that ill-adjusted confidence can cause cognitive biases (Rollwage et al., 2020), well-tuned assessments of decision quality are important for the guidance of decisions, especially in the absence of external feedback.

The degree to which confidence tracks objective performance is known as metacognitive sensitivity (Fleming and Lau, 2014). Although confidence judgments are also naturally subject to physical uncertainties and various biases, previous studies have established that humans and some non-human animals can give a reasonable estimation of the accuracy of their own performance (Gigerenzer et al., 1991; Kepecs et al., 2012). Admitting that metacognitive sensitivity is not always perfect and contains some systematic discrepancies, well-calibrated confidence ratings still support the view that our brains are capable of calculating the noisy probabilities and analyzing their validity.

In this manner, an important distinction appeared between Type-I judgments, which the decision-making process leads to; and Type-II judgment, which estimates the accuracy of the Type-I decision (Clarke et al. 1959; Galvin et al., 2003).

1.1 Computational Models Proposed for Type-I Decisions

Before moving on to our research question, which examines the effect of the post-decisional evidence on the Type-I and Type-II judgments, it is important to present the most accepted mechanism for Type-I decisions proposed in the literature, and its extension to account for the post-decisional evidence accumulation.

The process underlying the construction of the Type-I decision has been explained in great detail by evidence accumulation models. These models suggest that noisy sensory information is sequentially gathered until a threshold is reached. This threshold is the decision bound: the point at which enough evidence has been gathered to favor one choice over another and the observer commits to that choice (Bogacz et al., 2006; Ratcliff et al., 2016). As a result, these sequential sampling models are able to explain both the choices given at the end of the accumulation and the decision time for committing to this choice. These sequential sampling models are also supported by psychophysics and neural data (Gold and Shadlen, 2007; Shadlen and Kiani, 2013; Hanks et al., 2015).

Notwithstanding, in these sequential sampling models, the evidence accumulation for the Type-I response was explained by a strict boundary that assumes no further evidence intake after that point. Yet, recent studies have demonstrated that the ongoing motor actions toward the choice (Resulaj et al., 2009) and the already-committed initial decision (Fleming et al., 2018) could be revised in the presence of additional evidence and lead to occasional response alterations to improve the performance.

So, this occasional mind-changing in the presence or even in the absence of new evidence (Resulaj et al., 2009) challenges the assumption of a strict bound in sequential sampling models. These researchers accounted for changes of mind by extending the model with two additional parameters: a change-of-mind bound and a change deadline.

1.2 Limits of the Previous Change of Mind Studies

Changes of mind, which principally serve for correcting the initial errors (Couchman et al., 2016), can be explicit, seen in the overt behavior as the already committed choice altered by the alternative, or implicit, seen by changes in internal decision states prior to decision commitment (for a recent review in the change of mind literature, see Stone et al., 2022).

In their seminal work, Resulaj and his colleagues (2009) proposed that during the response preparation and the motor response execution stages, the continuous accumulation of the extra perceptual evidence can lead to a response change. Yet, in this study, there was no stimulus alteration that required a systematic update of the already accumulated evidence. Thus, the response change observed in this study was occasional and not determined by the quality of the evidence received after the decision. However, to see the systematicity of the mind-changing, the experimental paradigm should carefully manipulate the presented evidence before and after the decision.

The effect of continued evidence accumulation on response updating has been studied throughout the decision by manipulating the evidence while the decision is still in progress (Holmes et al., 2016). In a recent study that tested the specific role of confidence in the selective integration of post-decisional evidence (Rollwage et al., 2020), additional evidence was given after the explicit commitment of the decision. However, in this study, the evidence presented after the decision only confirmed the evidence that preceded the decision, leaving our question open: what happens if the post-decisional evidence disproves the evidence that preceded the decision. This open question leads to two distinct predictions proposed by reflecting and absorbing boundaries (Feller, 1968; Zhang et al., 2009): absorbing boundaries predict that the choice is dominated by the pre-decision evidence and accumulation terminates strictly, while reflexive boundaries argue that the choice can be affected by the post-decisional evidence that allows response changes in the presence of contradictory evidence even after decision commitment.

A promising study was already conducted by Fleming and his colleagues (2018) in which the observers were presented with pre-decision evidence until the required evidence was accumulated and post-decisional evidence after they committed to their initial choice and confidence ratings. This design can be easily used to compare pre- and post-decisional accuracy and metacognitive sensitivity in the case of extra contradictory evidence presented after the decision. A potential limitation of this design, however, is that the observer may consider the post-decisional evidence as a new trial and not compare this new evidence with the initial decision. So, instead of a free-response paradigm where the evidence accumulation is terminated by the observer, an interrogation paradigm where the evidence is given for a predefined quantity of evidence (Bogacz et al., 2006) could be a good protocol to test these two different predictions on the post-decisional evidence accumulation. A critical aspect in this paradigm would be to change the stimulus within a trial (as proposed by Zhang et al., 2009) and ensure that the decision boundary is already reached before this stimulus changes to guarantee that enough contradictory evidence is available as post-decisional evidence. Additionally, to prevent that observers notice the change in the stimulus after the decision, this change needs to be gradual rather than abrupt or discontinuous, which is difficult to ensure in the classic random-dot paradigm that is often used in the abovementioned experiments testing the continuous evidence accumulation.

To the best of our knowledge, no existing study tested the effect of newly received evidence on the already committed responses by providing disconfirming evidence within the trial but after the observers committed to their decision. Thus, there is still a lack of knowledge on whether the Type-I evidence accumulation inevitability ends in the decision boundary and completely ignores the evidence presented after that point (prediction by the absorbing boundaries), or instead the accumulation restarts systematically at some point in the presence of enough evidence that contradicts the evidence gathered for the already committed decision (prediction by the reflexive boundary).

1.3. Computational Models Proposed for Type-II Decision

Given the high benefit of the accompanying of confidence in effective decision-making, there has been a growing interest in understanding the underlying computations of confidence in a decision-making framework. There is already a significant number of promising studies that build different computational models for explaining how confidence is formed (e.g., Kiani, Corthell, and Shadlen, 2014; Pleskac and Busemeyer, 2010.). Yet, their respective assumptions vary, and they originated from different models of decision-making.

These models date back to Vicker's Balance of Evidence model (BoE), which explains confidence as the distance between the accumulated evidence for the losing race and the decision bound in an independent race model (Vickers, 1979). Following Vickers' proposal, some largely debated cognitive theories imply that confidence and decision occur at the same time and are based on the same evidence: observers' choices, decision times, and judgments on the choice accuracy are grounded in the same cognitive process (Kiani et al., 2014; Kiani and Shadlen, 2009). According to these theories, when the accumulated evidence reaches the decision bound and the decision-maker ends the evidence deliberation, they also quantify their confidence based on the same evidence that was used for the decision (Fetsch et al., 2014). Following this ’shared evidence’ assumption, at the point when the decision-maker commits a decision between two alternatives, the cumulative evidence in favor of the chosen alternative will always be the same, which should always result in the same confidence for that given choice.

However, this assumption cannot explain the examples where confidence judgments are not formed based on the full quality and quantity of the evidence used in the perceptual decision. For example, the evidence used to make decisions might be subjected to additional noise, causing additional variability in confidence ratings. Likewise, observers’ confidence might incorporate additional evidence such as their own decision time (Kiani et al., 2014). This additional evidence that is used in the confidence reports but inaccessible to decision-making may lead to error detection (Rabbitt et al., 1977). Furthermore, human observers have a tendency to exaggerate their confidence in a decision based on noisy information (Drugowitsch et al., 2014; Kiani et al., 2014), and they also can report low-level confidence while performing their decision above chance (Kunimoto et al., 2001). Additionally, these theories are not capable of explaining the mind-changing after the initial choice by integrating additional information after the initial decision. Observers also change their confidence judgments even if they did not receive extra information about their task and even if they did not receive any further information about how well they had done the task (van den Berg et al., 2016). Thus, these single-stage theories, which propose that confidence is formed by the same information that underlies the decision, do not suffice to explain the additional noise and evidence that confidence ratings are subject to and mind changing in the decision and confidence judgments.

All these examples call for a different and more precise explanation for confidence. Considering this uncertainty regarding the nature of the confidence judgments, alternative theoretical models extended the existing models by focusing on a new assumption that evidence accumulation may continue following the decision commitment. These alternative models thus qualified the metacognitive judgments by considering additional post-decisional evidence accumulation (Moran et al., 2015; Pleskac and Busemeyer, 2010; Yu et al., 2015).

The most promising extension was made by Pleskac and Bussemeyer (2010) in the Two-Stage Dynamic Signal Detection Theory (2DSD), in which they proposed that confidence judgments are derived from evidence collected up to the time of choice, plus evidence collected by a second stage of accumulation after making a choice that builds on the already accumulated evidence in the first stage. This notion of post-decisional evidence accumulation proposed in the 2DSD theory is supported by neurophysiological work (Murphy et al.,2015) and by behavioral results (e.g., Resulaj et al., 2009).

In this 2DSD model, confidence is formed by the so-called ‘time-based stopping rule’ in which the confidence is evaluated after a fixed time following the decision. In a more recent study that again supports the post-decisional evidence on the confidence construction, an alternative ‘evidence-based stopping rule’ is proposed (Desender et al., 2021). Still, no matter which stopping rule is implemented, such theories deny the previous theories proposing that confidence is driven solely by the information used in the decision-making process and support the key insight that the choice boundary is not necessarily the threshold for the accumulation, but evidence accumulation can continue even after the choice is given in the decision boundary and form the confidence judgments.

To summarize, although there is still a lack of agreement on the precise mechanism behind the construction of confidence judgments (Yeung et al., 2014), a majority of successful recent models agreed there is a role of post-decisional evidence integration in the confidence construction.

1.4. Online metacognitive monitoring during the decision making

Despite the post-decisional accumulation that focused on the recent models mentioned in the previous session, this does not necessarily mean that observers are entirely lacking the metacognitive monitoring during the pre-decisional stage. In fact, there is recent evidence emphasizing the online metacognitive experience that emerges from the decision-making process itself (Dotan et al., 2018).

So, the key assumption of this view is that confidence is not the pure result of the evidence evaluation until the decision point nor the result of evidence solely collected after the choice, but it is rather an online signal that reflects the probability of being correct at any time point and helps the observer understand how much information they need to reach the decision. Accordingly, the accumulation termination rule is controlled by confidence itself. This online monitoring role that confidence is in charge of helps the observer to regulate their decision, terminate the evidence accumulation, and calculate how much information they need to reach the decision during the choice formation process itself (Balsdon et al., 2020). This dynamic role of confidence in the evidence accumulation framework can explain the decision to withhold and search for more information in the decision-making process (Meyniel et al., 2015; Desender et al., 2018). Thus, the online and the post-decisional explanation of confidence are complementary. Even supposing the subjective estimation of the validity of the decision is calculated online, it can still be subjected to different post-decisional evolutions to become precise enough to report.

The relationship between Type-I and Type-II judgments and the online and post-decisional accounts of confidence in the decision-making framework was very recently investigated by Balsdon et al.'s (2020) study by asking observers to make their Type-I and Type-II judgments in three conditions differentiated with the quantity of presented evidence. This study suggests that evidence accumulation ends up with a strict boundary for the decision-making leading observers to commit to their decisions early, before the accumulation of all the available evidence, whilst the accumulation continues to monitor additional evidence for evaluating confidence. However, in this previous study, the quality of post-decision evidence was not specifically manipulated. To maintain a reasonable metacognitive efficiency, confidence judgments also need to be updated in sync with the newly presented evidence by comparing them with the initial decision and re-calculating the validity of the initial decision in this new context. In order to see this revision in confidence about an already completed decision, the new evidence needs to be quantitatively and qualitatively compared with the already committed decision.

1.5. Current Project

Altogether, this previous literature shows us that, while the processes underlying evidence integration have been explained in great detail, and although some of these models are detailed enough or successfully extended to also explain evidence accumulation for metacognitive judgments, still little is known about the accumulation of post-decisional evidence that comes after the decision and how this extra information can shape the Type-I and Type-II decisions separately.

1.5.1. Research Question

The objective of this project is to investigate how Type-I and Type-II judgments are affected by post-decisional information, especially if this additional information conflicts with information used in the initial decision. Our research question thus asks about the role of post-decisional information on Type-I and Type-II judgments.

To examine our research question, the limits were pushed by providing observers with either a lot of supportive evidence in favor of the correct choice or a lot of evidence against the correct choice within a trial but after the Type-I decision has already made.

1.5.2. Hypothesis

Based on findings in the literature that suggest that Type-II evidence accumulation continues even after Type-I evidence accumulation reaches its bounds, our main hypothesis is that in a visual decision-making task, disconfirming post-decisional evidence will reduce confidence but not revise the initial decision. Still, the studies showing the revision of decision in light of the extra information leads us to a second alternative hypothesis that the observer will also change their initial decision when provided with sufficient counter-evidence. Finally, the null hypothesis is that post-decisional contradictory evidence has no effect on the Type-I performance nor on Type-II judgments.

To test our hypothesis, a motion direction discrimination task followed by a confidence judgment was used in the present study. First, each observer's required evidence for a specific decision was estimated. Then, the same evidence was presented in five different Type-II contexts in which the quantity and/or quality of presented evidence was manipulated within the trial. In terms of quality, a lot of extra information that either supports or contradicts the correct response was presented after this estimated decision point.

2. METHODS

2.1. Subjects

A total of 20 participants (13 females; age: M = 31 years, range = 23-55 years) with normal or corrected-to-normal vision were recruited via the RISC mailing list (the French "Relais d'Information en Sciences de la Cognition"). The sample size was selected on the basis of pilot experiments and a power analysis by aiming to detect a moderate effect with a power of 0.8 (alpha = 0.05). Participants signed a consent form and received written information prior to running the experiment, and they received monetary compensation at the end of each session. The experiment protocol was approved by the "Conseil en éthique pour les recherches en santé" (CERES) and pre-registered on the Open Science Framework platform that can be accessed at (https://osf.io/9dp6u), in accordance with the Declaration of Helsinki (2013).

2.2. Apparatus

The experiment was run on a 17-inch LCD monitor (ViewSonic) at a frame rate of 75 Hz with a resolution of 1280x1024 pixels. The experiment was programmed and presented using MATLAB (Mathworks) and the Psychtoolbox-3 (Kleiner et al., 2007; Brainard, 1997; Pelli, 1997) running on a Mini Mac. The distance between the subject's eye and the computer screen was approximately 40 cm, stabilized by a chin rest.

2.3. Stimulus

The stimulus was a global-motion stimulus based on the original study by Amano et al., (2009). In this stimulus, global motion is inferred from the combination of multiple local motions of patterns (Gabor patches) presented at random orientations. We chose to have 330 drifting Gabor patches with a Michelson contrast of 0.5, arranged in a regular 0.75x0.75 grid within a 2D torus, extending 2-16° in diameter (see Figure 1.a). A fixation point was presented in the centre. Each Gabor patch had a spatial frequency of 3 c/deg and a random orientation (between 0° and 180°). The phase of the Gabors were updated on each frame to mimic a global drift at a speed of 1 deg/s. A global direction of motion was defined on each trial, but the local direction of each Gabor could be consistent with the global direction (Pro-signal Gabor), opposite to the global direction (Anti-signal Gabor), or a random direction (Noise Gabor). The task was to report which of two alternatives (rightward and leftward directions, the directions of the Pro-signal and Anti-signal Gabors) the global motion direction corresponded to.

The local drifting speed of each Gabor patch was then calculated by a cosine function of the angle between the carrier orientation of the individual Gabor patches and the 2D vector that was given depending on the status of the Gabor patch, namely

$$S_C = S_G \cos(\theta - \theta_G)$$

where S_C is the local drift speed the individual Gabor patch required to be consistent with the given 2D motion, S_G is the global motion speed (always 1 deg/s in this study), θ is the local orientation of an individual Gabor element, and θ_G is the global direction that depends on the status of the Gabor patch on the particular trial (see Figure 1.b for the creation of pro-signal, anti-signal, and noise Gabors patches).

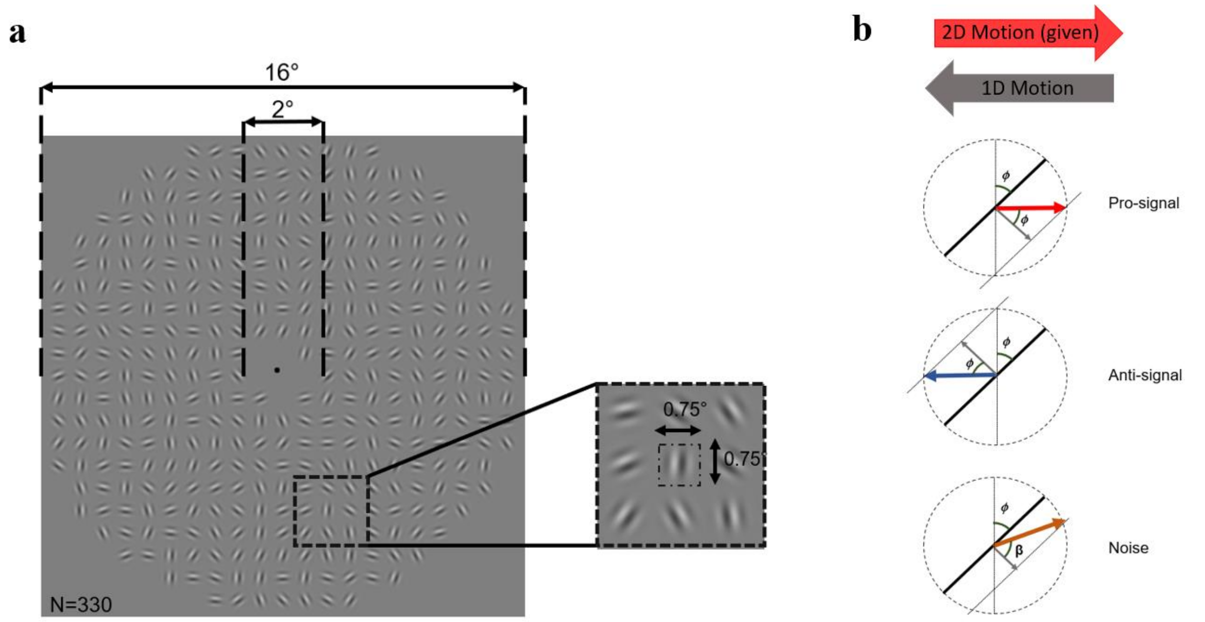


Figure 1: **Global Motion Stimulus**. **a.** Stimulus used in the experiment with an enlarged view showing the regular 0.75x0.75 grid of individual Gabor patches. **b.** Illustration showing how the pro-signal, anti-signal, and noise Gabors were created. The grey dashed circle represents the area of a Gabor patch; the central solid grey line represents the orientation of the Gabor patch, the red arrow serves as the 2D vector given to the Pro-signal Gabors (rightward in this diagram), the blue arrow shows the 2D vector for the Anti-signal Gabors (opposite of the pro-signal 2D vector-leftward in this diagram), the orange arrow represents the 2D vector for an example noise Gabor, and the gray arrows show the 1D component vectors orthogonal to the orientation for each Gabor.

Each individual Gabor patches provided certain evidence for the global signal direction in the trial, depending on their orientation and their status of being pro-signal, anti-signal, or noise Gabor patches that determine the direction of the 2D vector

$$E_S = S_G.\cos(\phi)^2$$

$$E_A = -S_G.\cos(\phi)^2$$

$$E_N = S_G.\cos(\phi).\cos(\beta)$$

where S_G is the magnitude of the 2D vector that gives the global motion speed, ϕ is the carrier orientation of the individual Gabor patches, and β is the angle between the local vector perpendicular to the carrier of an individual Gabor and the 2D noise vector that could be between 0° and 360° (see Figure 1.b).

The first cosine in each formula is to weight the magnitude of the 1D velocity consistent with the 2D vector of the global direction depending on the orientation of each Gabor element

(ϕ). The second cosine is to weigh the contribution of the 2D vector direction to the global trial direction. Since the anti-signal elements have the opposite global direction of the pro-signal elements, the contribution of the anti-signal element to the global trial direction is just the negative of the evidence contribution of the same pro-signal element to the global trial direction. For the evidence coming from the noise Gabors, since the noise elements are consistent with random global directions, the second cosine in the E_N formula calculates the contribution of the 2D noise vector to the direction of the 2D signal vector in that trial.

Thus, the total evidence in one frame is the mean evidence provided for the trial direction by each of the pro-signal Gabors (E_S), anti-signal Gabors (E_A) and noise Gabors (E_N) in that frame:

$$E_{frame} = f_S.E_S + f_A.E_A + f_N.\ E_N$$

where f_S is the fraction of pro-signal Gabors, f_A is the fraction of anti-signal Gabors, f_N is the fraction of noise Gabors in that frame.

The fraction of pro-signal, anti-signal, and noise elements was not stable during the stimulus presentation. The status of each Gabor patch of being pro-signal, anti-signal, or noise elements was changed in every frame throughout the stimulus duration and controlled by a three-state Markov chain, where the three states are the fraction of the pro-signal elements, the fraction of the anti-signal elements and fraction of the noise elements in that particular frame (for a detailed explanation of this three-state Markov chain and the stimulus generation, see Supplementary Note-1). In order to generate a gradual transformation, rather than a sudden switch, within a trial, the chain included a constraint on the fraction of Gabors that were allowed to change their status each frame. We chose this constraint to be 0.1, so that 10% of Gabor patches could change their status, irrespective of whether they were pro-signal, anti-signal, or noise elements. In order to give enough variability in the stimulus evidence, the desired fraction of pro-signal and anti-signal elements was not set to a fixed number but instead resampled at each frame following the constraint of the Markov chain. For creating different difficulty levels for detecting the global motion direction, the desired pro-signal fraction (f_S) was chosen from 6 different ranges (each level is divided into 0.15 steps, between 0.45 and 0.85). The anti-signal fraction (f_A) was randomly chosen from the range between 0.15 and 0.30, stable for all difficulty levels. Finally, the remaining fraction (1-f_S-f_A) gave the fraction of noise elements (f_N), and an increase in the noise fraction made it difficult to see the coherent global motion, so the trial became more difficult.

A total of 60 unique trials (10 at each difficulty level and with an equal number of trial directions (5 right 5 left) at each difficulty level) were pre-generated. The trials were defined for a maximum duration of 3 seconds, but could loop for a maximum of 30 seconds until the participant responded (rarely more that 3 seconds). All participants were repeatedly presented with these 60 pre-generated trials in a random order over the experiment.

On a practical note, the phase updates of each of the 330 Gabor patches, which gives the local velocity of each Gabor patch in each frame, and the orientation of each Gabor patch which is consistent during the whole trial, were saved for presenting the pre-generated trial on the screen. Additionally, the coherent global direction of each pre-generated trial, the status of each Gabor patches in each frame, the ranges from which the pro-signal fraction was randomly selected, and the random number generator seeds that provide the random number for the sampling of pro-signal Gabor fraction and anti-signal Gabor fraction within the given range were saved in order to re-create the same trials in the second session.

2.3.1. Relevance of the stimulus for answering the research question

The stimulus chosen for the study was highly suitable for answering the research question because this stimulus enabled the presentation of pro-signal, anti-signal, and noise elements simultaneously. This simultaneous presentation gave the flexibility to manipulate the balance between pro-signal and noise elements to create different difficulty levels and, at the same time, to gradually change the ratio of pro-signal and anti-signal elements after the decision time, which in the end enabled us to present a lot of supportive or contradictory evidence in that limited time without it being noticeable for the observer.

Thus, instead of the classic Random-Dot Kinematograms, in which one-dimensional dots are consistent with only one motion direction (or multiple directions would create motion transparency), Gabor patches in which the motion direction is consistent with multiple directions were preferred to be used in this study. Moreover, instead of a stimulus where Gabors are presented sequentially, which would enormously extend the trial duration in the current experimental design (many sequential Gabor patches would be needed in the More conditions), the global motion stimulus was preferred to provide multiple evidence simultaneously.

2.4. Procedure

The task was a motion direction discrimination task in which observers were asked to indicate which of the two possible directions (rightward or leftward) corresponded to their perceived global motion. Observers were instructed to press the left arrow key for the trials in which the perceived motion appeared to cohere leftward and the right arrow key for the trials in which the motion appeared to cohere rightward. All participants first completed the Free-Task session in which they entered their responses as soon as they felt ready. Then all participants ran the Replay-Task session in which ¾ of the trials had a fixed duration, and participants were cued to enter their final response after that duration. In the remaining ¼, the trials were identical to the Free Task, and participants were free to respond when they felt ready. The first session lasted approximately 40 min (including the training session), and the second session lasted approximately 1 hour. We now describe these sessions and tasks in detail.

2.4.1. Free Response Task

In the first session, a total of 600 trials (10 repetition of 60 pre-generated trials) were presented in random order. The session was divided into 20 blocks with 30 trials in each. Participants were instructed to give their responses as soon as they reached a decision, and the stimulus was presented until the response was given (up to 30 seconds, 10 repetitions of 3 seconds loops). Participants did not receive any feedback on their accuracy but received a reminder to respond faster if they took more than 1.5 sec to respond. Once the participant gave the response, the stimulus disappeared, and a new trial began after a 500 ms inter-stimulus interval (Figure 2).

The aim of running the Free Task is to estimate each observer's default decision time for a particular pre-generated trial. This estimated decision time was found by subtracting the estimated non-decision time (100 ms) from the median reaction time of the responses given to the 10 repetitions of the same trial. These unique decision times for each pre-generated trial determined the stimulus duration for each condition in the second session of the experiment, the Replay Task, where the main hypothesis was tested.

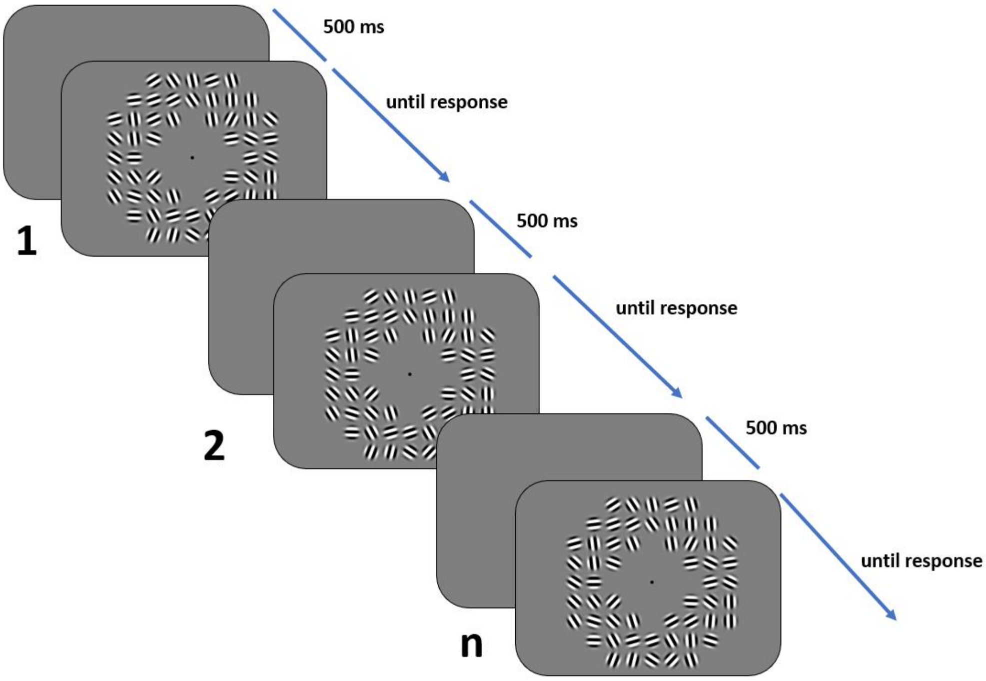


Figure 2: **Procedure in Free Response Task** Experimental paradigm in the Free Response Task in which the evidence sequence was terminated by the participant's response, and the new trial began after 500 ms. inter-stimulus interval. Note that the stimulus presented here is an illustration only.

2.4.2. Replay Task

In the Replay Task, in 3/4 of the trials, the stimulus was presented for a fixed number of frames depending on the decision time for each trial estimated in the Free Response session (see Figure 3).

The task from the observer's perspective was the same as the Free task: they were instructed to report the motion direction as soon as they decided. However, in three novel conditions, the stimulus continued after the initial response and observers were asked for a final response after the fixed trial duration ended. This provided us with the guarantee that the observers saw any extra evidence that was presented after the decision time, which was crucial for testing our hypothesis, even if they had already made a decision and gave responses before the end of the trial. To make sure that participants gave their responses as they did in the Free Task, they were asked to give their responses as soon as they decided, but they were instructed that sometimes the stimulus would be presented very briefly, and sometimes the stimulus would continue to be presented even if they give a response, and in that case, they would need to give a final answer at the end. Importantly, observers were explicitly instructed *a priori* that they were free to change their initial response or not when they were asked for a final response.

There were a total of five different conditions in this task: one “less”, one “free-in-replay”, and three “more” conditions. In the *Less condition*, the exact same pre-generated trials that were presented in the Free Response session were presented but for only half the number of frames required by the observer to respond to that particular trial (half of the estimated decision time which was calculated based on the median reaction time in the 10 repetitions of the responses given to the same trial in the Free Task). In the *Free-in-Replay condition*, the trial was the exact same trial presented in the Free Response Task: the pre-generated trials were presented for up to 30 seconds by the 10 repetitions of 3 seconds loops. Unlike the other conditions, in the Free-in-Replay condition, which is embedded in the Replay task to encourage observers to give their responses in the same way they did in the original Free Task. Once the observer gave a response, the stimulus disappeared immediately.

In the three *More conditions*, the exact same trials that were presented in the Free Task were presented up to the default decision frame calculated for that specific trial, and then the stimulus was continued to be presented for more time (twice the number of decision frames plus 200 ms, i.e. ~15 additional frames). For example, if the median number of frames that an observer requires to give a response to a specific pre-generated trial is 65, the default decision frame would be approximately 58 (median reaction time - non-decision time , ~7 frames), the number of frames in the Less condition would be around 29 (half of the default decision frame), and the number of frames in the More conditions would be approximately 131 (twice the decision frame plus 15 frames).

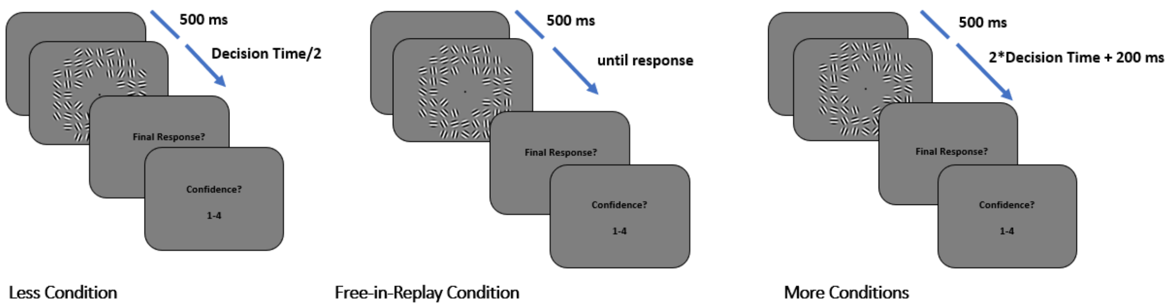


Figure 3: **Procedure in Replay Task.** The stimulus duration varied depending on the condition, but observers were still free to give their responses when they felt ready, which immediately terminated the trial in the Free-in-Replay (and Less conditions in the case where the stimulus had not already stopped) but did not terminate the stimulus presentation in the More conditions if the response was given earlier than the fixed stimulus duration time. In that case, the stimulus continued to be presented, and observers were asked to give their final response at the end of this fixed time. Note that the stimulus presented here is an illustration only.

In all three More conditions, the pre-generated trial was presented again up to the decision time, and then after the decision time, some extra stimulus was presented. Thus, the total trial duration was the same for all the More conditions of a given pre-generated trial. However, the evidence that was presented in the extra time (the time after the decision time and before the end of the trial) either contradicted, supported, or had approximately the same evidence as the evidence that was already presented up to the decision time.

Depending on the quality of the evidence that was presented after the decision frame, the More conditions were divided into three conditions. In the *More Neutral condition*, the continuation of the trial that was presented in the more time was generated by generating a new fraction pro-signal from the same range used in the pre-decision. In the *More Supportive condition*, the extra information was generated by selecting the fraction pro-signal from a new range in which the minimum and maximum points of the pre-decision pro-signal fraction range increased by 0.15. In the *More Contradictory condition*, in the extra frames, the fraction pro-signal was chosen from the range of 0.15-0.30. In all More conditions, despite this change in the pro-signal fraction range, the noise fraction was kept the same as in the Free Task. Thus, the manipulations in the pro-signal fraction range automatically affect the range of the anti-signal fraction, and so the ratio between pro-signal and anti-signal elements in the More time of the More Contradictory condition was the reversed version of the pro-signal/anti-signal ratio in the pre-decision stage. For example, if the fraction pro-signal in a given trial was chosen from a range of 0.45-0.60 in the pre-decision stage of the trial, the new range for the post-decisional stage of the same trial became 0.60-0.75 in the More Supportive condition, 0.15-0.30 in the More Contradictory condition and stayed stable in the 0.45-0.60 in the More Neutral condition. (see Table 1 in Tables for the fraction of pro-signal, anti-signal, and noise for the post-decisional stage of all three more conditions).

As a result of these manipulations in the upper and lower limit of the ranges, if the large fraction of Gabor patches have a rightward direction up to the decision time (this means the global trial direction is rightward for that case), in the More Contradictory condition, the fraction of this rightward Gabor patches (Pro-signal Gabors in that case) was decreased while the fraction of the Gabors patches that have a leftward 2D direction (Anti-signal Gabors in that case) was increased in the more time. By contrast, in the More Supportive condition, this fraction of pro-signal Gabor was increased even more in the more time, and the fraction of anti-signal Gabors was decreased. Finally, in the More Neutral condition, the fraction of pro-signal and anti-signal Gabor remained, on average, the same. Thus, the average evidence for the given trial global direction in the more time remained similar to the average pre-decision evidence in the More Neutral condition, increased in the More Supportive Condition, and decreased in the More Contradictory condition (see Figure 4 for the time-by-time evidence change in all 5 conditions of Replay Task for a particular pre-generated trial).

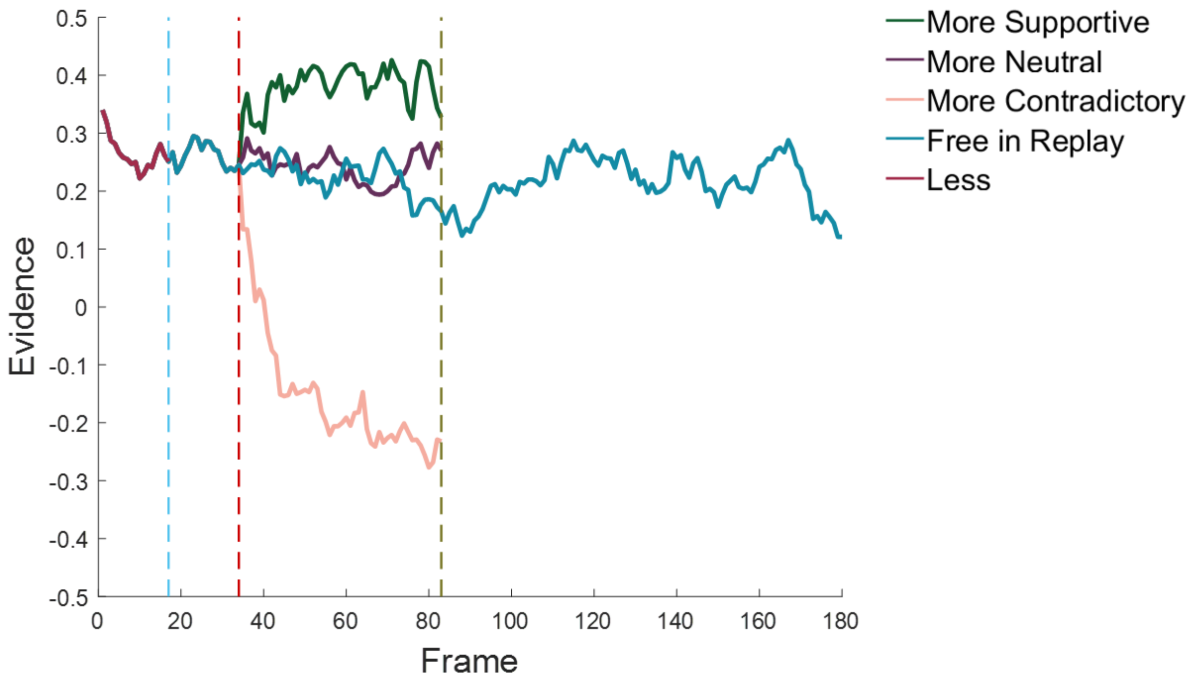


Figure 4: **Sequence of Evidence Frame by Frame (for 180 Frames).** The Replay Task had 5 conditions, and depending on the condition, the total evidence differed. The graph represents the evidence sequence during the first 180 Frames of each of the 5 conditions for one particular pre-generated trial. Horizontal solid lines represent the evidence in a particular frame, and vertical dashed lines represent a fixed number of frames where the evidence sequence terminated or changed. In the Less Condition (dark blue solid line), the evidence was presented up to the Less Frame (vertical light blue dashed line), which is the half of the decision frame calculated for this particular trial based on the responses given in the Free Task. In the Free in Replay condition (solid pink line), the exact same evidence that was presented in the Free Task for that trial was presented up to the response. In the More Conditions, the same evidence was presented up to the Decision Frame (vertical red dashed line), and starting at this decision frame, the evidence was increased for the More Supportive Condition (solid cream line), kept stable for the More Neutral Condition (solid green line) and decreased for the More Contradictory Condition (solid black line). The total evidence was presented up to the more frame (vertical light green dashed line) calculated for this particular trial based on the responses given in the Free Task.

Thanks to the constraint in the Markov chain for the proportion of Gabor patches that can change status over time with the constraint that only a fraction (0.1) can change, the evolution of evidence before and after the decision time in the More Contradictory and More Supportive conditions was presented gradually. This gradual change ensures that despite the huge evidence change between the starting and ending points of the post-decision evidence, the transformation was smooth and not too salient.

To be used in the analysis, all the initial responses that were given in the More conditions were also saved—recording these initial responses allowed us to compare the initial responses and final responses in conditions where observers received extra evidence after their default decision time. This recording also enabled us to see whether the observer gave their responses in the More conditions in the same way they did in the Free Task and in the Free in Replay condition.

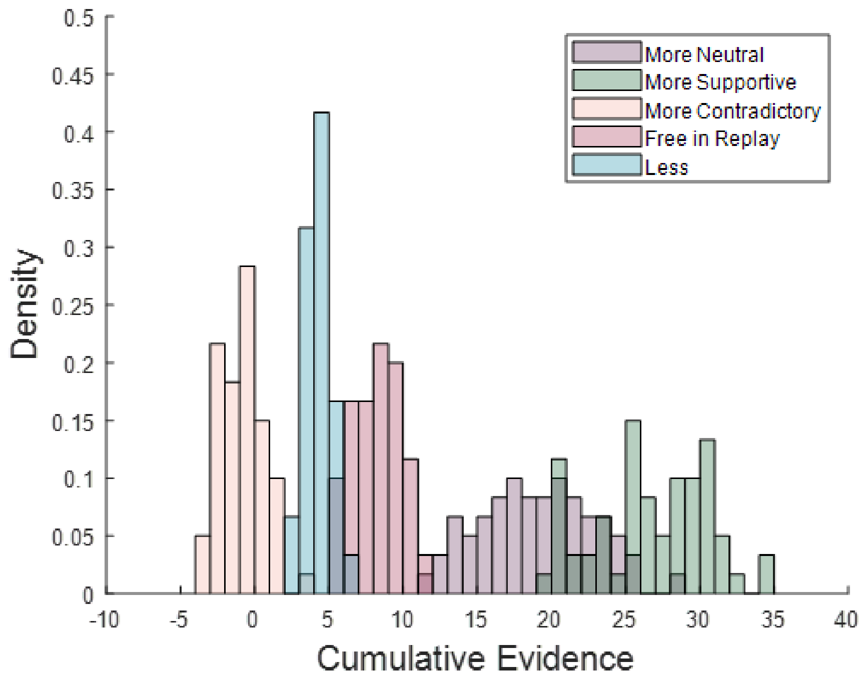


Figure 5: **Cumulative Evidence Distributions in the Replay Task.** The distribution of cumulative evidence in all trials presented for all five conditions for a hypothetical observer that has on average 500 ms of decision time. The blue distribution represents the cumulative evidence for the Less condition (that was only available until the first half of the decision time), the red distribution represents the cumulative evidence in the Free-in-Replay condition (that was available until the observer gave their responses), and three remaining colors represent the cumulative evidence in each of the More conditions that were available until the end of the fixed time. The cumulative evidence in the More Contradictory condition was less than in the other conditions, and in half of the More Contradictory conditions, the cumulative evidence at the end of the trial caused a direction change compared to the Free-in-Replay condition in which the same evidence was presented until the first response was received.

In the Replay Task session, the pre-generated 60 trials were repeated 3 times for the Less condition, 3 times for the Free in Replay condition, and 2 times for each of the More conditions. There were a total of 720 trials (60 pre-generated trials x 12 repetitions) that were presented in random order over 20 blocks of 36 trials.

All trials were followed by a confidence rating concerning the response given last. The confidence rating was on a scale of 1-4 (1 indicating the minimum confidence level and 4 indicating the maximum confidence level). Participants were instructed to give their confidence ratings as 1 if they felt completely unsure about their response, 4 if they felt they were 100% correct on their response, and to give 2 and 3 in between these two ends.

2.5. Analysis

2.5.1. Behavioral Analysis

First, for all the experimental conditions and the two separate sessions, the raw data of each participant was used to calculate the proportion correct, sensitivity (d'), meta-sensitivity (meta d'), and the median reaction time of the responses given to the 10 repetitions of the same trials.

To see whether the observer gave their responses in the same way across conditions and across sessions, the reaction times were compared in the Free Response Session, in the Free-in-Replay condition, and in the initial responses given to the More conditions by applying non-parametric within-subject statistics.

The main analysis was conducted on the Replay Task to examine the effect of extra post-decisional evidence that contradicted the pre-decisional evidence (More Contradictory Condition) on performance and on confidence. This effect was examined first by comparing the number of response changes from the initial and last responses between all more conditions. This comparison was statistically tested by conducting a Kruskal-Wallis followed by Wilcoxon sign rank tests between the proportion of response change trials in the trials in which there were at least two responses. Then to examine how the evidence accumulated up to the initial response predicts the initial response and how the evidence accumulated until the last response predicts the last response, we used a binomial Generalized Linear Model (GLM) with a probit link function. Beta weights for initial and final responses were then extracted and input into a repeated measure ANOVA with the beta weights predicting initial and final responses as within-subjects factors and 3 different More conditions as between-subjects factors.

The dependency of the confidence ratings on the presented evidence up to the final response was examined by conducting a GLM analysis with an identity link function on the confidence ratings separately for each more conditions and for the free in replay condition. As the choices, an ANOVA test was conducted on the beta weights obtained from the GLM analysis.

The average proportion correct, sensitivity (d'), metacognitive sensitivity (meta-d'), and meta efficiency (meta-d' / d') across conditions were also presented in the Results.

When more than one statistical test was performed per hypothesis and the obtained p-values were less than 0.05, these p-values are reported with a Bonferroni correction, while non-significant p-values are given uncorrected.

3. RESULTS

3.1. Free Task

We first examined how observers' decisions changed according to the presented evidence in the Free task. All the trials in the Free Task were first sorted according to the mean evidence that was presented up to the response. The sorted trials were then divided into three groups representing three difficulty levels (the initial six groups of difficulty levels, which were grouped according to the pro-signal fraction range that was randomly chosen from 6 different groups were reduced to 3 difficulty levels). The mean evidence presented up to the decision in the difficulty level 1, difficulty level 2, and difficulty level 3 averaged across all the participants were 0.18, 0.23, and 0.28, respectively.

In each difficulty level, participants scored an average proportion correct [95% between-subjects CI] of 0.77 [±0.049], 0.83 [±0.048], and 0.87 [±0.040]. These proportions correspond to a Type-I sensitivity (d') of 1.6 [±0.04], 2.1 [±0.04], and 2.5 [±0.04] for each group of difficulty levels. We found a significant increase in d' between the 3 difficulty levels by using a Wilcoxon sign rank test: Z (Difficulty Level 1 vs Difficulty Level 2) = 3.1, p_{bonf*2} = 0.0015; Z (Difficulty Level 2 vs Difficulty Level 3) = 3.4, p_{bonf*2} = 0.0113 with these p-values Bonferroni corrected for two comparisons (Figure 6.a). This result confirms that the task performance increased as the mean evidence presented until the response increased.

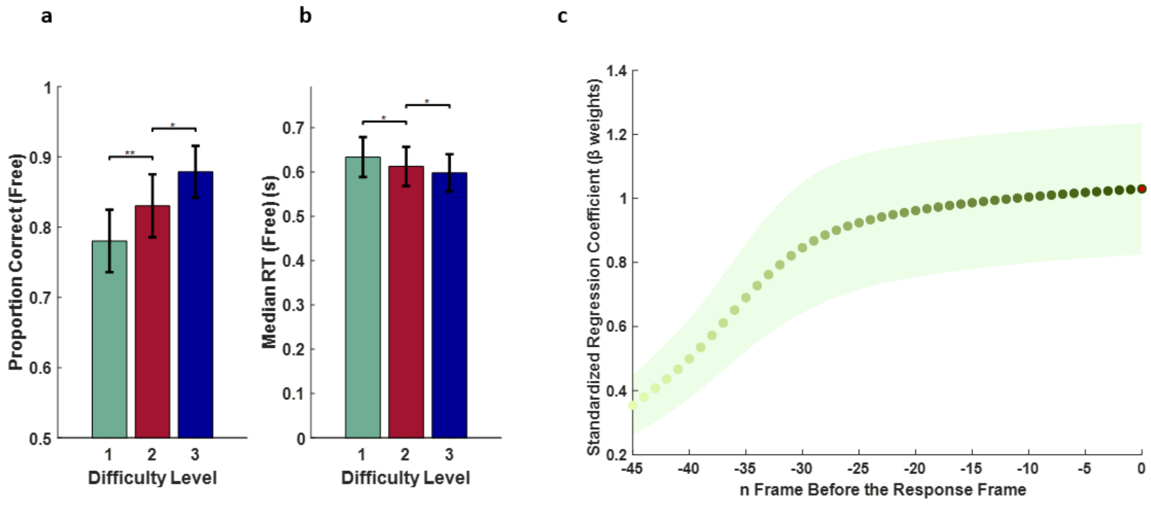


Figure 6: **Free Task Results a.** Average proportion correct in three difficulty levels in Free Task. Error bars show between-subjects confidence intervals n = 18. **b.** Median Reaction Time in Three Difficulty Levels in Free Task Error bars show between-subjects confidence intervals N=20. Both in plot **a** and plot **b** stars indicate a significant difference between difficulty levels with p <0.05:*, p <0.01:**, p <0.001:*** **c.** Comparison of the standardized correlation coefficient (beta weights) between the presented evidence and the responses within the trials on different time courses. The red dot on the right represents the standardized correlation coefficient between the response and the cumulative evidence up to the response frame, and the green dots represent the standardized correlation coefficient between the response and the cumulative evidence up to n frames before the response frame (From light green to dark green, the frame approaches the real response frame). Green shaded area show between-subjects confidence intervals N = 20.

Participants also committed to their decision earlier on less difficult trials (as the mean evidence presented up to the decision time got larger). The averaged median of the normalized reaction time data (in seconds) are 0.45 [±0.7], 0.50 [±0.7], and 0.53 [±0.7]; Z (Difficulty Level 1 vs Difficulty Level 2) = 3.29, pbonf*2 = 0.012, Z (Difficulty Level 2 vs Difficulty Level 3) = 2.46, pbonf*2= 0.0137; Figure 6.b).

To examine the time course of correlation between the accumulated evidence for different responses, we performed a general linear regression (assuming a binomial distribution) within each trial separately on the 45 frames leading to the response, including the response frame. The correlation between the proportion of responses in a trial and the (z-scored) cumulative evidence up to the specific frame was fitted by binomial GLM with a probit link function (the subject-level GLM fits for every 45 frames are available in the Supplementary Fig 1.). The slope of the GLM fit averaged across all participants separately for every 45 subsequent frames was then plotted to show that the dependency of the Type-I response in the Free Task to the accumulated evidence decreases as the frame moves away from the response frame (Figure 6. c). This increasing slope suggests observers were integrating all the evidence presented up until their response.

3.2. Replay Task

In order to examine whether the reaction time of observers was similar when they encountered the same trials in different conditions, we first compared the reaction times of the Free Task with the reaction time in the Free in Replay condition and with the reaction time of the initial responses given in the more conditions when there was more than one response. A Kruskal-Wallis Test was conducted to examine the differences in the median of the normalized reaction time data between the conditions. No significant differences (X^2 (2, 57) = 1.03, p = 0.59) was found among the three conditions (Free Task, Free in Replay Condition, Initial responses in the More conditions). Likewise, none of the subject-level data showed a significant difference in the reaction time in these three conditions (for the subject-level comparison of the reaction time, see Supplementary Fig 2).

We then compared performance in the last response of the 5 different conditions of the Replay Task with the performance in the Free Task. Before calculating the proportion correct, the correct response (only) in the trials where the cumulative evidence at the end of the trial is negative (see Figure 5 for the distribution), the correct response defined at the beginning of the trial was reversed at the last response. In this way, the correct response always stands for the cumulative evidence up to the response. Note that this reversion of the correct response was only performed for the trials in the More Contradictory condition and only when the cumulative evidence was negative. Performance in the Free in Replay condition was on par with the performance in the Free task (mean proportion correct = 0.77 [±0.05]; Z (Free in Replay vs. Free d') = 0.89, p = 0.3703). Together with the comparison of reaction times reported above, these results suggest that observers were not substantially changing their decision-making strategies across experimental sessions and testing contexts.

In the Less condition, halving the required stimulus presentation time for each observer had no significant effect on the performance (mean proportion correct = 0.79 [±0.05]; Z (Less vs. Free d') = 0.97, p = 0.3317). Likewise, in the More Neutral condition the extra evidence presented after the decision time did not significantly affect the performance within-subjects (mean proportion correct = 0.74 [±0.04]; Z (More Neutral vs. Free d') = 2.39, p_{bonf}*5 = 0.0675). In the More Contradictory condition, the contradictory evidence presented after the decision time leads to a substantial decrease in performance within-subjects (More

Contradictory d' = -0.06; Free d' = 1.83; Z= 3.92, pbonf*5 = 0.0004), but in the More Supportive condition the supportive evidence presented after the decision time did not significantly improve the performance within-subjects (mean proportion correct = 0.81 [±0.05]; Z (More Supportive vs. Free d') = 1.41, p = 0.1560) (Figure 7.a).

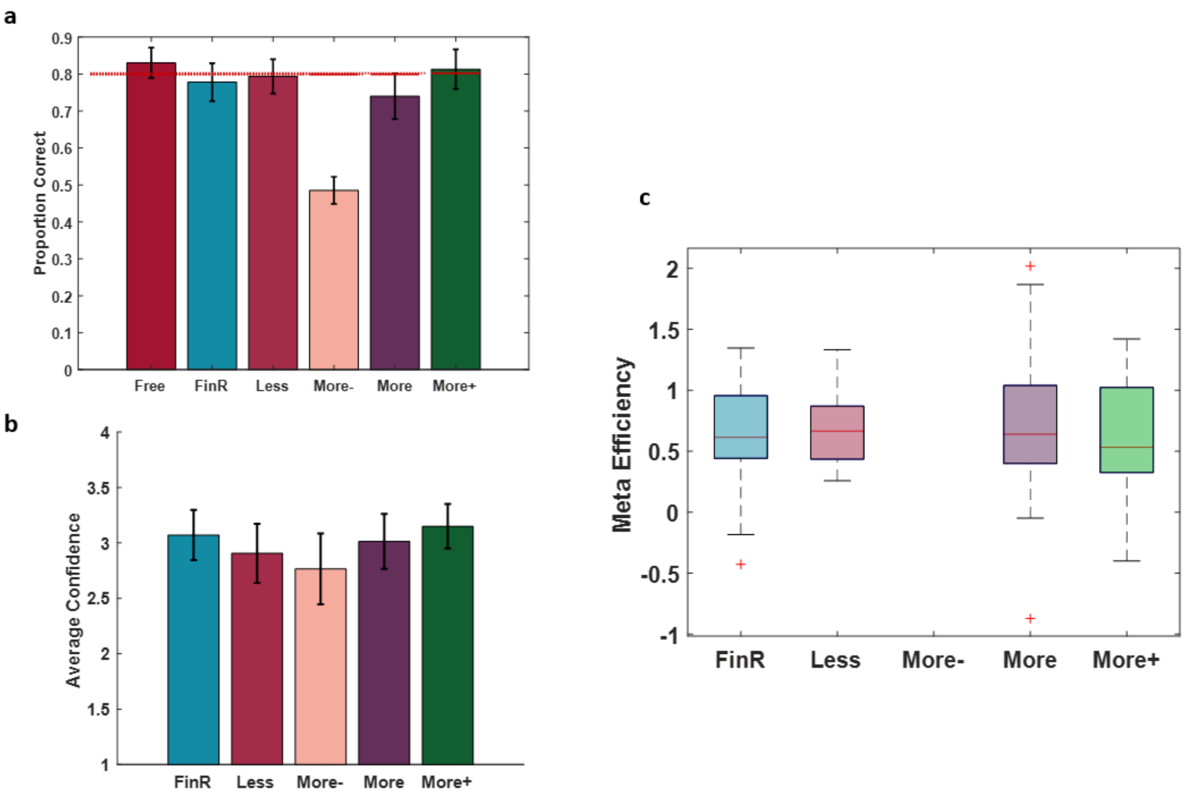


Figure 7 **Replay Task Results a.** Average proportion correct in Free Task and in the five conditions of Replay Task (for three more conditions, the bars represent the average proportion correct for the last responses given). The horizontal red lines represent the averaged proportion correct in the initial responses for the three more conditions. **b.** Average confidence ratings in the five conditions of Replay Task. **c.** Average meta-efficiency in the five conditions of Replay Task. For the More Contradictory condition, since the proportion correct is close to 0.5, the d' which is close to zero was not meaningful in our case to report, so we left it blank in the figure. In all plots, error bars show 95% between-subjects confidence intervals n = 20. Note that starting from this figure, more conditions are indicated in the figure with their abbreviated names (More Supportive Condition= More+, More Neutral Condition= More, and More Contradictory Condition= More-).

As shown in Figure 7.b, while the average confidence in the More Neutral condition (3.02) is on par with the averaged confidence in the Free Task (3.07), it decreased slightly in the More Contradictory condition (2.76) and increased slightly in the More Supportive condition (3.15). As a usual practice in metacognitive research, the confidence ratings given in the Replay Task were used to compute Type-II sensitivity (meta-d') for each observer in the 4 different conditions. To be able to take the variable Type-I sensitivity into account, we computed metacognitive efficiency by dividing the meta-d' by d'. The average Type-II efficiency was 0.62 [±0.202] in the Free in Replay Condition, 0.66 [±0.126] in the Less Condition, 0.73 [±0.292] in the More Neutral Condition, and 0.59 [±0.218] in the More Supportive Condition (see Figure 7-c).

The true direction at the beginning and at the end of the trial changed on average in 70% of the trials in the More Contradictory condition (for a detailed explanation, see Methods). Hence, the correct response within the trial changed in the More Contradictory condition, making the proportion correct, d', meta-d', and the meta efficiency ill-defined for this condition. Thus, comparing the More Contradictory condition with other conditions by referencing the d' and meta-d' was not useful for testing our main hypothesis (For the d' and meta-d' comparison between the conditions, see Supplementary Fig 3).

3.2.1. Effect of Extra Contradictory Evidence Presented After the Decision on the Type-I Responses

To test our first hypothesis and to examine the effect of extra contradictory evidence presented after the decision on the Type-I responses, we focused on the More trials. If the estimated decision time was calculated close to the real decision time of the subjects, we assumed that observers must have already given an initial response before the end of the more trials. In this manner, these more conditions are ideal for comparing initial and last response conditions as they allow us to see the effect of extra evidence (with different qualities) within the same trial. Thus, we first calculated in each of three More conditions the proportion of trials in which subjects gave at least two responses (on average, subjects gave more than one response in 0.94 [±0.0082], 0.93 [±0.0095], and 0.94 [±0.0091] of the trials in the More Contradictory, More Neutral and More Supportive conditions respectively).

Within these proportions, we divided more than one response trial into two categories: either the accumulated evidence until the initial response and the accumulated evidence until the last response was corresponding to the same direction (E1=E2) or to different directions (E1≠E2). Within these two main groups, trials were also divided into 4 subgroups. The first group holds the trials in which both the initial and last responses were given according to the accumulated evidence up to the response points (R1=E1 & R2=E2). The second group contains the trials in which only the initial response was given according to the cumulative evidence up to the initial response, but the last response did not reflect the cumulative evidence up to the last response (R1=E1 & R2≠E2). The third group contains the trials in which only the final response was given according to the cumulative evidence up to the last response (R1≠E1 & R2=E2), and the fourth group contains trials in which neither the initial nor the final response reflects the accumulated evidence up to the response points (R1≠E1 & R2≠E2) (Figure 8). (See Supplementary Fig 4 for the averaged confidence in all 3 more conditions for these eight different cases.)

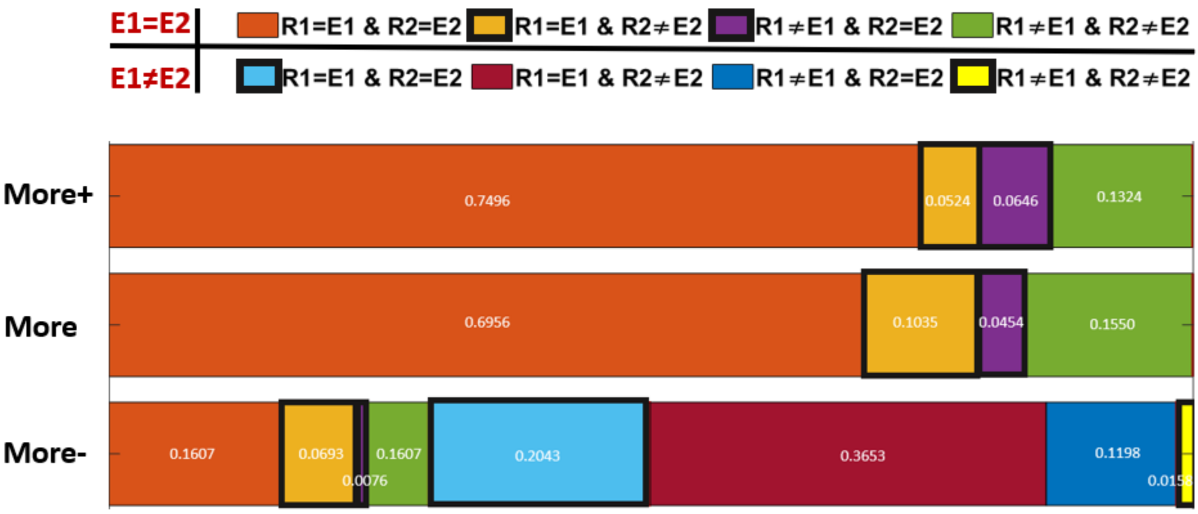


Figure 8: **Proportion of trials in all 3 more conditions for 8 different cases**. Orange, light orange, pink and green areas represent the proportion of trials in which the cumulative evidence presented up to the initial and last responses stand in the same direction. The light blue, red, dark blue, and yellow areas represent the proportion of trials in which the cumulative evidence presented up to the initial and last responses stand for the different directions. Note that for the More Supportive (More+) and More Neutral (More) conditions, there was no E1≠E2 case (no correct response reversal), so it was not shown in the figure. The black thick outlined areas stand for the cases where the response changed.

It is important to consider the cases when observers changed their response. This occurred when (E1=E2) and (R1≠E1 & R2=E2) or (R1=E1 & R2≠E2) and when (E1≠E2) and (R1=E1 & R2=E2) or (R1≠E1 & R2≠E2). Hence, for all three More conditions, these proportions of trials in which the initial responses and final responses were different (response changed) were extracted. A Kruskal-Wallis test showed that there was a statistically significant difference in the proportion of response change between the More conditions (X^2 (2, N=20) = 12.83, $p < 0.01$) , with a mean proportion response change of 0.30 for More Contradictory, 0.15 for More Neutral and 0.12 for More Supportive conditions. Post hoc analysis of the mean differences by paired Wilcoxon sign rank test showed that only the More Contradictory condition significantly differed from More Neutral; Z (More Contradictory vs More Neutral) = 3.5, $p_{bonf}*2 = 0.001$ and from More Supportive conditions; Z (More Contradictory vs More

Supportive) = 3.3, *pbonf**2 = 0.001 with these p-values Bonferroni corrected for two comparisons (*pbonf**2 < 0.05, corrected) (Figure 9). This significantly greater response change in the More Contradictory condition tells us that the observers were not completely blind to the extra evidence that they received after their initial response, and if they received enough contradictory post-decisional evidence, they used this information to change their minds. This first part of the results makes us reject the null hypothesis, suggesting that the post-decisional evidence has no effect on the Type-I and Type-II responses.

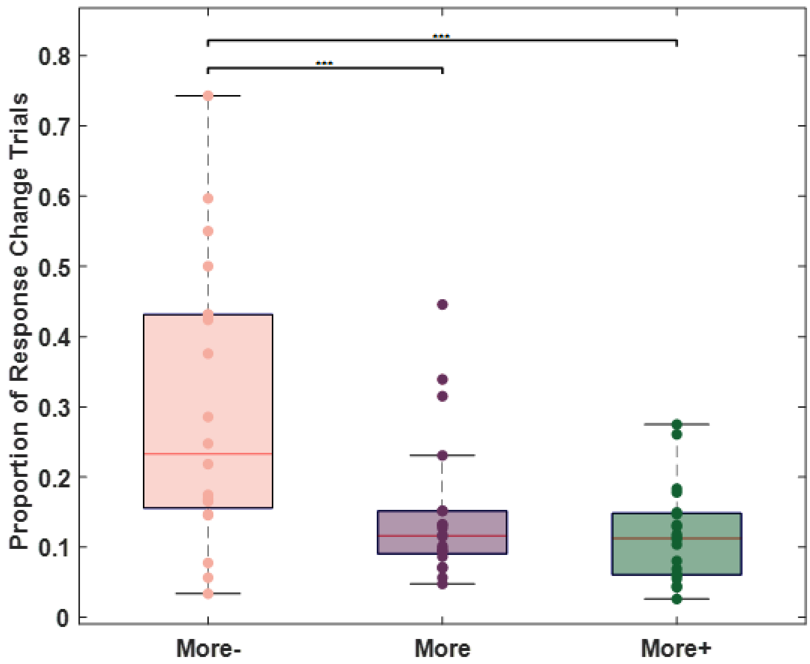


Figure 9: **Response Change Trials.** Proportion of trials in which the response changed in all 3 more conditions. Error bars show 95% between-subjects confidence intervals n = 20. Stars indicate significant difference in the number of response change trials between conditions with p <0.05:*, p <0.01:**, p <0.001:***.

We then tested a side hypothesis derived from these results: Do observers use all the evidence presented after their initial response in the same way that they did in the pre-decisional stage? To examine the dependency between the presented evidence up to the initial and final responses and the responses given at these times, a binomial GLM with a probit link function was performed separately for the initial and final responses. Beta weights were then input into a repeated measures ANOVA with the beta weights predicting initial and final responses as within-subjects factors and 3 different More conditions as between-subjects factors. The ANOVA demonstrated a significant interaction of initial and final responses [$F(2, 38) = 36.23$, $p < 0.01$]. Post hoc paired t-tests showed that, in the More Contradictory condition, the beta weights were significantly greater in the initial responses compared with the last responses [t(19) = 6.33, pbonf*3< 0.01; while in the More Supportive condition, the beta weights were significantly greater in the final responses compared with the initial responses [t(19) = 2.91, pbonf*3 < 0.0090] (Figure 10). These results, together with response changing results reported above, suggest that although the observers used the post-decisional evidence, unlike the pre-decisional stage, they were not accumulating all the evidence presented after their initial responses.

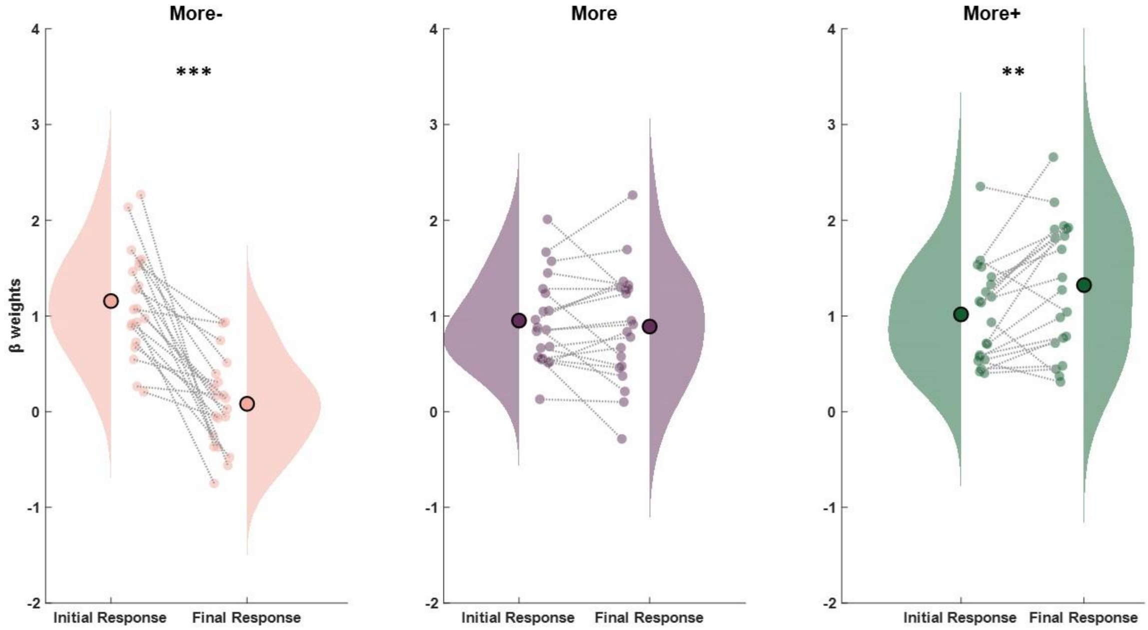


Figure 11: **Beta weights comparison for the Type-I responses.** Beta weights extracted from the binomial GLM analysis for each 3 more conditions. Left distributions subtend the standardized regression weights on how much the initial response depends on the evidence up to the initial response in the given condition. Right distributions subtend the standardized regression weights on how much the final response depends on the evidence up to the final response in the given condition. Dashed lines represent the difference between the initial and final response beta weights within subjects. Stars indicate significant difference between initial and final beta weights with p <0.05:*, p <0.01:**, and p <0.001:***.

3.2.2. Effect of Extra Contradictory Evidence Presented After the Decision on the Type-II Responses

To test the effect of post-decisional evidence on the Type-II responses, we examined the dependency between the standardized presented evidence up to the final response and the standardized confidence rating for this response by conducting a GLM. For this GLM analysis that was conducted separately for each of the more conditions and for the Free in Replay condition, we assumed a normal distribution with an identity link function on the confidence ratings. ANOVA tests were first performed on the extracted beta weights of three more conditions to see the effect of extra contradictory evidence on Type-II responses. The ANOVA revealed that there was a significant difference in confidence beta weights between more conditions [$F(2,38) = 7.72$, $p = 0.0015$].

The repeated measure ANOVA with the confidence beta weights in the Free in Replay condition and in the three different More conditions as within-subjects factors and 3 different More conditions as between-subjects factors was then performed to see how different the beta weights in each more condition were from the beta weights in the Free in Replay condition. The ANOVA demonstrated a significant interaction between the distance from the confidence beta weights in the Free in Replay condition and the more conditions [$F(2, 38) = 7.72$, $p = 0.0015$]. Post hoc paired t-tests showed that the beta weights for the More Contradictory condition were significantly lower compared with beta weights in the Free in Replay condition [$t (19) = 2.40$, $p_{bonf}*3 = 0.0399$]. But the beta weights in the Free in Replay condition do not significantly differ from the beta weights in the More Neutral condition [$t (19) = 1.66$, $p = 0.1122$] and from the More Supportive Condition [$t (19) = 1.39$, $p = 0.1780$] (Figure 11). This result shows us that the contradictory evidence presented after the initial response reduces the power of the cumulative evidence up to the last response to predict the confidence ratings given by the observer.

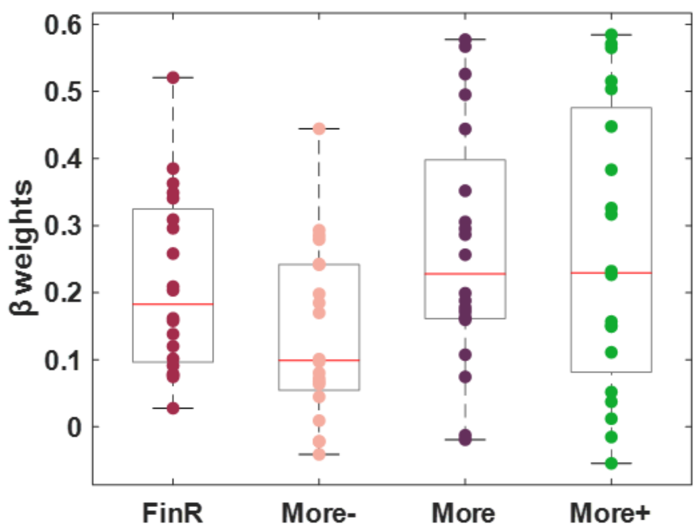


Figure 11: **Beta weights comparison for the Type-II responses**. Beta weights for confidence ratings were extracted from the normal GLM analysis for each 3 more conditions. Each point represents the beta weight for one observer in the given condition.

A GLM was also applied to see the dependency of the presented evidence up to the final response and the given confidence rating for the trials with response change and for the trials with no response change separately (see Supplementary Figure 5). ANOVA tests were again applied separately to these two groups. The ANOVA for the response change group revealed that there was no significant difference in confidence beta weights between conditions [$F(2,57) = 0.31$, $p = 0.73$], while the ANOVA for the no response change group demonstrated a significant difference in confidence beta weights between conditions [$F(2,57) = 8.44$, pbonf*3 < 0.001]. Post hoc paired t-tests showed that, for the no response change group in the More Contradictory condition, the beta weights were significantly lower compared with beta weights in More Neutral condition [$t(19) = 4.53$, pbonf*2 < 0.001] and compared with beta weights in More Supportive Condition [$t(19) = 3.95$, pbonf*2 < 0.001] with these p-values Bonferroni corrected for two comparisons. This result demonstrates that the decrease in the confidence beta weights of the More contradictory conditions reported in the previous result is mostly due to trials in which the contradictory evidence was not enough to change the initial response.

4. DISCUSSION

The current project was designed to investigate the effect of contradictory post-decisional evidence on perceptual and confidence decisions separately. For this purpose, we exposed our participants to a lot of supportive or disconfirming sensory evidence after their estimated decision time. By comparing Type-I decisions in three different conditions with variable quality of the post-decisional evidence, we found that observers were more likely to change their already committed decisions when they received enough contradictory post-decisional evidence compared to the cases in which they received post-decisional evidence confirming the pre-decision period. These results imply that observers continued to accumulate evidence in the post-decisional stage, which leads them to change their minds by integrating this additional contradictory evidence they encounter after their initial decision. The second key finding of the present study is that the extra post-decisional evidence, which contradicts the pre-decision stage, reduces the sensitivity of the metacognitive performance to the accumulated evidence compared to the condition in which no post-decisional evidence was presented. Put together; the results are consistent more with our second alternative hypothesis, which indicates that the contradictory evidence presented after the initial decision has an effect on the evidence accumulation process both for the perceptual and confidence decisions.

With these key results, our study also provided new data pointing out the continued evidence accumulation in the post-decisional period. This continuous accumulation of the evidence during the post-decisional period makes us reject the assumption of absorbing boundaries, which states that the observer ignores any additional evidence received after their decision. This rejection brings an additional question: What is the accumulation strategy observers used in this post-decisional stage? Or in other words, when observers are not free to control the amount of evidence they accumulate, do they accumulate all the provided evidence they were presented (including any post-decisional ones) and give their final response accordingly, or do they continue to accumulate evidence in the post-decisional period only if the extra evidence contradicts with their decision? The former extreme case implies that the decision boundary merely triggers the response but does not stop or alter evidence accumulation, and the latter case is rather close to the assumptions of reflexive boundaries.

In order to examine if observers continue to accumulate the post-decisional evidence in the same way that they accumulated the pre-decisional evidence, a GLM analysis was conducted on the initial and final responses. Suppose the same accumulation exists both in the pre-decision and post-decisional stages. In that case, we should expect the same beta weights for the initial and final responses in the GLM analysis. However, our analysis revealed that the evidence accumulated up to the initial response could predict the initial response significantly better than the evidence accumulated until the final response predicted the final response in the More Contradictory condition. These smaller beta weights on the final responses enabled us to reject the extreme hypothesis, which states that observers accumulate all the evidence presented after their initial decision commitment. Instead, it implies that observers accumulate at least some of the evidence coming after their initial decision. These findings leave us with the reflexive boundaries as a better explanation of the behavioral results. Still, the ability of each boundary model to explain the behavioral data needs to be compared by appropriate computational modeling. Together, these findings suggest that although the contradictory evidence presented after the initial decision was enough to change observers' minds in some trials, still this evidence was not accumulated in the same way as that accumulated in the pre-decision period.

For the confidence judgments, our findings are in accordance with two-stage dynamic models (Moran et al., 2015; Pleskac & Busemeyer, 2010), which declare the idea that the construction of the confidence is based on the additional evidence gathered after the decision.

Yet, the results also reported that the extra supportive post-decisional evidence did not lead to a significant enhancement in the beta weights for the confidence judgments. So, although the post-decisional information is not completely ignored in the construction of the confidence decisions, it does not mean that the confidence uses all the available evidence. Otherwise, the evidence up to the final response in the More Supportive condition should be a better predictor of confidence than the evidence up to the final response in the Free in Replay condition. Thus, the results of the present study also add a new dimension that the evidence exposed after the decision is not blindly accumulated as in the 2DSD model. Instead, the confidence ratings, or more generally, the Type-II judgments, are reweighted depending on whether the post-decisional evidence is consistent or not with the initial decision. Still, to precisely test how the evidence gathered up to the initial and final response predicted the confidence rating given for the initial and last decision, our experimental design would have needed to be updated to also record the confidence level for the initial choices. Recording the confidence in the initial responses would also allow us to examine the role of post-decisional accumulation on confidence judgment changing.

We also wanted to examine the effect of choice-changing on metacognitive performance. As shown in the results, there was no difference between more conditions in terms of the confidence beta weights for the response change groups. This result shows us that the mind-changing based on the additional disconfirming evidence is not an impairing factor for the confidence. Instead, findings suggest that when observers accumulate enough contradictory evidence to change their initial response, they can still give a well-calibrated confidence judgment for their new choice based on this post-decisional evidence. So, the post-decisional evidence that leads to a mind-changing synchronously affects the confidence judgments. However, the contradictory post-decisional evidence that does not lead to a mind change reduces the sensitivity of the Type-II judgments to the presented evidence. So, in some trials, the contradictory evidence gathered after the initial decision was not enough for observers to change their minds but still sufficient to reduce the sensitivity of the confidence judgments to the total accumulated evidence. This clearly shows us that the confidence judgment is constructed by the additional information that was not used in the perceptual decision. These results also sign the potential dissociation between the evidence used for decision making and confidence, or more generally for the Type-I and Type-II judgment that is well suited to the results of a recent study showing these distinct processes at the neural level (Balsdon et al., 2021).

The novelty of the present study derived from the gradual manipulation of the stimulus within the trial, which enabled the representation of all the evidence fluently until the end of the trial without causing any interruption in the accumulation. Additionally, by implementing the Free trials in the Replay condition, which enabled observers to terminate the evidence themselves, the observers were forced to accumulate the evidence in the Replay Task in the same way that they accumulated in the Free Task. Very recently, it was found that metacognitive performance decreases when the observer is free to select when to respond (Free Task) compared to making a decision based on a set of evidence predefined by the experimenter (Replay Task), even if the amount of presented evidence is equal in these two designs (Rosenbaum et al., 2022). Hence, the implementation of Free trials in the Replay task in the present study ruled out this potential confounding effect on the metacognitive performance that different accumulation strategies would cause. The last and yet very important strength of the present study comes from the design that enabled us to estimate each observer's individual decision time. This gave us the opportunity to carefully present the post-decisional evidence after the initial decision was already made.

Still, one of our findings also suggests that people do not always change their minds in the presence of extra contradictory evidence, as established by the proportion of the trials in which the observers stayed in their initial response being significantly larger than the trials in which they changed their responses in the More Contradictory condition. The GLM analysis results already established the non-ideal accumulation of the post-decisional evidence. Now we may ask what else may drive this divergence from the ideal information accumulation in the post-decisional stage. Or in other words, is there a systematic difference between the trials where the response changed and not changed in the More contradictory condition.

This tendency on the observers to stick to their initial response even in the presence of strong disconfirming evidence can be explained by the well-known psychological phenomenon termed 'confirmation bias'. According to this bias, the decision-maker searches for and weighs evidence supporting their choice while underweighting those that disconfirm their belief (Nickerson, 1998). An essential facet of confirmation bias that has been well studied in recent years is its interactive relation with decision confidence. As already discussed in the introduction, confidence plays an important role in monitoring the decision (Bogacz et al., 2010; Folke et al., 2016). Thus, it is likely that confidence in the initial decision may guide the selective information accumulation after the choice commitment. Indeed, recent studies demonstrate this confirmatory evidence accumulation in the guidance of the confidence in the pre-decision period (Desender et al., 2019) and in the post-decisional stage (Rollwage et al., 2020; Kaanders et al., 2022). Hence, the experimental design we used can be developed in a way to also record the confidence ratings in the initial choices to see the effect of confidence on the selective gathering of post decisional evidence. For this, recent mind-changing studies (e.g., van den Berg et al., 2016) can be imitated by using a handle that enables the recording of the decision and the confidence ratings simultaneously.

Finally, the present study raises several outstanding questions that might be interesting to ask in further research: Whether the stopping rule for the evidence accumulation in the post-decisional stage is evidence-based or time-based. In the present study, neither the amount of the evidence nor the time of the post-decisional stage in the More conditions was manipulated systematically. Thus, this could be a potentially good step to examine further. Another question to investigate would be the relation between the post-decisional evidence and the mind changing in the confidence judgments. Therefore, the experimental paradigm we used in the present study can be developed in a way to motivate and answer future scientific questions.

References

Amano, K., Edwards, M., Badcock, D. R., Nishida, S. Y. (2009). Adaptive pooling of visual motion signals by the human visual system revealed with a novel multi-element stimulus. Journal of Vision, 9(3):4, 1-25.

Balsdon, T., Wyart, V., & Mamassian, P. (2020). Confidence controls perceptual evidence accumulation. Nature communications, 11(1), 1-11.

Balsdon, T., Mamassian, P., & Wyart, V. (2021). Separable neural signatures of confidence during perceptual decisions. *Elife*, *10*, e68491.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. Psychological review, 113(4), 700.

Boldt, A., Blundell, C., & De Martino, B. (2019). Confidence modulates exploration and exploitation in value-based learning. Neuroscience of consciousness, 2019(1), niz004.

Brainard, D. H. (1997). The psychophysics toolbox. Spatial Vision, 10(4), 433–436.

Carlebach, N., & Yeung, N. (2020). Subjective confidence acts as an internal cost-benefit factor when choosing between tasks. Journal of Experimental Psychology: Human Perception and Performance, 46(7), 729.

Clarke, F. R., Birdsall, T. G., & Tanner Jr, W. P. (1959). Two types of ROC curves and definitions of parameters. The Journal of the Acoustical Society of America, 31(5), 629-630.

Couchman, J. J., Miller, N. E., Zmuda, S. J., Feather, K., & Schwartzmeyer, T. (2016). The instinct fallacy: The metacognition of answering and revising during college exams. Metacognition and Learning, 11(2), 171-185.

Desender, K., Boldt, A., & Yeung, N. (2018). Subjective confidence predicts information seeking in decision making. Psychological science, 29(5), 761-778.

Desender, K., Murphy, P., Boldt, A., Verguts, T., & Yeung, N. (2019). A postdecisional neural marker of confidence predicts Information-Seeking in Decision-Making. *Journal of Neuroscience*, *39*(17), 3309-3319.

Desender, K., Donner, T. H., & Verguts, T. (2021). Dynamic expressions of confidence within an evidence accumulation framework. Cognition, 207, 104522.

Dotan, D., Meyniel, F., & Dehaene, S. (2018). On-line confidence monitoring during decision making. Cognition, 171, 112-121.

Drugowitsch, J., Wyart, V., Devauchelle, A. D., & Koechlin, E. (2016). Computational precision of mental inference as critical source of human choice suboptimality. Neuron, 92(6), 1398-1411.

Fetsch, C. R., Kiani, R., & Shadlen, M. N. (2014, January). Predicting the accuracy of a decision: a neural mechanism of confidence. In Cold Spring Harbor symposia on quantitative biology (Vol. 79, pp. 185-197). Cold Spring Harbor Laboratory Press.

Fleming, S. M., & Lau, H. C. (2014). How to measure meta- cognition. Frontiers in Human Neuroscience, 8, Article 443. https://doi.org/10.3389/fnhum.2014.00443

Fleming, S. M., Van Der Putten, E. J., & Daw, N. D. (2018). Neural mediators of changes of mind about perceptual decisions. Nature neuroscience, 21(4), 617-624.

Galvin, S. J., Podd, J. V., Drga, V., & Whitmore, J. (2003). Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. Psychonomic bulletin & review, 10(4), 843-876.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: a Brunswikian theory of confidence. Psychological review, 98(4), 506.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.*, *30*, 535-574.

Hanks, T. D., Kopec, C. D., Brunton, B. W., Duan, C. A., Erlich, J. C., & Brody, C. D. (2015). Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature*, *520*(7546), 220-223.

Holmes, W. R., Trueblood, J. S., & Heathcote, A. (2016). A new framework for modeling decisions about changing information: The Piecewise Linear Ballistic Accumulator model. Cognitive psychology, 85, 1-29.

Johnston, A., & Scarfe, P. (2013). The role of the harmonic vector average in motion integration. Frontiers in computational neuroscience, 7, 146.

Kaanders, P., Sepulveda, P., Folke, T., Ortoleva, P., & De Martino, B. (2022). Humans actively sample evidence to support prior beliefs. *Elife*, *11*, e71768.

Kepecs, A., & Mainen, Z. F. (2012). A computational framework for the study of confidence in humans and animals. Philosophical Transactions of the Royal Society B: Biological Sciences, 367(1594), 1322-1337.

Kiani, R., Corthell, L., & Shadlen, M. N. (2014). Choice certainty is informed by both evidence and decision time. Neuron, 84(6), 1329-1342.

Kiani, R., & Shadlen, M. N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. science, 324(5928), 759-764.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. Perception, 36(14), 1-16.

Kunimoto, C., Miller, J., & Pashler, H. (2001). Confidence and accuracy of near-threshold discrimination responses. Consciousness and cognition, 10(3), 294-340.

Mamassian, P. (2016). Visual confidence. Annual Review of Vision Science, 2, 459-481.

Meyniel, F., Schlunegger, D., & Dehaene, S. (2015). The sense of confidence during probabilistic learning: A normative account. PLoS computational biology, 11(6), e1004305.

Moran, R., Teodorescu, A. R., & Usher, M. (2015). Post choice information integration as a causal determinant of confidence: Novel data and a computational account. Cognitive psychology, 78, 99-147.

Murphy, P. R., Robertson, I. H., Harty, S., & O'Connell, R. G. (2015). Neural evidence accumulation persists after choice to inform metacognitive judgments. Elife, 4, e11946.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, *2*(2), 175-220.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spatial Vision, 10(4), 437–442.

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. Psychological review, 117(3), 864.

Rabbitt, P., & Rodgers, B. (1977). What does a man do after he makes an error? An analysis of response programming. *Quarterly Journal of Experimental Psychology*, *29*(4), 727-743.

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.

Resulaj, A., Kiani, R., Wolpert, D. M., & Shadlen, M. N. (2009). Changes of mind in decision-making. Nature, 461(7261), 263-266.

Rollwage, M., Loosen, A., Hauser, T. U., Moran, R., Dolan, R. J., & Fleming, S. M. (2020). Confidence drives a neural confirmation bias. Nature communications, 11(1), 1-11.

Rosenbaum, D., Glickman, M., Fleming, S. M., & Usher, M. (2022). The Cognition/Metacognition Trade-Off. *Psychological Science*, *33*(4), 613-628.

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

Stone, C., Mattingley, J. B., & Rangelov, D. (2022). On second thoughts: changes of mind in decision-making. Trends in cognitive sciences.

Takeuchi, T. (1998). Effect of contrast on the perception of moving multiple Gabor patterns. Vision Research, 38(20), 3069-3082.

Van den Berg, R., Zylberberg, A., Kiani, R., Shadlen, M. N., & Wolpert, D. M. (2016). Confidence is the bridge between multi-stage decisions. Current Biology, 26(23), 3157-3168.

van Den Berg, R., Anandalingam, K., Zylberberg, A., Kiani, R., Shadlen, M. N., & Wolpert, D. M. (2016). A common mechanism underlies changes of mind about decisions and confidence. Elife, 5, e12192.

Veenman, M. V., Wilhelm, P., & Beishuizen, J. J. (2004). The relation between intellectual and metacognitive skills from a developmental perspective. Learning and instruction, 14(1), 89-109.

Vickers, D. (2014). Decision processes in visual perception. Academic Press.

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. Philosophical Transactions of the Royal Society B: Biological Sciences, 367(1594), 1310-1321.

Yu, S., Pleskac, T. J., & Zeigenfuse, M. D. (2015). Dynamics of postdecisional processing of confidence. Journal of Experimental Psychology: General, 144(2), 489.

Zhang, J., Bogacz, R., & Holmes, P. (2009). A comparison of bounded diffusion models for choice in time controlled tasks. Journal of mathematical psychology, 53(4), 231-241.

Tables

Free Response Task & More Neutral Condition

Range Number	Pro-Signal Fraction Range	Noise Fraction Range	Anti-Signal Fraction Range
1	0.45-0.60	0.25	0.15-0.30
2	0.50-0.65	0.20	0.15-0.30
3	0.55-0.70	0.15	0.15-0.30
4	0.60-0.75	0.10	0.15-0.30
5	0.65-0.80	0.05	0.15-0.30
6	0.70-0.85	0	0.15-0.30

More Contradictory Condition

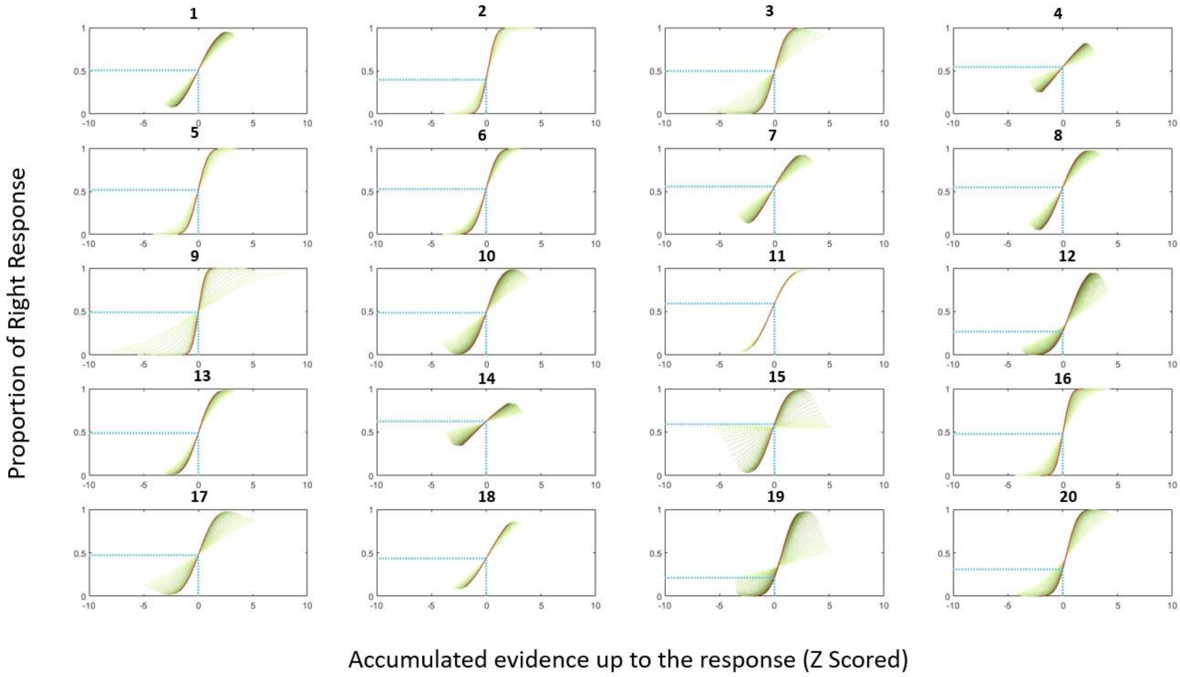
Range Number	Pro-Signal Fraction Range	Noise Fraction Range	Anti-Signal Fraction Range
1	0.15-0.30	0.25	0.45-0.60
2	0.15-0.30	0.20	0.50-0.65
3	0.15-0.30	0.15	0.55-0.70
4	0.15-0.30	0.10	0.60-0.75
5	0.15-0.30	0.05	0.65-0.80
6	0.15-0.30	0	0.70-0.85

More Supportive Condition

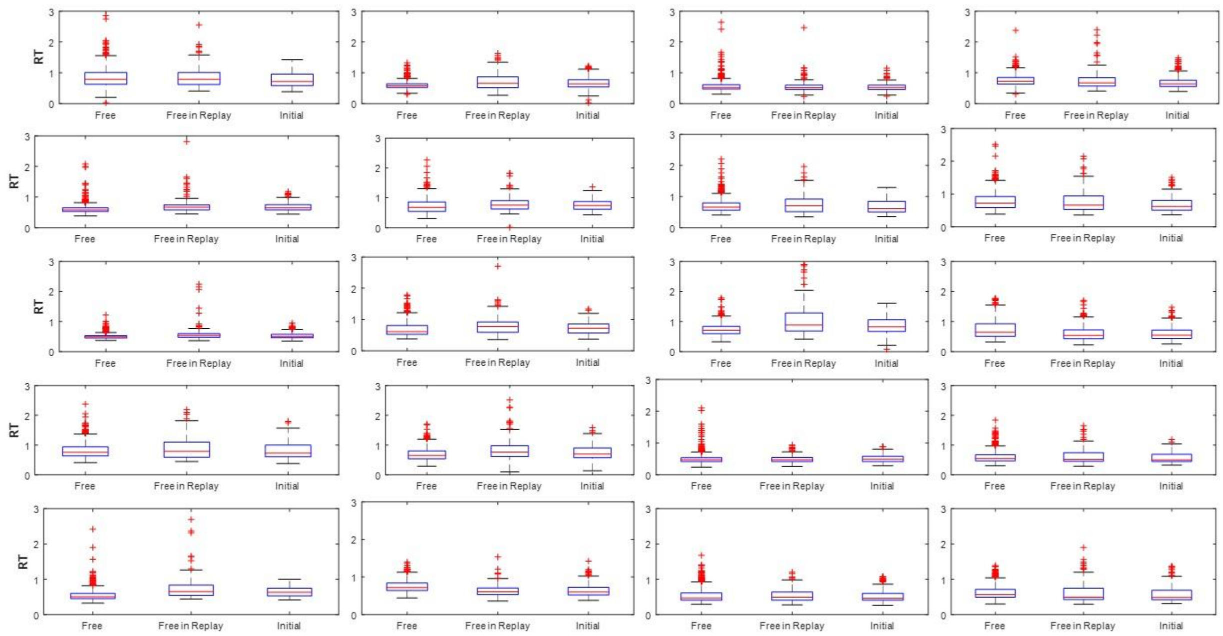
Range Number	Pro-Signal Fraction Range	Noise Fraction Range	Anti-Signal Fraction Range
1	0.60-0.75	0.25	0-0.15
2	0.65-0.80	0.20	0-0.15
3	0.70-0.85	0.15	0-0.15
4	0.75-0.90	0.10	0-0.15
5	0.80-0.95	0.05	0-0.15
6	0.85-1.00	0	0-0.15

Table 1: **Fraction Range Table:** Table for the Range of pro-signal, anti-signal, and noise for all six difficulty levels in the Free Response Task and in the post-decisional stage of the More Neutral (same as the Free Response Task), More Contradictory, and More Supportive conditions.

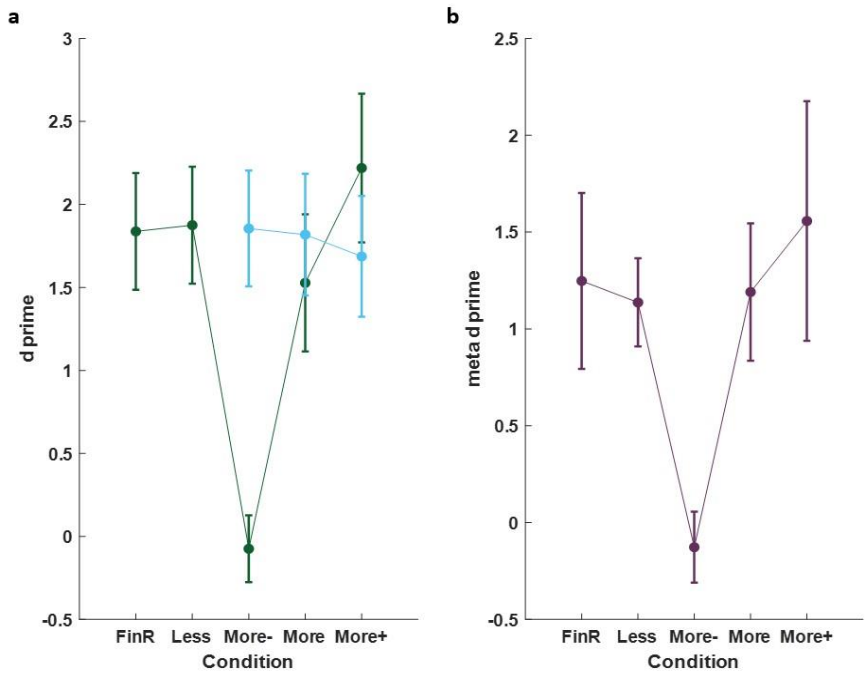
Supplementary Material



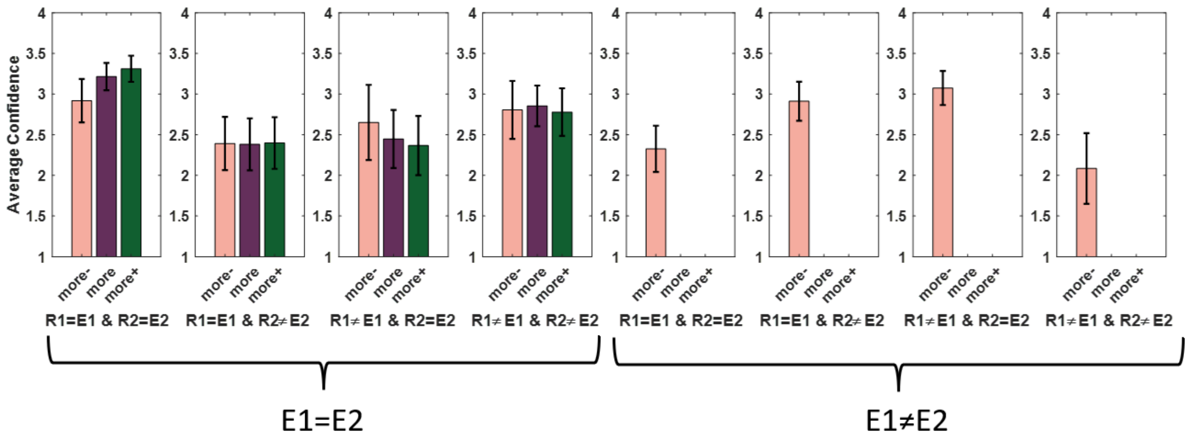
Supplementary Fig 1 The binomial Generalized linear models (GLM) with a probit link function fit in every 45 subsequent frames for each subject. The X-axis is the z-scored accumulated evidence up to the frame in question (negative values stand for the evidence in the leftward direction and positive values stand for the evidence in the rightward direction). Y-axis is the proportion of the rightward response. Red fits subtend to the GLM fit on the dependency of the responses to the evidence accumulated up to the response. From light green to dark green, the GLM fits approaches the real response frame. For all the subjects, the steeper slope was detected from the GLM fit in the real response frame (red fit). The closer the frame to the real response frame, the more the response is predictable by the evidence accumulated up to that point. The dashed blue line represents the proportion of right responses when there is equal number of evidence directing leftward and rightward motion (0 cumulative evidence).



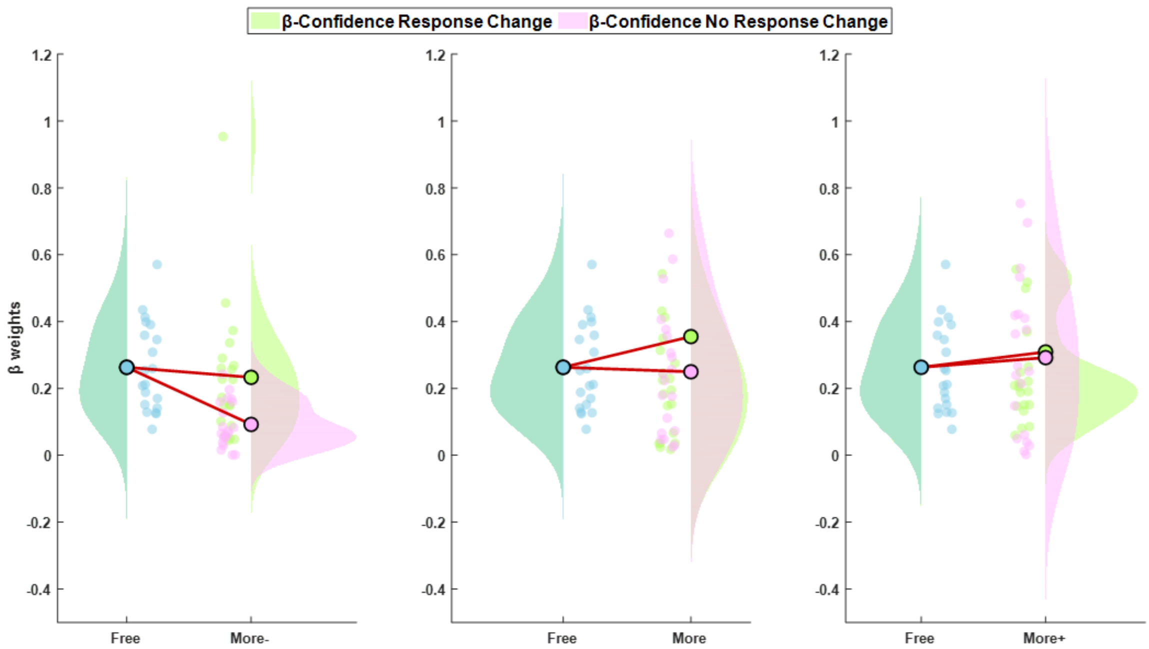
Supplementary Fig 2 Subject-level reaction time (in seconds) in the Free Task (leftmost), in the Free in Replay trials (middle), and in the initial responses in all three More conditions when there was more than one response (on the far right). Within the subject, there was no significant difference in reaction time in these three conditions, meaning that the reaction time data matched in all three conditions. Hence, when the participants saw the same trial, they committed their decision at about the same time interval.



Supplementary Fig 3 a. d' comparison between 5 different conditions in Replay Task. Green dots subtend the averaged d' for the final responses, and the blue dots represent the averaged d' for the initial responses in more conditions. **b.** meta-d' comparison between 5 different conditions in Replay Task. Error bars show 95% between-subjects confidence intervals n = 20.



Supplementary Fig 4 Averaged confidence in all 3 more conditions for 8 different cases. Error bars show 95% between-subjects confidence intervals n = 20.



Supplementary Fig 5 Beta weights for confidence ratings extracted from the GLM analysis for each 3 more conditions for the response change and not change trials separately.

Supplementary Note-1

Generation of a stimulus that has any desired fraction of pro-signal, anti-signal, and noise, and at the same time where the fraction of Gabors that change status (pro-signal, anti-signal, or noise) is controlled throughout the duration of the stimulus.

General idea: Create a three-state Markov chain, where the 3 states are {pro-signal, anti-signal, noise}. Determine a set of constraints for the transitions between the different states, and look for the transition probabilities that best match these constraints.

Details:

The stimulus is composed of N=330 drifting Gabors, where a fraction moves in one global direction (pro-signal), another fraction moves in the opposite direction (anti-signal), and the remaining Gabors move in random directions (noise). Let P_1, P_2, and P_3 be the probability that any Gabor is a pro-signal, anti-signal, or noise Gabor, respectively. By definition,

$$P_1 + P_2 + P_3 = 1 \; . \tag{1}$$

Let a_{ij} be the transition probability to go from state $\{i\}$ to state $\{j\}$. These transition probabilities define a 3x3 transition matrix U

$$U = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} . \tag{2}$$

To find the values of this transition matrix, we can write a set of constraints. The first 3 constraints come from the fact that the chain is closed (when we are in 1 state, there are only 3 options, either stay in that state, or move to one of the other 2 states)

$$\begin{cases} 1 = a_{11} + a_{12} + a_{13} \\ 1 = a_{21} + a_{22} + a_{23} \; . \\ 1 = a_{31} + a_{32} + a_{33} \end{cases} \tag{3}$$

The next 3 constraints come from the stationary regime (if we let the system go on forever, it should stabilize in some desired probabilities of the 3 states)

$$\begin{cases} P_1 = a_{11}P_1 + a_{21}P_2 + a_{31}P_3 \\ P_2 = a_{12}P_1 + a_{22}P_2 + a_{32}P_3 \; . \\ P_3 = a_{13}P_1 + a_{23}P_2 + a_{33}P_3 \end{cases} \tag{4}$$

And finally, we can add one constraint on the "lifetime" of the Gabor. This lifetime represents the probability τ that a Gabor will change its state at each time step

$$\tau = a_{11}P_1 + a_{22}P_2 + a_{33}P_3 \ . \tag{5}$$

We can write these 7 constraints in a matrix format,

$$\begin{cases} 1 = 1\,a_{11} + 1\,a_{12} + 1\,a_{13} + 0\,a_{21} + 0\,a_{22} + 0\,a_{23} + 0\,a_{31} + 0\,a_{32} + 0\,a_{33} \\ 1 = 0\,a_{11} + 0\,a_{12} + 0\,a_{13} + 1\,a_{21} + 1\,a_{22} + 1\,a_{23} + 0\,a_{31} + 0\,a_{32} + 0\,a_{33} \\ 1 = 0\,a_{11} + 0\,a_{12} + 0\,a_{13} + 0\,a_{21} + 0\,a_{22} + 0\,a_{23} + 1\,a_{31} + 1\,a_{32} + 1\,a_{33} \\ P_1 = P_1 a_{11} + 0\,a_{12} + 0\,a_{13} + P_2 a_{21} + 0\,a_{22} + 0\,a_{23} + P_3 a_{31} + 0\,a_{32} + 0\,a_{33} \\ P_2 = 0\,a_{11} + P_1 a_{12} + 0\,a_{13} + 0\,a_{21} + P_2 a_{22} + 0\,a_{23} + 0\,a_{31} + P_3 a_{32} + 0\,a_{33} \\ P_3 = 0\,a_{11} + 0\,a_{12} + P_1 a_{13} + 0\,a_{21} + 0\,a_{22} + P_2 a_{23} + 0\,a_{31} + 0\,a_{32} + P_3 a_{33} \\ \tau = P_1 a_{11} + 0\,a_{12} + 0\,a_{13} + 0\,a_{21} + P_2 a_{22} + 0\,a_{23} + 0\,a_{31} + 0\,a_{32} + P_3 a_{33} \end{cases} . \tag{6}$$

or

$$B = C\,A \ , \tag{7}$$

where

$$B = \begin{pmatrix} 1 \\ 1 \\ 1 \\ P_1 \\ P_2 \\ P_3 \\ \tau \end{pmatrix}, C = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ P_1 & 0 & 0 & P_2 & 0 & 0 & P_3 & 0 & 0 \\ 0 & P_1 & 0 & 0 & P_2 & 0 & 0 & P_3 & 0 \\ 0 & 0 & P_1 & 0 & 0 & P_2 & 0 & 0 & P_3 \\ P_1 & 0 & 0 & 0 & P_2 & 0 & 0 & 0 & P_3 \end{pmatrix}, \text{and } A = \begin{pmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{21} \\ a_{22} \\ a_{23} \\ a_{31} \\ a_{32} \\ a_{33} \end{pmatrix}. \tag{8}$$

In general, there is no unique set of 9 parameters in A that satisfy perfectly the 7 constraints because we cannot invert Equation 7, but we can look for an approximation A^*. For instance, we can search for A^* that has values between 0 and 1 and that minimizes

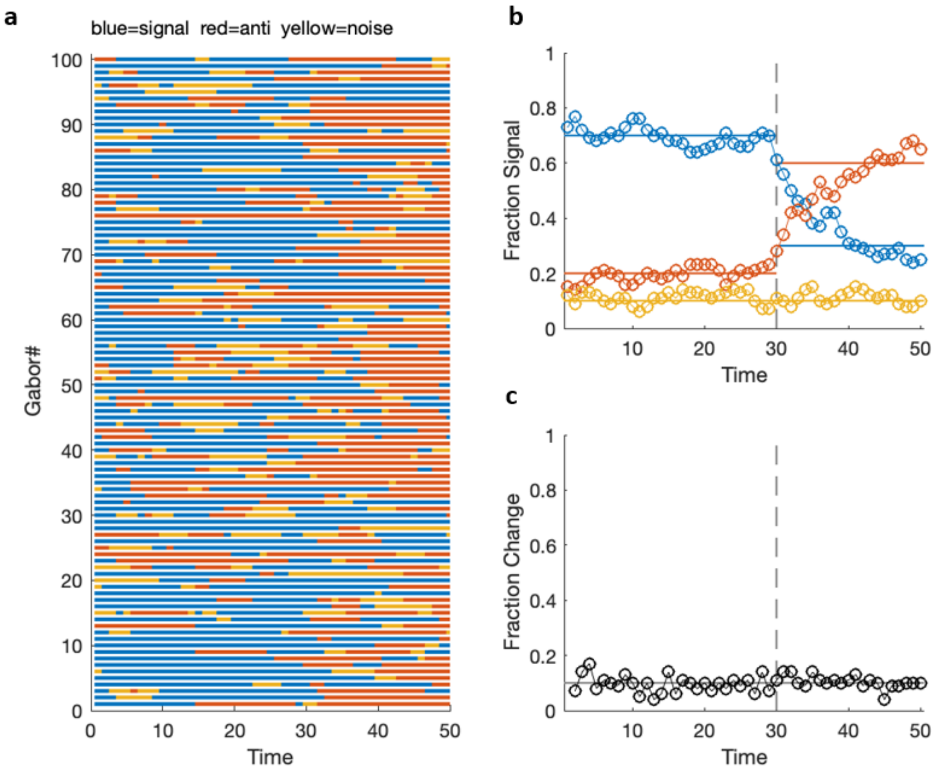
$$\min_x \left(\frac{1}{2}\|C\,x - B\|_2^2\right) . \tag{9}$$

We used the matlab function '**lsqlin**' to solve this constrained linear least-squares problem.

Simulations:

As an illustration, we simulate a set of N=100 Gabors that can be either pro-signal, anti-signal, or noise. After a set time (t_0=30), there is a transition from one set of probabilities for these 3 states (P_1=0.7,P_1=0.2,P_3=0.1) to another set of probabilities (P_1=0.3,P_1=0.6,P_3=0.1).

The important variable τ that we called 'lifetime' represents the probability that any Gabor (pro-signal, anti-signal, or noise) will change status. If it is 0.9 (which is actually used lifetime in the present study), then there are on average only 10% of Gabors that change from frame to frame. If this value is too high, then it will take some extended time to see a desired change in probabilities for (pro-signal, anti-signal, and noise). (The simulations are shown in the Supplementary Fig 6 below)



Supplementary Fig 6: Simulation. a) The simulation of 100 Gabors across time, where the color represents the current status of a Gabor (blue for pro-signal, red for anti-signal, and yellow for noise). **b)** The fraction of (pro-signal, anti-signal, and noise) Gabors across time (as if one were to take vertical slices of the figure on the left). **c)** The fraction of Gabors that change their status, over time. Importantly, for that simulation, even though there is a sudden change in the probabilities of (pro-signal and anti-signal) Gabors at time ($t_0 = 30$), there is not a sudden increase of the fraction of change at that time.