

Computational and neural
mechanisms underlying
decision-making
in humans

Thesis by
Ryo Adachi

In Partial Fulfillment of the Requirements for
the degree of
Doctor of Philosophy

The Caltech logo, featuring the word "Caltech" in a bold, orange, sans-serif font, centered within a light yellow rectangular background.

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2018
Defended January 19, 2018

© 2018

Ryo Adachi

ORCID: 0000-0003-0239-5694

ACKNOWLEDGEMENTS

Firstly, I would like to thank my advisor, John O’Doherty, for his generous support during my research. His optimistic encouragement definitely helped to move the research forward whenever I was deeply stuck in a problem. Since joining his lab in October 2012, I was lucky enough to work with not only bright and talented but also thoughtful and caring students and postdocs in the lab. First, I would like to thank Vikram Chib. His project on social facilitation was the very first fMRI project that I was involved in. He guided me through essential skills and techniques related to fMRI experiments, which eventually served as the foundation to my research for the thesis. Second, I would like to show appreciation to Shinsuke Suzuki. He has always been very helpful whenever I consulted with him about practical questions ranging from computational models to fMRI analysis. Also, it was an honor to work with him on his state-of-the-art research on consensus decision-making. Finally, I would like to show my deepest gratitude to Simon Dunne. We spent a significant amount of time together talking about our research and lives while walking around campus. His sharp insights and considerate attitude toward others helped me (try at least) to not only be a better scientist but also a nicer individual.

I would also like to thank Colin Camerer, Ben Gillen, and Antonio Rangel for being part of my committee, and for all I have learned from their classes and lab meetings. Their practical and accurate advice significantly improved the quality of the content of this thesis.

Studying with friends for preliminary exams is an unforgettable memory. I made great friends among classmates: Rahul Bhui, Nick Broten, Liam Clegg, Yifei Huang, Sergio

Montero, Gerelt Tserenjigmid, Jay Vilorio, Teagan Wall, and Jackie Zhang. Collaborating on and discussing course materials and beyond was more than valuable to me.

Living far from home was not easy for me at first but I eventually had a enjoyable life here in Pasadena thanks to all the friends in Japan and the Caltech community that I had fun with chatting, playing tennis, eating out, drinking, and more. To Albert Chung, Aika Gunji, Takuya Higo, Chika Horie, Yu Horie, Noriko Imai, Taisuke Imai, Koichiro Kajikawa, Mai Kajikawa, Yutaka Kayaba, Emiko Kimura, Seoyeon Lee, Maiko Obana, Tomoyuki Oniyama, Saneyuki Ono, Atsushi Shibukawa, Minako Suzuki, and Yodai Takei: thank you and I will certainly miss you all.

I would like to acknowledge the support of grants from the Nakajima Foundation. They not only supported me financially but also did care about my well-being while studying and living abroad.

Finally and most of all, I would like to thank my parents, Michiko and Seiichi, and my younger brother Yu for their support throughout my life. I could not have successfully earned the degree without you. I love you guys.

PUBLICATIONS AND CONTRIBUTIONS

- [1] Adachi R, Jessup R, O'Doherty JP (in preparation) Neural correlates of binary choice guided by visual attention.
R.A. analyzed the data and participated in the writing of the manuscript.
- [2] Adachi R, Suzuki S, O'Doherty JP (under review) Neural computations mediating temporal change detection in the human brain.
R.A. participated in the collection of the data, analyzed the data, and participated in the writing of the manuscript.
- [3] Chib VS, Adachi R, O'Doherty JP (under revision) Neural substrates of social facilitation effects on incentive-based performance.
R.A. participated in the collection of the data and in the writing of the manuscript.
- [4] Suzuki S, Adachi R, Dunne S, Bossaerts P, O'Doherty JP (2015) Neural mechanisms underlying human consensus decision-making. *Neuron* 86:591-602.
doi: <http://dx.doi.org/10.1016/j.neuron.2015.03.019>
R.A. participated in the collection of the data and in the writing of the manuscript.

ABSTRACT

How do we make economic decisions in everyday life? How do we make decisions in the face of uncertainty regarding the statistics of the environment? These are the questions that played a pivotal role in the formation of the field “Decision Neuroscience”. In each chapter of this thesis, we investigated the computational and neural mechanisms to tackle these questions using behavioral and neural data acquired through fMRI experiments.

In the first chapter, we investigated the computational and neural basis of economic decision-making in a binary choice task between two food items. We analyzed behavioral and neural data in a task where participants conducted a sequence of binary choices under the manipulation of fixation-based attention. We developed and calibrated a computational model based on evidence sampling and accumulation to show that the model not only accurately captured basic properties such as choice and reaction time (RT) but also the effect of attentional manipulation in participants’ behavior. We found that the evidence accumulation process predicted by the model to drive a decision was implemented in the areas of frontoparietal network including dmPFC and IPS. These regions also exhibited increased functional connectivity with the activity in vmPFC during choice period where sampled evidence was represented. Our results suggest the involvement of these areas in value-based binary choice.

In the second chapter, we examined the computations involved in the decision making under uncertainty. In particular, we aimed to pin down the computations related to temporal change detection. Temporal change detection is the capacity to detect change in the

statistics that govern the timing of occurrence of events. We analyzed behavioral data from a novel task where participants observed a sequence of images presented at irregular timings and tasked to detect a change in the frequency of image presentations. We developed and compared computational models from Bayesian to heuristic models and found that all the models captured quantitative aspects of participants' behavior equally well despite the difference in their computational complexity. Thus, we could not distinguish computations involved in temporal change detection solely from the behavioral data.

In the third chapter, we aimed to elucidate the computations involved in temporal change detection from the perspective of neural implementation using fMRI data by leveraging the computational models examined in the previous chapter. We found that the key variable to guide a decision derived from a computationally frugal heuristic model correlated with the activity of the frontoparietal network including dlPFC and IPS, while similar variables derived from more computationally taxing Bayesian models did not show significant correlation with any of the brain regions. Our results suggest that humans might be relying on a simple heuristic model to implement temporal change detection.

TABLE OF CONTENTS

Acknowledgements.....	iii
Publications and contributions.....	v
Abstract.....	vi
Table of Contents.....	viii
List of Illustrations and/or Tables.....	ix
Nomenclature.....	xi
Chapter I: Overview	1
Chapter II: Neural correlates of binary choice guided by visual attention.....	5
Introduction	5
Materials and methods	8
Results	20
Discussion	25
Chapter III: Computational account of temporal change detection in the human behavior	41
Introduction	41
Materials and methods	43
Results	55
Discussion	61
Chapter IV: Neural computations mediating temporal change detection in the human brain	78
Introduction	78
Materials and methods	79
Results	84
Discussion	86
Bibliography	98

LIST OF ILLUSTRATIONS AND/OR TABLES

<i>Figure</i>	<i>Page</i>
1. Illustration of task structure	29
2. Computational model.....	30
3. Behavioral results.....	32
4. Fixation-driven attentional effect on the stimulus value representation in vmPFC	33
5. Neural correlates of the time course of accumulator value and PPI result	34
6. Illustration of task structure	65
7. Behavioral results.....	66
8. Brain areas showing greater activity associated with button presses in the main task than in the control task	90
9. Neural correlates of the time course of the decision variable derived from each of the three computational models of temporal change detection.....	91
10. Brain regions that uniquely correlated with the time course of the decision variables derived from the TAM.....	92

<i>Supplemental figure</i>	<i>Page</i>
1. Joint distribution of choice and RT of participants' behavior.....	35
2. Attentional bias in participants' behavior.....	37
3. Experimental design.....	68
4. Performance metric.....	69
5. Participants' behavior across runs.....	70
6. Participants' behavior across conditions.....	71
7. Performance of the optimal Bayesian model.....	72
8. Performance of computational models.....	73
9. Best fitting parameter values for the best sub-model of each computational model of participants' behavior.....	74
10. Performance of the computational model based on delta-rule.....	76
11. Illustration of the decision variable regressor.....	93

<i>Supplemental table</i>	<i>Page</i>
1. Details of the model fitting procedure.....	38
2. Effect of item values on participants' choice behavior.....	39
3. MNI coordinates of peak activation voxels from the contrasts in GLM1.....	40
4. Details of model fitting and comparison.....	77
5. Properties of the decision variables.....	94
6. Average correlation between the time courses of the decision variables.....	94
7. MNI coordinates of peak activation voxels from the contrasts in GLM1.....	95
8. MNI coordinates of peak activation voxels from the contrasts in GLM2-GLM4.....	96
9. MNI coordinates of peak activation voxels from the contrasts in GLM5.....	97

NOMENCLATURE

- BMS.** Bayesian model selection
- BOLD.** Blood-oxygen-level-dependent
- dIPFC.** Dorsolateral prefrontal cortex
- dmPFC.** Dorsomedial prefrontal cortex
- EEG.** Electroencephalography
- fMRI.** Functional magnetic resonance imaging
- GLM.** General linear model
- HRF.** Hemodynamic response function
- IPL.** Intraparietal lobule
- IPS.** Intraparietal sulcus
- LIP.** Lateral intraparietal
- MEG.** Magnetoencephalography
- MNI coordinate.** Montreal Neurological Institute coordinate
- PFC.** Prefrontal cortex
- PPI.** Psychophysiological interaction
- ROI.** Region of interest
- RT.** Reaction time
- SPM.** Statistical parametric mapping
- TE.** Echo time
- TR.** Repetition time
- vmPFC.** Ventromedial prefrontal cortex

*Chapter 1***OVERVIEW**

“Decision Neuroscience” is an interdisciplinary field in neuroscience that bridges ideas from economics, psychology, artificial intelligence, and cognitive neuroscience to understand computational and neural mechanisms underlying human decision-making. The key questions include but are not limited to how we make economic choices (Krajbich et al., 2010; Rangel and Hare, 2010; Hare et al., 2011; Hunt et al., 2012), how we make decisions in the face of uncertainty regarding the statistics of the environment (Behrens et al., 2007; Rushworth and Behrens, 2008; Nassar et al., 2010, 2012; Payzan-LeNestour et al., 2013; McGuire et al., 2014), how we infer the intention of other individuals in the social interaction (Hampton et al., 2008; Yoshida et al., 2008, 2010; Dunne and O’Doherty, 2013; Suzuki et al., 2015), and how we engage in and arbitrate between habitual and goal-directed control of behavior (Daw et al., 2005; Lee et al., 2014; O’Doherty et al., 2017).

In this thesis, I tackle some of these questions by conducting cognitive neuroscience experiments when scanning participants in the fMRI scanner and analyzing behavioral and neural data using the method called “Model-based fMRI analysis” (O’Doherty et al., 2007). In this approach, a rigorous computational model that captures quantitative aspects of the behavioral data is first developed and fitted to participants’ behavior. Then, computational variables in the model that play a crucial role implementing a decision are correlated against BOLD signal. By using this approach we can do more than test the areas of the brain that are involved in a particular cognitive process: we are capable of identifying which areas of the brain implement an exact computation prescribed by the computational model of the decision making. The structure of this thesis is as follows.

In Chapter 2, we investigate our decision making in an economic context. More specifically, we focus on the binary choice between two food items. Previous studies showed that sequential sampling models such as drift diffusion model that are based on evidence sampling and evidence accumulation process describe quantitative aspects of participants' behavior in value-based binary choice tasks. These models were first fitted to participants' behavior, and then indirect model-related stimulus or a model-derived quantity, which served as a good proxy for but does not necessarily represent the evidence accumulation process, were correlated against BOLD signal. Here, we directly correlated the time course of evidence accumulation process predicted by a fitted sequential sampling model against BOLD signal. In order to obtain a reliable time course of accumulator values, acquiring the information of participants' attention guided by visual fixations is crucial. We achieved this in a task where we externally manipulated the fixation to the items by sequentially displaying the two items. The information of fixation-based attention is utilized in a novel computational model we developed called attentional neural drift diffusion model (anDDM). The time course of accumulator value estimated by anDDM correlated with the activity in the areas of the frontoparietal network including dmPFC and IPS. Moreover, these regions showed increased functional connectivity during choice period to vmPFC where item values used as the input to the accumulators are represented. Our results corroborate the involvement of the regions in the frontoparietal network in the evidence accumulation process in value-based binary choice.

In Chapter 3, we investigate our decision making under uncertainty about the statistics of the environment. Specifically, we examine the computational mechanisms underlying temporal change detection in human behavior. Temporal change detection is the capacity to detect changes in the statistics governing the timing of the occurrence of events. This ability is crucial for producing adaptive behavior in a volatile environment, yet little is known about its underlying computational

and neural mechanisms. Here, we applied computational models of varying degrees of sophistication to the behavioral data from human participants performing a novel temporal change detection task when they were scanned with fMRI. We found that a parsimonious heuristic model performed equally as well as more sophisticated Bayesian models. The result is interesting in its own right in that the computational models we tested explain variance in participants' behavior equally well regardless of the difference in the computational complexity across models, and also in that we cannot distinguish computational models solely from behavioral data. This result leads to our analysis of fMRI data in Chapter 4 to differentiate computational models using neural data.

In Chapter 4, we investigate the neural computations involved in temporal change detection. We apply the computational models that explained participants' behavior equally well in Chapter 3 to fMRI data in order to identify which model is most plausible from the perspective of neural implementation. We found that the heuristic model correlated with activity in the regions of frontoparietal network including dlPFC and IPS, which furthermore accounted for activity in these areas significantly better than the Bayesian models. Our results suggest that the brain implements an approximate strategy to detect temporal change points and that the computations behind the model are implemented in brain areas also implicated in evidence accumulation in perceptual and value-based decision making.

*Chapter 2***NEURAL CORRELATES OF BINARY CHOICE
GUIDED BY VISUAL ATTENTION**

Adachi R, Jessup R, O’Doherty JP. Neural correlates of binary choice guided by visual attention. R.J. and J.P.O. designed the experiment. R.J. collected the data. R.A. and R.J. performed the data analysis. R.A., R.J, and J.P.O. wrote the manuscript. J.P.O. supervised the research.

INTRODUCTION

How do we choose between two alternatives? Elucidating the neural computations underlying binary choice has been a key challenge in neuroscience in both perceptual and value-based decision making domains (Gold and Shadlen, 2007; Heekeren et al., 2008; Rangel and Hare, 2010; Summerfield and Tsetsos, 2012). In both humans and non-human primates, a widely acknowledged mechanism posits that evidence in favor of one choice over the other is sampled from stimulus properties and accumulated until a decision criterion is met. This mechanistic account is compatible with the computations underlying sequential sampling models (Ratcliff et al., 2016), which have been shown to accurately describe quantitative aspects of participants’ behavior in binary choice tasks (Usher and McClelland, 2001; Bogacz et al., 2006; Bogacz, 2007; Ratcliff and McKoon, 2008). In recent years, neural correlates of the computational variable associated with sequential sampling models have been examined in humans using fMRI, MEG, and EEG (Mulder et al., 2014; Forstmann et al., 2016). In particular, there has been growing interest in identifying brain regions involved in the evidence accumulation process, which is the key computation of the model.

To find brain regions implementing the accumulation of evidence in a binary choice task, the time course of accumulator value predicted by a sequential sampling model could be directly correlated against BOLD signal. Alternatively, previous studies correlated other indirect model-related stimulus (Heekeren et al., 2004; Basten et al., 2010; De Martino et al., 2013a) or a model-derived quantity (Hare et al., 2011) that served as a proxy for but does not necessarily represent evidence accumulation process. This is partly because a reliable estimate of the time course of accumulator value within a trial cannot be attained in a binary choice task where two options are presented simultaneously without additional information of attention driven by visual fixations. The information of fixation-driven attention enables the computational model to incorporate the knowledge of which of the two options is more relevant to the accumulation process for the participants at each time point during the trial. Therefore, the description of how the decision unfolds over time within a trial can be estimated at a higher precision.

The importance of attention in a choice task has been shown in a number of previous reports (Armel et al., 2008; Krajbich et al., 2010; Krajbich and Rangel, 2011; Lim et al., 2011; Towal et al., 2013; Tavares et al., 2017). In the choice process, we shift our gaze between the two alternatives back and forth and the pattern of fixation exerts a significant effect on the final choice. We refer to this fixation-driven attentional effect on choice as attentional bias. There are two approaches to obtain the information regarding the fixation-driven attention of the participants: first, we present the options in the choice set simultaneously and measure visual fixation directly using an eye-tracking device when participants are freely moving their fixations (Krajbich et al., 2010; Krajbich and Rangel, 2011; Towal et al., 2013; Tavares et al., 2017) or when participants are instructed to fixate on a particular item (Lim et al., 2011). Second, we explicitly control fixations by displaying only one option at a time and achieve comparison by presenting alternatives in the choice set sequentially

(Armel et al., 2008). The latter approach is easier to implement and exogenous control of the fixated option enables us to obtain the time course of the accumulator value at a decent temporal resolution.

The information of fixation-based attention has been incorporated in a variant of sequential sampling model called attentional drift diffusion model (aDDM, Krabich et al., 2010). The model has been shown to capture attentional bias in participants' behavior in various experiments (Krajbich et al., 2012; Tavares et al., 2017). However, this is a psychological model intended to describe only the behavioral data and its equivalent cortical model that has the capacity to describe both the firing rate of cortical neurons and the behavioral data has not yet been explored (Bogacz, 2007).

The goal of the present study is two-fold: first of all, we aimed to develop a neurally plausible computational model of binary choice in the value-based decision making domain that could capture attentional bias in participants' behavior. Secondly, we aimed to identify brain regions representing quantities associated with the computational model. In particular, we tested neural correlates of the time course of accumulator value during trials.

To address these questions, we developed a behavioral task and ran participants through this task while they were scanned with fMRI. In the task, participants conducted a sequence of binary choices between two food items. To manipulate the fixation-driven attention of the participants and control which item is more relevant for the evidence accumulation process to guide a choice, we alternated the item presented on the screen every 1 second. This information is reflected in a variant of sequential sampling model we developed such that it flexibly changes the sampled evidence every 1 second depending on which item is currently presented on the screen.

Behaviorally, we hypothesized that the participants' choice behavior will show a strong attentional bias as a result of sequential presentation of the two items under comparison. At the neuroanatomical level, we first hypothesized that the ventromedial prefrontal cortex (vmPFC) would be involved in representing item values as shown in a number of previous reports (Chib et al., 2009; Clithero and Rangel, 2013; McNamee et al., 2013) and the activity in this area would be modulated by participants' attention guided by visual fixations. Secondly, the activity in the dorsomedial prefrontal cortex (dmPFC) and intraparietal sulcus (IPS) would reflect the evidence accumulation process predicted by the computational model. These are the findings from previous studies of value-based binary choice mechanisms under voluntary visual fixations (Hare et al., 2011; Lim et al., 2011), and yet we hypothesized that the same mechanisms would be involved in our task of binary choice under exogenous manipulation of visual fixations.

MATERIALS AND METHODS

Participants

21 right-handed participants performed the experiment (18-26 years old, 14 males). The participants were prescreened to exclude those with a prior history of neurological or psychiatric illness. The School of Psychology Research Ethics Committee at Trinity College Dublin approved the experiment and all the participants gave informed consent.

Experiment tasks

The experiment contained the following three phases in order. In the first stage participants made a sequence of binary choices of food items that they were willing to consume at the end of the experiment (choice phase; Figure 1A). There were 65 trials in the choice phase and the item pair used for each trial was selected randomly without replacement from 130 different food items

including snacks, candies, and condiments. We selected food items that were easily obtainable at a local grocery store so that the participants were familiar with them. The set of items was fixed for all the participants but the item pair used in each trial was randomized for each participant. At the beginning of each trial, one item was shown on either the left- or right-hand side of the screen for 1 second, followed by the presentation of the paired item on the other side of the screen for 1 second. In the current paper the items presented first and second are referred to as item1 and item2, respectively. The side of item1 presentation was randomized across trials for each participant. The presented item and its location on the screen alternated between the two every 1-second until the participant made a choice with the button press. There was no restriction on reaction time (RT). The choice was enabled after 1 second from the start of a trial, which coincided with the timing of the first item2 presentation, to ensure that the participant went through the comparison process of the two items to make a choice. The button press was made with either the left or right index finger to choose the item presented on the corresponding side of the screen. Importantly, the choice of either item was possible irrespective of which item was displayed on the screen at the time of button press. The presentation of a fixation cross (6-18 seconds, mean 12 seconds) immediately following the button press separated the trials.

In the second stage the participants reported their willingness-to-pay (WTP) for the 130 food items used in the choice phase (pricing phase; Figure 1B). We adopted the Becker-DeGroot-Marschak (BDM) auction paradigm to measure the valuation of items specific to each participant (Becker et al., 1964). Importantly, the participants did not have incentives to deviate from reporting their true WTP in this paradigm. At the beginning of each trial, the participant was given an endowment of 3 euros that could be spent on purchasing the item involved in the trial. The item was presented on the screen until the participant reported the amount of money he/she would be willing to pay in exchange for the item by an integer amount between 0 and 3 euros. There was no restriction on RT

and each trial was separated by the presentation of a fixation cross (2-6 seconds, mean 4 seconds). The order of items used in the trials was randomized for each participant. At the end of the experiment, we randomly selected a trial; if the reported WTP for the item in the trial was greater than the number drawn from the uniform distribution between 0 and 3, the participant paid the amount of money equal to the number drawn from the endowment of 3 euros in exchange for the item and kept the remaining amount. Otherwise, the participant did not get the item but kept the endowment of 3 euros. In the current paper we use the terms WTP and item value interchangeably.

In the third stage the participants rated how familiar they were with each of the 130 food items used in the choice and pricing phase (familiarity phase). The paradigm of the task was the same as the pricing phase except that the participants were rating their familiarity with each item on an integer scale from 1 (not familiar) to 5 (very familiar). We obtained familiarity ratings to ascertain that the participants were familiar with the items and uncertainty related to unfamiliar items when making binary choices and reporting WTPs would be negligible (Plassmann et al., 2007).

At the end of the experiment, the participants could consume the food item obtained in the trials randomly selected from each of the choice phase and the pricing phase (1 or 2 food item(s) depending on the outcome of BDM auction in the pricing phase). We instructed the participants not to eat for 4 hours prior to the experiment to increase the overall valuation of the items. Since it was crucial for our experiment that the participants reported their preferred item in every trial in the choice phase and the true valuation of each item in the pricing phase, we also instructed the participants that they would have to stay in the lab for up to an hour after the experiment and that the only foods they could consume in the waiting period were those they obtained from the choice phase and the pricing phase. Further, we explained the structure of BDM auction extensively to the participants to make sure that they understood that it was optimal for them to report their true WTP

of each item in the pricing phase. The participants conducted the choice phase and the pricing phase in the fMRI scanner but the scanning data from only the choice phase was used for the fMRI analysis.

Computational model

We developed a computational model to describe participants' behavior quantitatively. We aimed to construct a model that is plausible at the neural level and capable of capturing participants' choice, RT, and attentional bias associated with external manipulation of visual fixation unique to our task. We combined the ideas from the model used in two previous studies. One was the neural drift diffusion model (nDDM), which is a neurally plausible implementation of the drift diffusion model that can accurately capture participants' choice and RT in a binary choice task (Hare et al., 2011). The other was the attentional drift diffusion model (aDDM) that can describe not only participants' choice and RT but also attentional bias in the choice behavior in a binary choice task (Krajbich et al., 2010). We refer to our model as attentional neural drift diffusion model (anDDM). We provide the description of anDDM here (Figure 2A).

The model contains two accumulators representing two pools of neurons, each of which accrue evidence toward the choice of item1 or item2. The input to each accumulator changes according to the attended item (i.e., on-screen item) at the time of accumulation. It scales with the value difference of the two items under comparison with a discount factor on the value of unattended item (i.e., off-screen item) capturing the effect of fixation-driven attention toward the attended item. For instance, when item1 is on the screen the input to the accumulator supporting the choice of item1 is the value of item1 minus the value of item2 discounted by the attentional bias parameter α (i.e., $V^1 - \alpha V^2$). The input to the accumulator supporting the choice of item2 is opposite in sign (i.e.,

$\alpha V^2 - V^1$). The interpretation of the attentional bias parameter α is that there is a full attentional bias toward the attended item when $\alpha = 0$ ignoring the value of unattended item in the sampled evidence. On the other hand when $\alpha = 1$, there is no attentional bias toward the attended item and the value of attended and unattended items are weighted equally. During the initial 1 second of a trial when the identity of item2 is not yet revealed, we set the value of item2 to be equal to its expected value of 1.5. Each of the two inputs is then perturbed by an independent identically distributed noise $\varepsilon \sim N(0, \eta^2)$ before its value is added to the corresponding accumulator. Also, there are inhibitory connections between the two accumulators with the magnitude of inhibition controlled by the parameter ν . In summary, the accumulator value for the attended item i and for the unattended item $-i$ at time step $t + 1$, which are denoted by A_{t+1}^i and A_{t+1}^{-i} respectively, is calculated as

$$\begin{aligned} A_{t+1}^i &= \max \{0, A_t^i + (V^i - \alpha V^{-i})\delta dT - \nu A_t^{-i} + \varepsilon^i \sqrt{dT}\} \\ A_{t+1}^{-i} &= \max \{0, A_t^{-i} + (\alpha V^{-i} - V^i)\delta dT - \nu A_t^i + \varepsilon^{-i} \sqrt{dT}\}. \end{aligned} \quad (1)$$

Here, δ is the drift rate stipulating the average slope of evidence accumulation per 1 ms for the unit input value. dT is the duration between two successive accumulation time steps in milliseconds and it is arbitrarily set to 100 ms. Another parameter κ was introduced to restrict the accumulator values computed in Equation (1) during the initial 1-second from the start of the trial such that

$$\begin{aligned} A_{t+1}^i &= \min \{\kappa, A_{t+1}^i\} \\ A_{t+1}^{-i} &= \min \{\kappa, A_{t+1}^{-i}\}. \end{aligned} \quad (2)$$

This restriction made the model prediction in accordance with our experimental manipulation that the choice of an item during the initial 1-second of a trial was not allowed. To justify the introduction of this parameter, we compared this model against a model without κ . It simply assumes that the accumulator value could not exceed the threshold value of 1 during the initial 1-second. The model with κ was preferred in terms of Bayesian Information Criterion (BIC=5582 against 5623 of the model without κ). The two accumulator values are set to zero at the start of each trial and they are updated every 100 ms until the value of either of the two accumulators exceeds the fixed threshold of +1, resulting in the choice of corresponding item.

Model fitting to participants' behavior

We calibrated anDDM to participants' behavior at the group level using maximum likelihood estimation by grid search. The model contains five free parameters: α , η , ν , δ , and κ . The model fitting procedure is as follows. First, we discretized the range of each parameter and set up a grid in the joint parameter space. The grid is the Cartesian product of the following range of five parameters: α : 0: 0.1: 0.5 , η : 0.005: 0.005: 0.04 , ν : 0: 0.1: 1.0 , δ : 2.0: 1.0: 7.0×10^{-4} , and κ : 0: 0.1: 1.0. Second, for each point on the grid defining a set of parameter values θ , we computed the empirical joint distribution of choice and RT (RT was categorized into one of 9 bins with the duration of 1 second each: from 1-2 seconds to 9-10 seconds) by simulating the model 5000 times for each item value pair. Note that the item value pair takes into account the information of which of the two items is presented first. Thus, there are 16 pairs of item1 and item2 values denoted by $V = (V^1, V^2)$ (item values V^1 and V^2 are integers between 0 and 3 euros measured in the pricing phase of the experiment). We did not differentiate whether item1 was presented on the left- or right-hand side of the screen because there was no significant difference between the probability of choosing item1 when item1 was presented on the left-hand side and the right-hand side of the screen ($p =$

0.31, Wilcoxon signed-rank test). All the simulation runs with RT exceeding 10 seconds were discarded from the calculation of joint probabilities to adjust for the range of RT observed in participants' behavior. Third, the log-likelihood value (LL) for a set of parameter values θ was calculated by summing the log-probability from the previous step across trials ($N = 1355$) using the information from participants' behavior of choice C_j and RT bin T_j given the item value pair $V_j = (V_j^1, V_j^2)$ in each trial j :

$$LL = \sum_{j=1}^N \log P(C_j, T_j | V_j, \theta).$$

Fourth, log-likelihood values for all the points on the grid were compared to set up a smaller grid to start another grid search in a more confined joint parameter space. We iterated these steps three times by narrowing the range of parameters tested to the values around the best fitting parameter values from the previous iteration. We obtained the best fitting parameter value of $\alpha = 0$, $\eta = 0.0175$, $\nu = 1.0$, $\delta = 4.75 \times 10^{-4}$, and $\kappa = 0.575$ with $-LL = 2773$ (See Table S1 for the detail of each iteration).

Creating the accumulator value regressor

To examine the brain regions implementing the accumulation of evidence to guide participants' choice, we estimated the time course of accumulator value during each trial at 100ms time resolution and correlated it directly against BOLD signal. The accumulator value time course that conforms to each trial's item value pair, choice (item1 or item2 chosen) and RT (RT was categorized into one of 9 bins with the duration of 1 second each: from 1-2 seconds to 9-10 seconds) was created as follows. For each of the 16 possible item value pairs of item1 and item2

(each item had an integer value between 0 and 3 euros measured in the pricing phase of the experiment), we first simulated the behavior of anDDM 100,000 times using Equation (1) and (2) with the set of best fitting parameter values obtained by the process of model fitting. Notice that the order of fixations to the two items under comparison was taken into account in the definition of item value pair: for instance, (1,2) and (2,1) value pairs were treated differently. Second, we categorized the time courses obtained from simulation runs by the combination of choice and RT bin. Finally, time courses in each category were averaged at every time point separately for accumulators accruing evidence toward the choice of item1 and item2. Thus, we created a total of up to 18 estimated average accumulator value time course pairs for each item value pair (See Figure 2B for examples of these time courses). Note that there were some combinations of value pair, choice, and RT bin where no simulation runs were categorized (e.g., for (3,0) value pair, no simulation runs resulted in the choice of item2 having the value 0 with RT of 8-9 seconds). Nonetheless, we could obtain the accumulator value time courses for all the combinations of value pair, choice, and RT observed in the participants' behavior due to an appropriate model calibration and a large number of simulation runs. In fMRI analysis we assumed that the two pools of neurons in anDDM, each implementing the accumulation process toward the choice of one item, would probably be spatially intermingled. Thus, the BOLD signal we measured would reflect the sum of activities of these two pools of neurons. Given this assumption we correlated the time course of the sum of two accumulator values predicted by anDDM against BOLD signal. For each trial the accumulator value time course for the trial's value pair, observed choice, and RT bin was used in the first-level participant-specific general linear models (GLMs) after being convolved with canonical hemodynamic response function (HRF) (See GLM1 in Materials and Methods).

fMRI data acquisition

The task was conducted in a Philips Achieva 3T scanner using a phased-array eight-channel head coil. The imaging data was acquired at a 30 degree angle from the anterior commissure-posterior commissure line in order to avoid the signal dropout in the orbitofrontal cortex (Deichmann et al., 2003), using a gradient echo T2* weighted echo planar imaging sequence. 39 axial ascending slices were acquired with a slice gap of 0.35mm with the following scan parameters: echo time 28 ms, repetition time 2000 ms, field of view 240×240 mm, and voxel size 3×3×3.55 mm. High-resolution T1 images were collected at the beginning of each participant's session.

fMRI data preprocessing

SPM8 (Wellcome Department of Imaging Neuroscience, London, UK; www.fil.ion.ucl.ac.uk/spm) was used to preprocess and analyze the fMRI data. The functional images for each participant were spatially realigned to the first volume using a 6-parameter rigid body spatial transformation and then slice time corrected to the middle (i.e., 20th) slice. The high-resolution structural image was then coregistered to the mean functional image generated from the realignment procedure. The realigned functional images were spatially normalized to the MNI template with a resample voxel size of 2 mm isotropic by applying the warping parameters from the segmentation process of the coregistered structural image. Finally, the normalized functional images were spatially smoothed using a Gaussian kernel of 8 mm isotropic.

fMRI statistical analysis

We estimated several participant-specific GLMs to test our hypotheses. All the regressors in the GLMs were convolved with a canonical HRF provided by SPM8. Participant-specific contrasts from the estimated GLMs were used for second-level random effect analyses. For the whole-brain analyses, unless otherwise noted, we applied the correction for multiple comparisons at the cluster

level $p < 0.05$ with the underlying voxel-wise threshold of $p < 0.005$ and the minimum spatial extent of 191 voxels, calculated by using a version of 3dClustSim in AFNI that does not assume Gaussian spatial correlations, and which is therefore not susceptible to inflated false positives (Cox et al., 2017). For display purposes all the statistical maps are presented at $p < 0.005$ uncorrected, with a more stringent $p < 0.001$ uncorrected overlaid. In the tables the coordinates reported are in MNI coordinates of the peak voxel in each of the activation clusters. All the GLMs included boxcar functions for excluded trials, 6 movement regressors generated from the realignment procedure in the image preprocessing accounting for the head motion, and a constant term as regressors of no interest. SPM default serial orthogonalization was applied to parametric modulators in each GLM.

GLM1

The first GLM was estimated to test for the regions correlating with (i) task relevant properties such as stimulus values and motor responses by the button press to report a choice, and (ii) the time course of accumulator value, which was the main effect of our interest. This GLM included (1) indicator function for item presentation on the left-hand side of the screen, parametrically modulated by (2) the corresponding item value, (3) indicator function for item presentation on the right-hand side of the screen, parametrically modulated by (4) the corresponding item value, (5) button press by the left index finger, (6) button press by the right index finger, and (7) indicator function of every 100 ms from the start of each trial to the button press, parametrically modulated by (8) the time course of accumulator value predicted by anDDM (See Creating the accumulator value regressor in Materials and Methods). The contrasts of interest were (i) stimulus value (Figure 4A), (2)+(4); motor response, (5)+(6); and (ii) the time course of accumulator value, (8) (Figure 5A). For small volume correction for contrast (ii), we used a sphere of 10 mm radius around average peak coordinates in left IPS ($x = -34, y = -55, z = 44$) found in relevant previous studies (Basten et al., 2010; Hare et al., 2011; Hunt et al., 2014). See Table S3 for all the activations.

ROI analysis

A previous study showed that the activity in vmPFC is modulated by fixation-driven attention and represents the difference between the value of attended item and the value of unattended item (Lim et al., 2011). Given these findings we conducted a set of region of interest (ROI) analyses (See GLM2 and GLM3 below) to investigate the attentional modulation of the value representation within an independent ROI in vmPFC: a 10 mm sphere centered at the peak activation coordinates ($x=-6, y=44, z=-5$) of a cluster showing correlation with food item value in a previous report (Hare et al., 2011). All the ROI analyses were implemented using Marsbar toolbox (Brett et al., 2002).

GLM2

We tested if the activity in vmPFC represented the value of attended item minus the value of unattended item as found in a previous study (Lim et al., 2011). This GLM contained the following regressors: (1) indicator function for item presentation on the left-hand side of the screen, parametrically modulated by (2) the value of left-hand side item minus the value of right-hand side item, (3) indicator function for item presentation on the right-hand side of the screen, parametrically modulated by (4) the value of left-hand side item minus the value of right-hand side item, (5) button press by the left index finger, and (6) button press by the right index finger. The contrasts of interest were the value of left- minus right-hand side item when fixating on the left-hand side item, (2), and the value of left- minus right-hand side item when fixating on the right-hand side item, (4) (Figure 4B).

GLM3

We investigated if the value difference representation in vmPFC found by GLM2 was driven by either the value of attended item, the value of unattended item or both. This GLM contained the following regressors: (1) indicator function for item presentation on the left-hand side of the screen,

parametrically modulated by (2) the value of left-hand side item and (3) the value of the right-hand side item, (4) indicator function for item presentation on the right-hand side of the screen, parametrically modulated by (5) the value of right-hand side item and (6) the value of left-hand side item, (7) button press by the left index finger, and (8) button press by the right index finger. The contrasts of interest were the value of attended item, (2)+(5), and the value of unattended item, (3)+(6) (Figure 4C).

Bayesian model selection analysis (BMS)

Previous studies reported that the activity in vmPFC represents the output of the choice process in the form of the value of chosen item minus the value of unchosen item (Boorman et al., 2009; Jocham et al., 2012), which tends to be highly correlated with the value of attended item minus the value of unattended item due to the observation that attended item tends to be chosen. Therefore, we conducted a Bayesian model selection analysis (Stephan et al., 2009) within an independent ROI in vmPFC (See ROI analysis in Materials and Methods) to test whether the activity in this area was better explained by (i) the value of attended and unattended item or (ii) the value of chosen and unchosen item. We used GLM3 for (i) and created another GLM4 for (ii). GLM4 included the following regressors: (1) indicator function for item presentation on the left-hand side of the screen, parametrically modulated by (2) the value of chosen item and (3) the value of unchosen item, (4) indicator function for item presentation on the right-hand side of the screen, parametrically modulated by (5) the value of chosen item and (6) the value of unchosen item, (7) button press by left index finger, and (8) button press by the right index finger. This BMS procedure yielded a voxelwise exceedance probability that assigned a probability to the event that the GLM3 containing the value of attended and unattended items explained the neural activity in vmPFC ROI better than the GLM4 containing the value of chosen and unchosen items.

PPI analysis

We conducted psychophysiological interaction (PPI) analysis to test the interaction between brain regions that represent quantities related to anDDM. At each time step, the model samples evidence from item values and adds it to the accumulators. Therefore, vmPFC representing item value should have increased functional connectivity with dmPFC and left IPS that are involved in the evidence accumulation process during the decision period. We examined this with a GLM that contained the following regressors: (i) BOLD signal time course within 5 mm sphere surrounding the vmPFC activation peak coordinates ($x=-6$, $y=36$, $z=-12$) obtained by the contrast of item value in GLM1, and (ii) interaction between BOLD signal time course in vmPFC and decision period modeled by a boxcar function. The model also included all the other regressors in GLM1. Notice that the psychological regressor (i.e., a boxcar function representing the decision period) was contained in the regressor of GLM1 ((1)+(3) of GLM1; see GLM1 in Materials and Methods). We report the parameter estimate in 5 mm sphere surrounding the peak activation coordinates within dmPFC and left IPS clusters that were significantly correlated with the time course of the accumulator value identified by GLM1 (Figure 5B).

RESULTS

Behavioral results

Average WTP across participants was 1.17 (SD = 0.94), and about 72% of the reported WTPs had positive values. The average familiarity rating was 3.82 (SD = 1.26), indicating that the participants were familiar with most of the items used in the choice and pricing phases. Among the total of 1365 trials in the choice phase from 21 participants, 10 trials with extreme RTs (\geq third quartile + $3 \times$ interquartile range) were excluded from further analysis. In the choice phase, the participants' choice behavior was significantly modulated by the value difference between the two

items under comparison and its interaction with the sum of item values (Table S2; $p < 0.001$ for value difference and $p = 2.79 \times 10^{-6}$ for interaction, mixed effect logistic regression). The result justifies our model calibration procedure based on item value pair (i.e., item1 and item2 value pair), rather than only on the value difference of the two items.

Model fitting to participants' behavior

To describe participants' behavior quantitatively, we developed attentional neural drift diffusion model (anDDM, Figure 2A; see Materials and Methods). The model was constructed to be plausible at the neural level and capture not only choice and RT but also attentional bias. This is achieved by the following two key components of the model: (i) the input to the accumulators changes according to which item is attended to at each step of accumulation, and (ii) the value of unattended item is discounted relative to the value of attended item so that the value of attended item has a greater influence on the sampled evidence added to the accumulators than the value of unattended item. We calibrated anDDM to participants' choice and RT at the group level by using maximum likelihood estimation procedure (See Materials and Methods). The model is flexible enough to capture quantitative aspects of participants' behavior (Figure 3A for choice and Figure 3B for RT; see Figure S1).

Attentional bias in participants' choice

Participants' choice exhibited a strong attentional bias as observed from greater than 0.5 probability of choosing attended item when the value difference between the two items was zero (Figure 3C; $p = 1.35 \times 10^{-4}$, Wilcoxon signed-rank test, see Figure S2). In the computational model the best fitting attentional bias parameter was $\alpha = 0$, suggesting that the value of unattended item was fully discounted in the sampled evidence. These observations indicate that the calibrated model

accurately captured the effect of fixation-driven attention in participants' behavior. To examine whether the attentional bias parameter α reflected the variability of the magnitude of attentional bias in the choice behavior across participants, we fitted anDDM to the behavior of each individual participant by varying this parameter alone. The other four parameters were fixed to the best fitting parameter values to the participants' group data. We found that as the value of attentional bias parameter decreased, meaning that the input to the accumulators was more dominated by the value of attended item, the probability of choosing the attended item increased (Figure 3D; $r = -0.50$, $p = 1.97 \times 10^{-2}$, Spearman's rank correlation).

fMRI results

Stimulus value representation in vmPFC was affected by attentional manipulation

In the choice phase the inputs to the two accumulators prescribed by anDDM necessitate the neural representation of the value of two items under comparison. In accordance with a number of previous studies (Chib et al., 2009; Clithero and Rangel, 2013; McNamee et al., 2013), a whole brain analysis revealed that the activity in vmPFC correlated with the value of the food item the participants were fixating at (i.e., attended item) (Figure 4A; $p < 0.05$ whole brain corrected at cluster level, see Table S3 for other activated areas, see GLM1 in Materials and Methods).

We then conducted a set of region of interest (ROI) analyses within an independent ROI in vmPFC to further investigate the effect of attentional manipulation of alternating the presented item every 1 second. First, we tested a claim from a previous study using a similar paradigm that vmPFC encoded the value of attended item minus the value of unattended item (Lim et al., 2011). We found that vmPFC activity was significantly modulated positively and negatively by the value of left item minus the value of right item when the participants were fixating on the left-hand side and the right-

hand side item, respectively (Figure 4B; $p = 7.1 \times 10^{-3}$ and $p = 9.6 \times 10^{-3}$ in order, Wilcoxon signed-rank test, see GLM2 in Materials and Methods). Second, we investigated an observation from our model fitting result that the inputs to the two accumulators were dominated by the value of the attended item at least at the group level (i.e., $\alpha = 0$). Therefore, we hypothesized that the modulation of vmPFC activity by the difference between the value of attended item and the value of unattended item was driven by the value of attended item under the strong attentional manipulation in our task. This ROI analysis revealed that the activity in vmPFC was indeed significantly correlated with the value of attended item but not with the value of unattended item (Figure 4C; $p = 2.50 \times 10^{-2}$ for attended item value and $p = 0.38$ for unattended item value, Wilcoxon signed-rank test, see GLM3 in Materials and Methods). As a robustness check, we changed the order of parametric modulators for attended and unattended items in GLM3 and confirmed that the effect of attended item was significant while the effect of unattended item was not ($p = 4.57 \times 10^{-2}$ for attended item value and $p = 0.41$ for unattended item value, Wilcoxon signed-rank test).

Attended and unattended item values better captured the vmPFC activity than chosen and unchosen item values

A number of previous studies have reported that the activity in vmPFC was representing the value of the chosen item minus the value of the unchosen item (Boorman et al., 2009; Jocham et al., 2012). As the participants had a tendency to choose the attended item (mean 78.7%, SEM 3.8% of trials), our findings could have been explained by the use of the value of chosen and unchosen items instead of the value of attended and unattended items. To examine this possibility, we performed a Bayesian model selection analysis (BMS) between these two models (See BMS in Materials and Methods). We found that the model with the value of attended and unattended items had the mean exceedance probability of 0.72 (SEM 0.01) within an independent ROI in vmPFC. The result

indicates that the value of attended and unattended items captures the activity in vmPFC better than the value of chosen and unchosen items.

Neural correlates of anDDM-derived time course of accumulator value in dmPFC and IPS

Using the value of attended item perturbed by noise as sampled evidence, anDDM accumulates the evidence in favor of each of the two alternatives until the decision criterion is met. Since we explicitly manipulated which of the two items the participant was attending to at each time point during the decision period, we could obtain a reliable estimate of the time course of accumulator value by simulating the model using the best fitting parameter values (Figure 2B; see Materials and Methods). We examined the neural correlates of the predicted time course of accumulator value to pin down areas implementing evidence accumulation to guide participants' choice (See GLM1 in Materials and Methods). We found that dmPFC and left IPS were involved in this process (Figure 5A; $p < 0.05$ whole brain corrected at cluster level for dmPFC and small volume correction (SVC) for left IPS, see Table S3 for other activated areas).

dmPFC and IPS exhibited model-predicted increased functional connectivity with vmPFC

Our model makes the following prediction about the functional connectivity between distinct brain regions: the areas of dmPFC and left IPS representing the time course of accumulator value should exhibit increased functional connectivity with vmPFC encoding the item value. This is because anDDM stipulates that the accumulators in the model accrue evidence sampled from item values. We conducted PPI analysis to test this prediction. We found that the areas of dmPFC and left IPS, which was identified to be involved in evidence accumulation process, showed increased functional connectivity with vmPFC during the decision period (Figure 5B; $p = 4.19 \times 10^{-4}$ for dmPFC and $p = 2.1 \times 10^{-3}$ for left IPS, Wilcoxon signed-rank test, see PPI analysis in Materials and Methods).

DISCUSSION

Here we investigated the computational and neural mechanisms of binary choice in the value-based decision making. We manipulated fixation-driven attention of the participants externally by alternating the item displayed on the screen every 1 second. This manipulation biased participants' choice behavior significantly such that the participants showed a tendency to choose the attended item. To describe participants' behavior quantitatively, we developed and fitted anDDM against the behavioral data. The model switches input flexibly depending on the item presented on the screen and the participants' attention is directed to. As a result, it could not only capture choice and RT but also attentional bias in participants' behavior. In the fMRI analysis we first confirmed that the item value representation in vmPFC reflected the effect of attentional manipulation: the activity in this area correlated with the value of attended item but not with the value of unattended item. Moreover, we directly correlated the time course of the accumulator value calculated using the fitted anDDM against BOLD signal. We found that this variable showed a significant correlation with the activity in dmPFC and left IPS.

Our findings that the participants' choice showed attentional bias and neural activity in vmPFC was modulated by fixation-driven attention are consistent with previous reports (Krajbich et al., 2010; Krajbich and Rangel, 2011; Lim et al., 2011; Towal et al., 2013; Tavares et al., 2017). However, the magnitude of attentional bias was greater in the present study: we found that anDDM fitted to participants' behavior produced a smaller attentional bias parameter value of $\alpha = 0$ compared to $\alpha = 0.3$ (Krajbich et al., 2010) and $\alpha = 0.36$ (Tavares et al., 2017). This implies that the value of unattended item was fully discounted and ignored in the sampled evidence while it was discounted and yet relevant in the previous reports. In the fMRI analysis we observed the neural effect of attentional manipulation that was consistent with the behavioral effect: the activity in vmPFC

showed a significant correlation with the value of attended item but not with the value of unattended item while it was modulated by both item values in a previous report (Lim et al., 2011). The observed discrepancy in the magnitude of attentional effect can be explained by the difference in the task structure: the attentional bias is stronger in a binary choice task when two items are presented sequentially and only one item is displayed on the screen compared to when two items are presented on the screen simultaneously. This is compatible with a recent report that the information acquired through peripheral vision has an effect on the choice process (Wästlund et al., 2018).

The time course of accumulator value predicted by anDDM was represented in dmPFC and left IPS. This result is consistent with previous reports that these areas of the frontoparietal network are involved in the evidence accumulation process in value-based decision making tasks (Basten et al., 2010; Hare et al., 2011; Gluth et al., 2012; Hunt et al., 2012; Rodriguez et al., 2015; Pisauro et al., 2017). However, our study is novel in terms of the analysis method that we directly correlated the time course of accumulator value against BOLD signal. We could utilize the time course of accumulator value because our model reflected which of the two items was more relevant for the accumulation process at each time point of the trial attained by explicitly controlling the attention of the participants. One advantage of correlating the time series of accumulator value directly is that the observed neural effect is expected to be less confounded by other effects such as trial difficulty (i.e., absolute difference of the two item values). In general, it is inevitable that the neural signature of accumulator related activity and trial difficulty are highly correlated because more difficult trials tend to result in a longer accumulation period to reach a decision. In previous studies, additional analysis methods such as PPI and BMS were combined to dissociate these two effects. In the present task, the correlation between the two signals was indeed lower ($r=-0.624$) than a previous report ($r=-0.930$; Hare et al., 2011). We can decrease the correlation further by increasing behavioral variability by increasing the number of trials in the experiment. In all, our result provides

additional evidence that the areas of the frontoparietal network are involved in the evidence accumulation process to guide our choices.

An important avenue to pursue in future research will be to elucidate the distinct role of dmPFC and IPS in value-based decision making. One possibility is that IPS is involved in the accumulation of evidence per se as analogous to the involvement of lateral intraparietal (LIP) neurons in this process in non-human primates (Shadlen and Newsome, 2001; Roitman and Shadlen, 2002) while dmPFC is encoding the consequence of choice process utilized to trigger a motor response. The idea comes from previous reports that the activity in dmPFC reflected the post decision signal of the value of unchosen item minus the value of chosen item (Wunderlich et al., 2009; Lopez-Persem et al., 2016). Moreover, another study combining methodologies from fMRI and EEG showed that the decision output signal of chosen value propagated from parietal to frontal region during a trial before the value difference signal was observed in dmPFC only in the later period of the trial (Larsen and O'Doherty, 2014). We examined this by including both the decision output signal and the time course of accumulator value in the same GLM (Incorporated regressors (7) and (8) of GLM1 into GLM4; see Materials and Methods). We found that the clusters of dmPFC and left IPS remained significantly correlated with the time course of accumulator value. The result supports the observation from previous reports that the evidence accumulation is implemented in parallel in these areas of the frontoparietal network in value-based decision making (Hare et al., 2011; Polanía et al., 2014, 2015). This parallel evidence accumulation process is also supported by the result of our PPI analysis that both dmPFC and left IPS showed increased functional connectivity during the decision period with vmPFC where the sampled evidence from the item value is represented.

In conclusion, the present study investigated the neural computations underlying binary choice in the value-based decision making domain. In a novel task, we manipulated the attention of the

participants by sequentially presenting the two items under comparison. We directly correlated the time course of accumulator value predicted by the computational model we developed that captured attentional bias in participants' choice behavior. Our results corroborate the involvement of vmPFC in representing input to the accumulators reflecting the effect of fixation-driven attention, and dmPFC and IPS in implementing the evidence accumulation process to guide the decision.

FIGURES AND TABLES

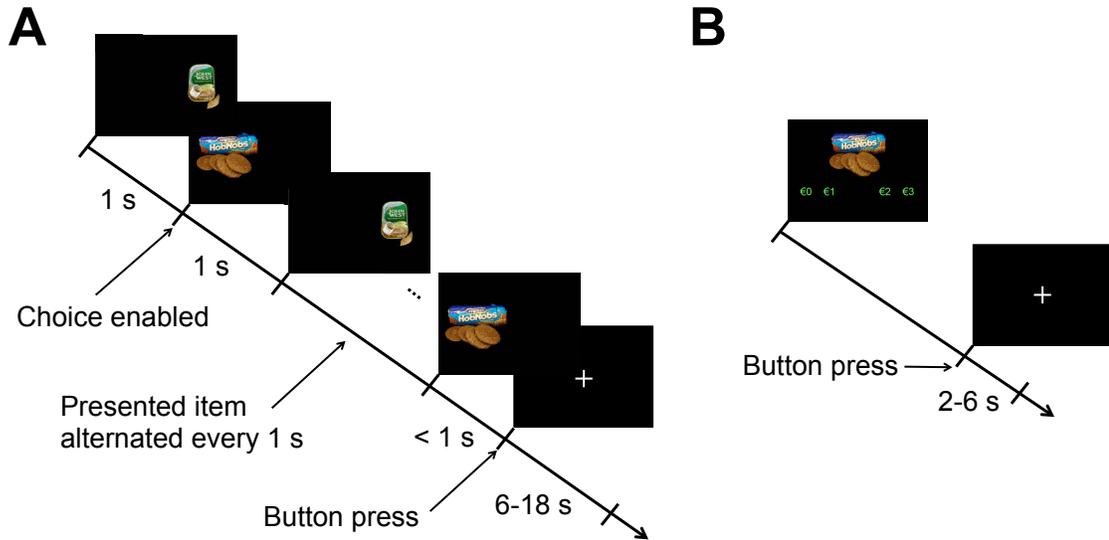


Figure 1. Illustration of task structure.

A. In the choice phase, participants engaged in 65 trials of binary choice with a free reaction time (RT). Two items were displayed on the screen sequentially until the participants indicated their choices by the button press. The presented item and its location on the screen (left or right) alternated every 1 second. The choice was enabled at the onset of second item presentation and the participants could choose either of the two items irrespective of the on-screen item at the time of button press.

B. In the pricing phase, participants reported their WTP for 130 items used in the choice phase by an integer amount between 0 and 3 euros. There was no RT restriction in each trial.

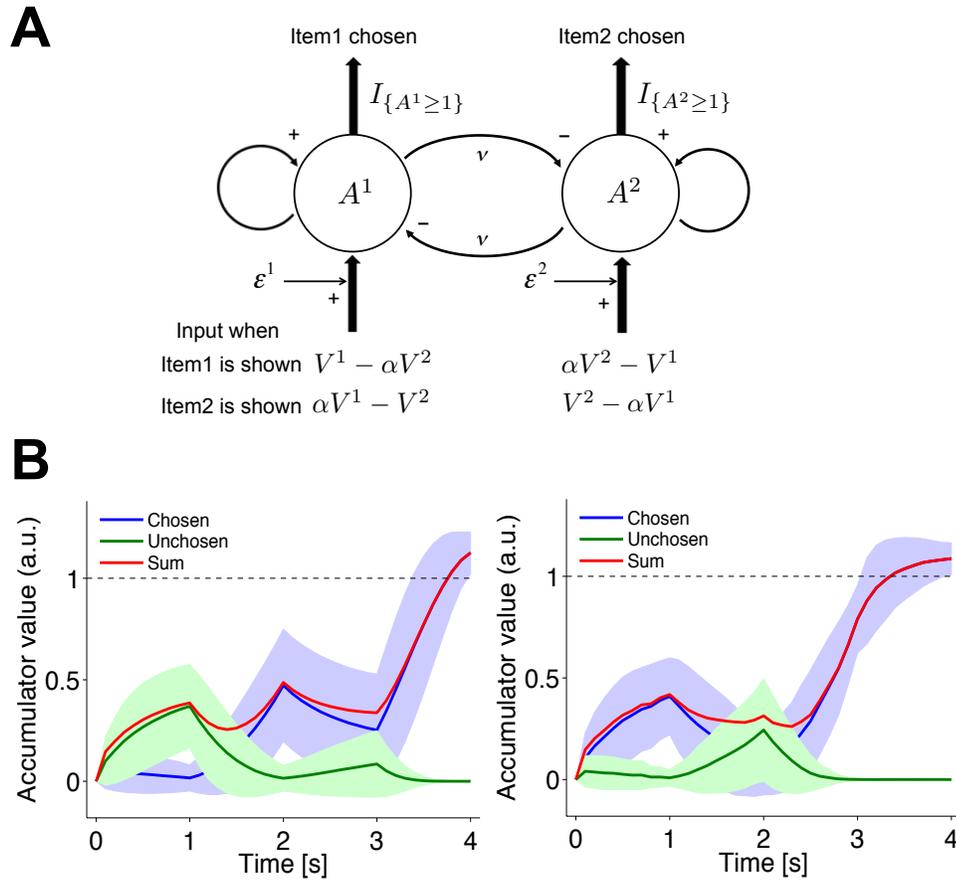


Figure 2. Computational model.

A. A graphic representation of attentional neural drift diffusion model (anDDM). Accumulator that accrue evidence supporting the choice of the firstly presented item (item1) and the secondly presented item (item2) are depicted by A^1 and A^2 , respectively. The sampled evidence serving as the input to the accumulator is the value difference of the two items under comparison and it changes according to which item is displayed on the screen at the time of accumulation. The effect of unattended item on the input is discounted by the attentional bias parameter α . The input is then perturbed by a Gaussian noise ε before being added to the accumulator. The accumulators have mutual inhibitory connections and the magnitude of inhibition is controlled by the parameter v . The choice of either of the two items is made when the corresponding accumulator value reaches the threshold of +1.

B. The time course of accumulator value prescribed by anDDM for trials with item value pair (1,2). (Left) Estimated time course when item2 with value 2 is chosen between 3-4 seconds of RT; the item is chosen when it is attended. (Right) Estimated time course when item1 with value 1 is chosen between 3-4 seconds of RT; the item is chosen when it is unattended. To obtain these time courses, we simulated the model 100,000 times using the best fitting parameters. The simulation runs were categorized by the combination of choice (item1 or item2 chosen) and RT (1 second duration bins from 1-2 seconds to 9-10 seconds). The time courses in each category were averaged at each time point of the simulation to give the average time course of accumulator value corresponding to the chosen item (blue), unchosen item (green), and sum of the two (red). Shaded regions represent standard deviation (omitted for the sum for display purpose). These raw time courses were convolved with canonical HRF and used as a regressor of interest in GLM for fMRI analysis (See GLM1 in Materials and Methods).

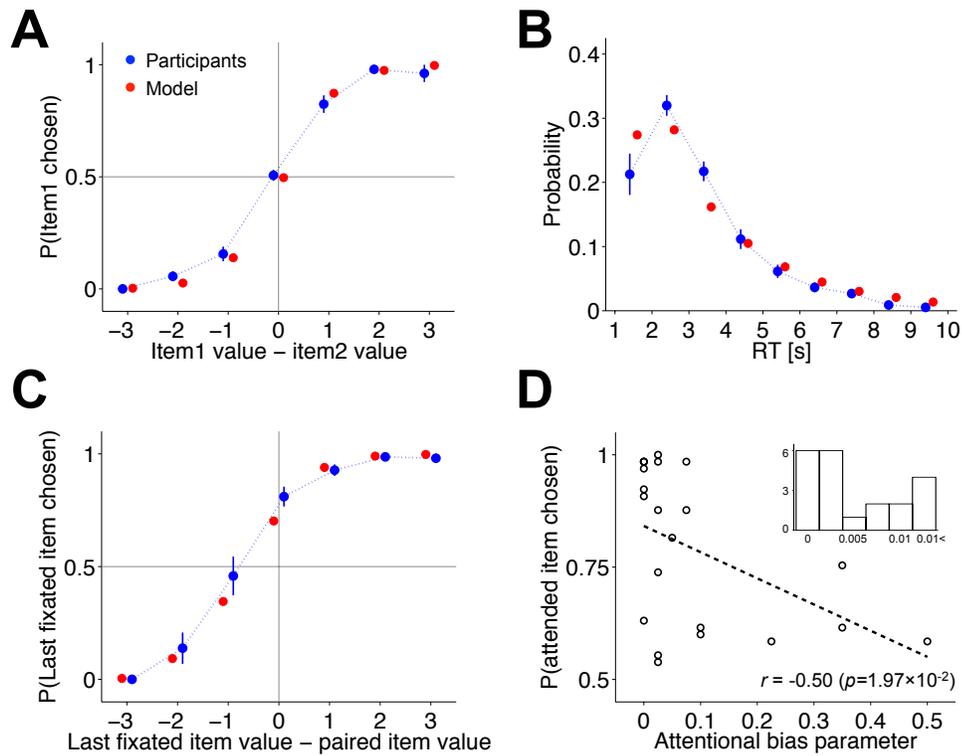


Figure 3. Behavioral results.

A. The probability of choosing item1 binned by the value difference between item1 and item2. Blue circles represent participants' behavior. Bars denote SEM across participants. The behavior of the fitted model is depicted in red circles.

B. RT distribution across participants. RT was sorted by 1-second duration bins from 1-2 seconds to 9-10 seconds. Bars denote SEM across participants. Blue: participants, Red: model.

C. The probability of choosing the last fixated item binned by the value difference between the last fixated item and paired item. Bars denote SEM across participants. Blue: participants, Red: model.

D. Across participants' correlation between the probability of choosing attended item and attentional bias parameter from model fitting to each participant. Circles represent each participant. Dashed line represents best-fit linear slope. The correlation coefficient r and associated p -value are noted (Spearman's rank correlation). Inset shows the histogram of the fitted attentional bias parameter values across participants.

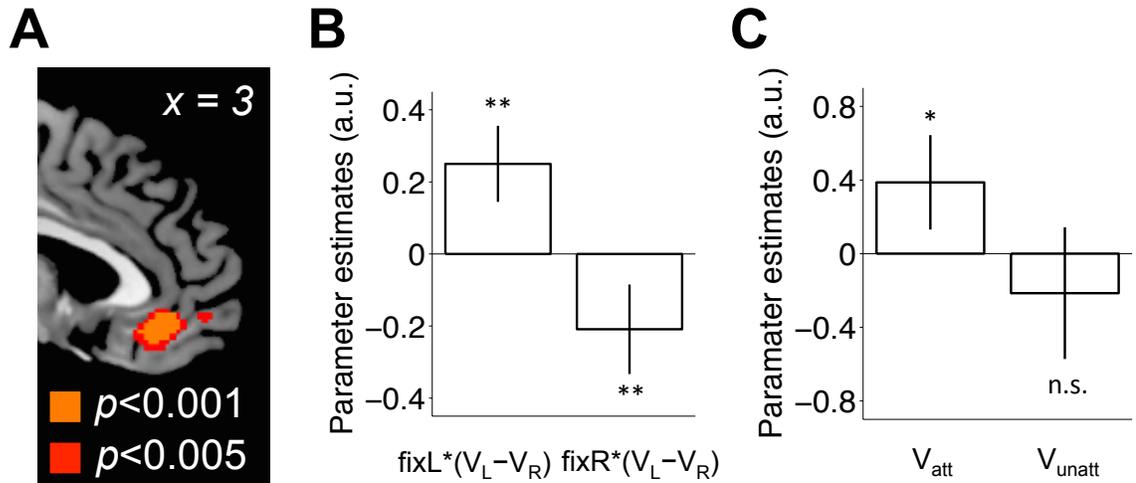


Figure 4. Fixation-driven attentional effect on the stimulus value representation in vmPFC.

A. Activity in vmPFC correlated significantly with the value of item presented on the screen ($p < 0.05$ whole brain corrected at cluster level; see Table S3 for other activated areas). Activation map is thresholded at $p < 0.001$ (orange) and $p < 0.005$ (red) for display purpose.

B. Parameter estimates within an independent ROI in vmPFC for the effect of value difference between left and right item for each side of fixation on the screen (**: $p = 7.1 \times 10^{-3}$ and $p = 9.6 \times 10^{-3}$ for fixL and fixR, respectively, Wilcoxon signed-rank test).

C. Parameter estimates within an independent ROI in vmPFC for the effect of value of attended (V_{att}) and unattended item (V_{unatt}) (*: $p = 2.50 \times 10^{-2}$, n.s.: $p = 0.38$, Wilcoxon signed-rank test).

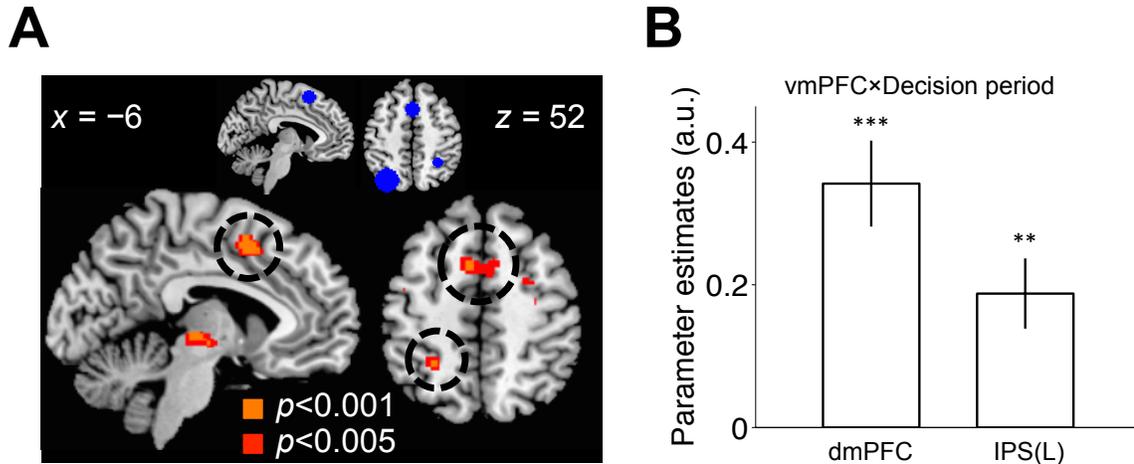


Figure 5. Neural correlates of the time course of accumulator value and PPI result.

A. Areas significantly correlated with the time course of accumulator value predicted by anDDM. dmPFC and left IPS are circled by the black dashed line ($p < 0.05$ whole brain corrected at cluster level; see Table S3 for other activated areas). Activation map is thresholded at $p < 0.001$ (orange) and $p < 0.005$ (red) for display purpose. Inset displays clusters in dmPFC and bilateral IPS that showed significant correlation with accumulator value signal identified in a previous study (Hare et al., 2011). Each cluster was approximated by a sphere surrounding peak activation coordinates with the radius computed from the cluster's voxel size reported.

B. Functional coupling between each of the two regions involved in evidence accumulation process (dmPFC and left IPS) and the region representing sampled evidence from item value (vmPFC) during the decision period (***: $p = 4.19 \times 10^{-4}$, **: $p = 2.1 \times 10^{-3}$, Wilcoxon signed-rank test).

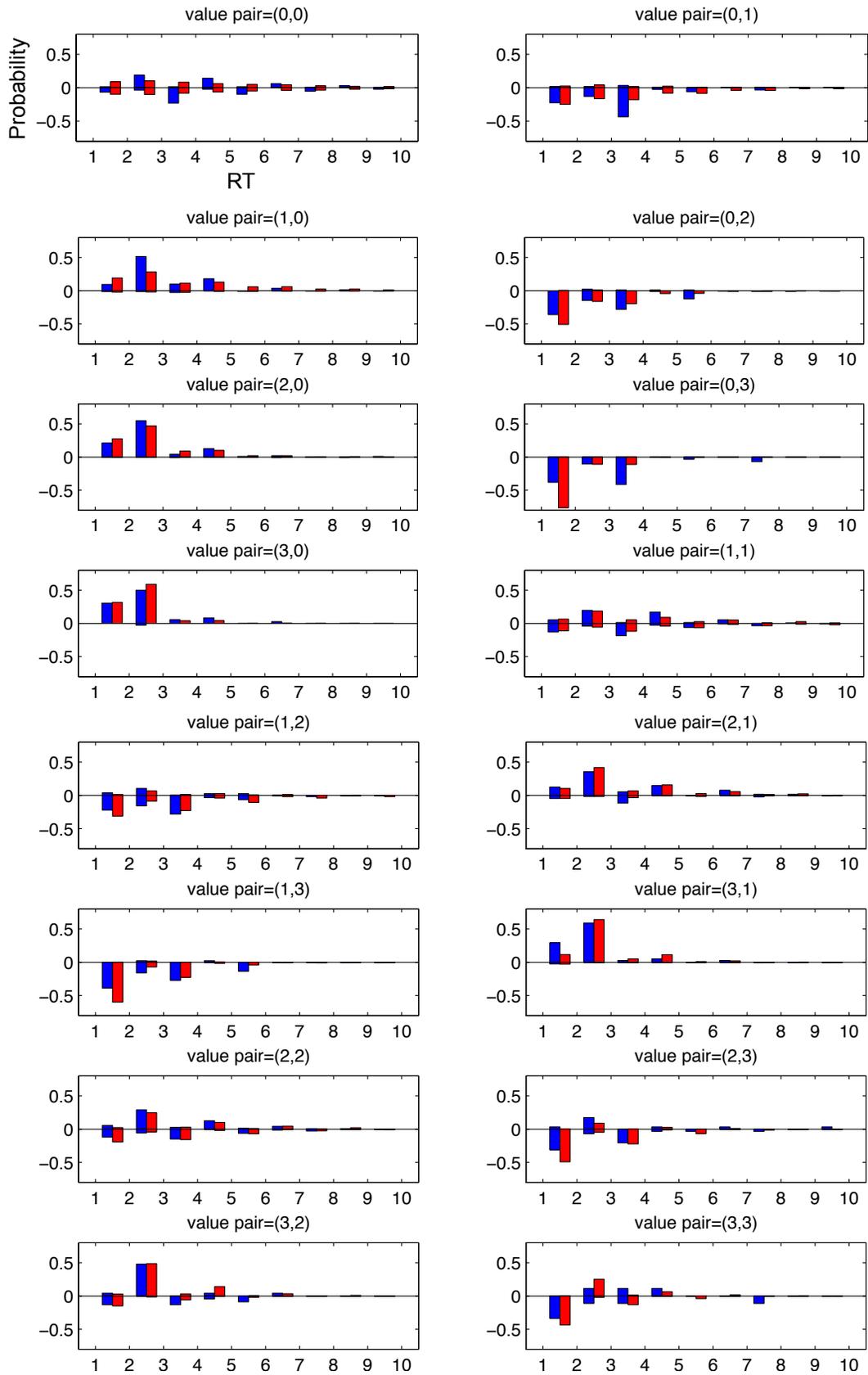


Figure S1. Related to Figure 3A and Figure 3B. Joint distribution of choice and RT of participants' behavior.

The joint probability distribution of choice and RT for each of the 16 possible value pairs. The choice of item1 and item2 are depicted in the positive and negative domain of the y-axis, respectively. RT was sorted by 1-second duration bins from 1-2 seconds to 9-10 seconds. Blue: participants, Red: model.

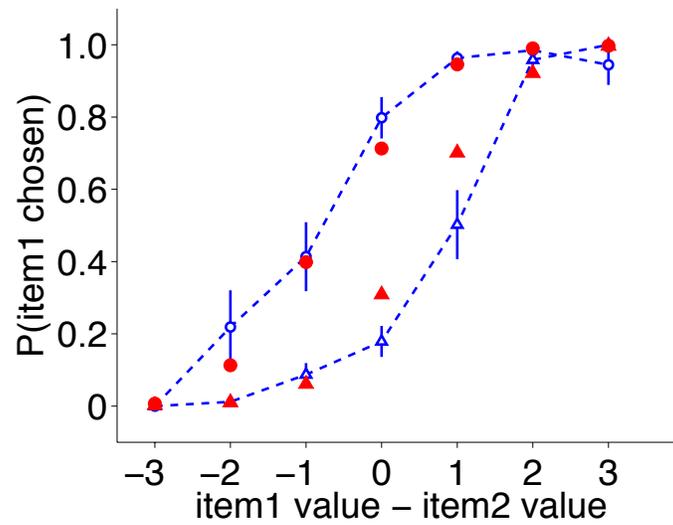


Figure S2. Related to Figure 3C. Attentional bias in participants' behavior.

The probability of choosing item1 binned by the value difference between item1 and item2 when item1 (circle) and item2 (triangle) are displayed at the time of choice, respectively. Blue symbols represent participants' behavior with bars denoting SEM across participants. The behavior of the fitted model is depicted in red symbols.

	$-LL$	α	η	ν	δ	κ
First grid		0:0.1:0.5	0.005:0.005: 0.04	0:0.1:1.0	2.0:1.0: 7.0×10^{-4}	0:0.1:1.0
Best parameters	2794	0	0.02	1.0	4.0×10^{-4}	0.5
Second grid		0:0.05: 0.15	0.0175:0.0025: 0.0225	0:0.05:1.0	3.5:0.05: 4.5×10^{-4}	0.45:0.05: 0.65
Best parameters	2778	0	0.0175	1.0	4.5×10^{-4}	0.6
Third grid		0:0.025: 0.075	0.01625: 0.00125:0.01875	0.475: 0.025:1.0	3.75:0.025: 4.75×10^{-4}	0.475:0.025: 0.625
Best parameters	2773	0	0.0175	1.0	4.75×10^{-4}	0.575

Table S1. Related to Figure 3. Details of the model fitting procedure.

We iterated the steps involved in the grid search procedure three times by narrowing the range of parameters tested to the values around the best fitting parameter values from the previous iteration. Here, we provide the information of the parameter grid used in each iteration and the resulting best fitting parameter values.

	Coefficient	Z score	<i>p</i> value
Intercept	1.68×10^{-3}	0.01	0.99
Value difference	2.54	10.91	<0.001
Sum of item values	0.02	0.45	0.65
Interaction	-0.35	-4.69	2.79×10^{-6}

Table S2. Related to Figure 3. Effect of item values on participants' choice behavior.

The result of mixed effect logistic regression using the participant group data is presented. Dependent variable is 1 for the choice of item1 and 0 for the choice of item2. Independent variables are the value difference of two items (item1 value minus item2 value), the sum of two item values, and their interaction term.

Contrast	Region	Hemi	Peak MNI coordinates			<i>t</i> score	Cluster size (voxels)
			<i>x</i>	<i>y</i>	<i>z</i>		
Stimulus value							
	Medial frontal gyrus (vmPFC)	L	-6	36	-12	6.03	660
	Parahippocampal gyrus	R	26	-24	-18	5.34	196
	Superior temporal gyrus	L	-40	-60	24	4.25	492
	Superior frontal sulcus	L	-26	32	46	3.95	289
Button press			<i>x</i>	<i>y</i>	<i>z</i>		
Left > Right	Cerebellum	L	-14	-52	-22	8.10	1683
	Precentral gyrus	R	38	-22	64	7.77	2948
	Pulvinar	R	20	-26	4	5.05	233
	Insula	R	40	-10	22	4.41	356
Right > Left	Precentral gyrus	L	-36	-26	64	8.46	3185
	Cerebellum	R	18	-52	-22	6.44	756
Time course of accumulator value			<i>x</i>	<i>y</i>	<i>z</i>		
	Fusiform gyrus	R	30	-58	-8	6.98	507
	Fusiform gyrus	L	-26	-60	-8	5.59	414
	Thalamus	L	-4	-22	0	5.20	387
	Medial central gyrus (dmPFC)	L	-6	10	48	5.17	250
	Precentral sulcus	R	32	-6	46	4.78	232
	Intraparietal sulcus (IPS)	L	-28	-50	50	3.22	51 (SVC)

Table S3. Related to Figure 4A and Figure 5A. MNI coordinates of peak activation voxels from the contrasts in GLM1.

Reported activation clusters survived $p < 0.05$ whole brain correction at cluster level with the height threshold of $p < 0.005$ and the minimum spatial extent of 191 voxels. The significance of the cluster in left IPS representing the time course of accumulator value was tested by small volume correction (SVC) in a 10mm sphere surrounding the mean peak activation coordinates from previous studies ($x = -34, y = -55, z = 44$).

*Chapter 3***COMPUTATIONAL ACCOUNT OF
TEMPORAL CHANGE DETECTION IN THE HUMAN BEHAVIOR**

Adachi R, Suzuki S, O’Doherty JP. Neural computations mediating temporal change detection in the human brain.

R.A. and J.P.O. designed the experiment. R.A. and S.S. collected the data. R.A. performed data analysis with feedback from S.S. and J.P.O. R.A. and J.P.O. wrote the manuscript. J.P.O. supervised the research.

INTRODUCTION

Our brains are constantly receiving streams of information varying in temporal structure. Detecting changes in the temporal structure of events in the environment can enable one to make inferences about changes in the underlying causal processes operating in the world. For instance, such a capability could underlie the ability of an animal to detect whether a predator is approaching or departing their hiding place, or enable human traders to determine whether a financial market is about to crash by keeping track of changes in the frequency of trades (De Martino et al., 2013b). Thus, the capacity to detect changes in temporal information streams is a crucial part of the cognitive machinery required by humans and other animals to produce adaptive behavior in a dynamical and rapidly changing environment.

A fundamental open question in cognitive and computational neuroscience concerns the nature of the computational algorithm that is actually used by the brain to infer changes in a generative process when there is no directly observable information explicitly signaling that change has

occurred. Here, we address this question with regard to change detection in the temporal domain by comparing and contrasting different computational strategies that can solve this particular computational problem. We analyze behavioral data in this chapter and fMRI data in Chapter 4 from human participants as a means to determine which of these computational strategies provides the best account of how temporal change detection is implemented at the neural levels in humans. Computational approaches to understanding the psychological and neural mechanisms of change detection more broadly have tended to utilize two distinct approaches: the first of these involves taking a normative perspective in which Bayesian models are constructed for the purpose of describing the theoretically optimal manner in which to detect changes in a given context (Yu and Dayan, 2005; Behrens et al., 2007; Wilson et al., 2010; Payzan-LeNestour et al., 2013). However, while such normative approaches have the merit of optimality, it is unclear to what extent the brain can plausibly implement a full Bayesian inference scheme involving the encoding of full probability distributions, particularly under situations that might require the simultaneous maintenance and updating of many-fold belief distributions over numerous possible environmental states. Furthermore, well-characterized instances of capacity constraints in human cognition might auger against the likelihood that humans can implement the computationally expensive process of a full Bayesian solution to change detection in anything but the simplest task situations.

As a result of these limitations, an alternative approach in the literature has involved developing strategies that can do reasonably well at detecting change points without incurring the representational and computational costs of the more optimal Bayesian approaches. These approaches have utilized simple model-free algorithms such as a delta-rule or other heuristics (Krugel et al., 2009; Nassar et al., 2010; Summerfield et al., 2011; McGuire et al., 2014).

The goal of the present study is to conduct a comprehensive assessment of which computational mechanism could best describe the computations involved in this form of change detection, by formally comparing optimal Bayesian models of change detection to simpler yet computationally more tractable approximate strategies.

To address this question we developed a novel behavioral task and ran participants through this task while they were scanned with fMRI. In the task, participants observed a sequence of images that contained abrupt changes in the underlying frequency of image presentations. The observed sequence was corrupted by noise, and the participants were instructed to report a change if they attributed the perceived change to a change in the underlying statistics of the sequence, rather than to mere stochastic noise.

MATERIALS AND METHODS

DATA COLLECTION

Participants

22 right-handed participants performed the experiment (18–35 years old; 12 males). The participants were prescreened to exclude those with a prior history of neurological or psychiatric illness. The data from one participant was excluded from both behavioral and fMRI analyses because of a large head displacement (>8 mm) during fMRI image acquisition. We also excluded the fMRI data of 1 out of 4 runs from two participants because of scanner failure in these runs. The Institutional Review Board of California Institute of Technology approved the experiment and all participants gave informed consent.

STATISTICAL ANALYSIS

Performance measure of participants' behavior

We measured the participants' behavioral performance using a distance metric – the “performance f measure” between participants' button presses and actual change points calculated as follows: first, we computed the precision and sensitivity of the participant's behavior. Precision and sensitivity were defined as $TP/(TP+FP)$ and $TP/(TP+FN)$, respectively (TP: true positive; FP: false positive; FN: false negative). TP is a count of the number of times a participant's button press was within a 5-second window after an objective change point (task determined) occurred. FP is a count of the number of times that a participant's button press occurred outside of a 5-second window after an objective change point occurred or when a participant's button press was within the time window but there was another button press in the same window already categorized as being a TP. FN is a count of the number of times when there was no participant's button press within a 5-second time window. By definition, precision described the fraction of the participant's button presses that conformed to change points while sensitivity captured the fraction of change points that were correctly captured by the participant. Finally, we put equal weight on precision and sensitivity and calculated the performance f measure as:

$$f \text{ measure} = 2/(1/\text{precision} + 1/\text{sensitivity}). \quad (3)$$

Comparison against a random agent

To exclude the possibility that the participants' behavior was random, we defined the behavior of agents that acted randomly and compared the agents' behavior to participants' behavior. The random agent made a random response in each bin using the rate of button presses equal to what was observed in individual participant's behavior. We simulated the random agent model for each participant 5000 times and evaluated the performance using the performance f measure.

Computational models

We tested three computational models of varying degrees of sophistication to describe participants' behavior quantitatively. Two of these models employed Bayesian methods. One was an implementation of the Bayesian online change point detection model (BOCPD, Adams and MacKay, 2007), and another was an implementation of the dynamic belief model (DBM, Yu and Cohen, 2009). We augmented each of the above original models by adding a decision rule to detect change points in our task. The third model we tested was a computationally frugal one that in contrast to the above models did not require the encoding of probability distributions or Bayesian updating, which we named the temporal averaging model (TAM). In all of these models, we utilized 120 bins with 0.25-second duration for each 30-second trial used in the data generating process of image sequence and the decision variable employed to detect change points was computed at this time resolution. The data in each bin was denoted by x and took the value of 1 when a sushi piece was produced at the bin and 0 when no sushi piece was produced at the bin. The index of each bin was subscripted by n ($1 \leq n \leq 120$). Below, we describe each of the three computational models in detail.

Bayesian Online Change Point Detection model (BOCPD)

The model tracked the posterior distribution over run length r , which was the number of bins since the most recent change point in the past. Suppose that the model was calculating the posterior run length distribution at the n th bin and let z ($z < n$) denote the bin number of the most recent change point. The value of z was initialized to $z = 0$ at the start of a trial. The range of run length the model considered increased as the bin number n increased (i.e., $r_n \in [0, n - z]$) and each run length r corresponded to a hypothesis that $(n - r_n)$ th bin was a change point. The posterior run length distribution at the n th bin given the past observations ($x_{1:n}$) was obtained as

$$P(r_n|x_{1:n}) = \frac{P(r_n, x_{1:n})}{P(x_{1:n})} = \frac{P(r_n, x_{1:n})}{\sum_{r_n} P(r_n, x_{1:n})},$$

where its numerator was calculated as

$$\begin{aligned} P(r_n, x_{1:n}) &= \sum_{r_{n-1}} P(r_n, r_{n-1}, x_{1:n}) = \sum_{r_{n-1}} P(r_n, x_n | r_{n-1}, x_{1:n-1}) P(r_{n-1}, x_{1:n-1}) \\ &= \sum_{r_{n-1}} P(r_n | r_{n-1}) P(x_n | r_{n-1}, x_{1:n-1}) P(r_{n-1}, x_{1:n-1}). \end{aligned}$$

Therefore, the joint distribution of run length and data had a recursive form and was computed by using the following two quantities:

$$P(r_n | r_{n-1}) = \begin{cases} h, & r_n = 0 \\ 1 - h, & r_n = r_{n-1} + 1 \\ 0, & \text{otherwise} \end{cases}$$

and

$$\begin{aligned} P(x_n | r_{n-1}, x_{1:n-1}) &= \int_0^1 P(x_n, \varepsilon | r_{n-1}, x_{1:n-1}) d\varepsilon = \int_0^1 P(x_n | \varepsilon) P(\varepsilon | x_{n-1}^r) d\varepsilon \\ &= \frac{\int_0^1 P(x_n | \varepsilon) P(x_{n-1}^r | \varepsilon) P(\varepsilon) d\varepsilon}{\int_0^1 P(x_{n-1}^r | \varepsilon') P(\varepsilon') d\varepsilon'}. \end{aligned}$$

Here, the abbreviation x_{n-1}^r denoted the past observations conforming to run length r_{n-1} at $(n-1)$ th bin (i.e., $x_{n-1}^r = \{x_{n-1}, x_{n-2}, \dots, x_{n-1-r_{n-1}}\}$). The Bernoulli probability of image presentation (production of a sushi piece) in each bin was represented by ε . We assumed that the unconditional

probability of a change point in each bin was given by a constant hazard rate h , which was equivalent to prescribing a geometric distribution over the number of bins between the two successive change points. Thus, the model correctly specified the form of generating process of change points. After obtaining the posterior run length distribution, the model computed the change point probability (CPP) by summing all the probability masses in the posterior run length distribution conforming to the existence of a change point within the last T seconds. The optimal value of this parameter was $T = 5$ because correct and incorrect button presses were rewarded and punished by an equal amount and any button presses after 5 seconds from a change point were deemed incorrect. We introduced this parameter to the model to take into account the following two possibilities: (1) the participants might have been using maladaptive values that varied across participants, and (2) even if the participants were subjectively using the optimal values, there is heterogeneous variability associated with the measurement of time (Jazayeri and Shadlen, 2010). The CPP served as the decision variable (DV) of this model and was formally written as

$$DV_n = \sum_{k=0}^{\min(\text{floor}(4T), n-z-1)} P(r_n = k | x_{1:n}). \quad (4)$$

If the decision variable was greater than the threshold value θ_n , a change was detected and z was updated to the location of the change point. We tested both a fixed threshold ($\theta_n = \theta$) and a collapsing bound having the form of Weibull function as

$$\theta_n = \theta - (1 - e^{-(n-z/\alpha)^\beta})(1 - \delta)\theta. \quad (5)$$

The initial value of the threshold after a change point was set to θ and three parameters α , β , and δ controlled the shape of the decay. The collapsing bound was tested in all the three models because the computed decision variable was noisy for some period after the start of a trial and after a change point because only a small number of data points were available in the new regime.

We tested the following parameter sets: P1) h, θ , P2) h, θ, T , P3) $h, \theta, \alpha, \beta, \delta$, and P4) $h, \theta, T, \alpha, \beta, \delta$. When not varied, we set T to its optimal value of $T = 5$. For P1 and P2, the threshold was fixed throughout the trials. After detecting a change point, the model needed to reset the current run length to compute CPP in the subsequent bins. We tried three methods for this: M1) reset to 0, M2) set to posterior mean, and M3) set to posterior mode. M1) assumed a change occurred at the current bin, and M2) and M3) assumed that a change occurred at a bin in the past for the number of bins equal to posterior mean and mode of the run length distribution, respectively. Finally, we tested two distributions of unconditional Bernoulli probability $P(\varepsilon)$ of image presentation in each bin: D1) a true prior which had equal point masses at $\varepsilon = 0.1, 0.15, 0.45, 0.5$, and D2) a uniform prior over $[0, 0.6]$. In all, we tested 24 sub-models (P1-4*M1-3*D1-2).

Dynamic Belief Model (DBM)

The model updated the posterior distribution over the Bernoulli probability of an image presentation (production of a sushi piece) $P(\varepsilon)$ after observing the data in each bin, while taking into account the possibility of a change point occurrence at the bin. It took the form of Bayesian updating as follows:

$$\begin{aligned} P(\varepsilon_n | x_{1:n}) &\propto P(x_n | \varepsilon_n) P(\varepsilon_n | x_{1:n-1}) = P(x_n | \varepsilon_n) ((1-h)P(\varepsilon_{n-1} = \varepsilon_n | x_{1:n-1}) + hP(\varepsilon_n)) \\ &= \varepsilon_n^{x_n} (1 - \varepsilon_n)^{1-x_n} ((1-h)P(\varepsilon_{n-1} = \varepsilon_n | x_{1:n-1}) + hP(\varepsilon_n)). \end{aligned}$$

Here, h was the probability that the observed data in the n th bin came from a Bernoulli distribution with mean ε_n , which was different from its value in the previous bin ε_{n-1} . Using the posterior distribution over Bernoulli probability incorporating the possibility of a change point, the model computed the maximum absolute difference of the posterior mean or mode (M1 or M2) between the current bin and any bin within the last T seconds after the most recent change point z . The value of z was initialized to $z = 1$ at the start of a trial. The decision variable (DV) was written as

$$DV_n = \max_{1 \leq i \leq \min(\text{floor}(4T), n-z)} |post_n - post_{n-i}|. \quad (6)$$

A change point was detected if $DV_n \geq \theta_n$ and the value of z was updated to the location of the change point (i.e., $z = n$). We tested both a fixed threshold ($\theta_n = \theta$) and collapsing bound having the form of Weibull function (See Equation (5) under BOCPD for description). In addition to h, θ, T , and three parameters α, β, δ related to the shape of threshold, another parameter k was considered. The parameter k changed the threshold value based on the observed frequency of image presentation as

$$\theta_n = \theta_n + k \times \min\{post_n, post_{n-i}\}, \tilde{i} = \underset{1 \leq i \leq \min(\text{floor}(4T), n-z)}{\text{argmax}} |post_n - post_{n-i}|.$$

A positive value of k complied with Weber's law, which states that any noticeable difference between two stimuli is proportional to the magnitude of the stimuli. We tested the following parameter sets: P1) h, θ , P2) h, θ, T , P3) h, θ, k , P4) h, θ, T, k , P5) $h, \theta, \alpha, \beta, \delta$, P6) $h, \theta, T, \alpha, \beta, \delta$, P7) $h, \theta, k, \alpha, \beta, \delta$, and P8) $h, \theta, T, k, \alpha, \beta, \delta$. We tried two distributions of unconditional Bernoulli probability $P(\varepsilon)$: D1) a true prior which has equal point masses at $\varepsilon = 0.1, 0.15, 0.45, 0.5$, and D2) a uniform prior over $[0, 0.6]$. In all, we tested 32 sub-models (P1-8*M1-2*D1-2).

Temporal Averaging Model (TAM)

This model compared the average image presentation rate in the two different time windows and reported a change if the absolute difference of the two quantities was greater than a threshold value θ_n . Because the computations it involves are arithmetic mean of stimulus values and subtraction, the model is highly generalizable to detecting changes in any domains. At the n th bin, the model first calculated the average frequency of image presentation (production of a sushi piece) per second from $(n - 4X + 1)$ th to n th bin (i.e., during X seconds) and from $\max(n - 4X - 4Y + 1, z)$ th to $(n - 4X)$ th bin (i.e., during up to Y seconds after the most recent change point z), which were denoted by \bar{X}_n and \bar{Y}_n , respectively. The value of z was initialized to $z = 1$ at the start of a trial. Second, it computed the decision variable (DV), which was the absolute difference between \bar{X}_n and \bar{Y}_n :

$$DV_n = |\bar{X}_n - \bar{Y}_n|. \quad (7)$$

The model detected a change point if $DV_n \geq \theta_n$ and the value of z was updated to the location of the change point (i.e., $z = n - 4X + 1$). We tested both a fixed threshold ($\theta_n = \theta$) and a collapsing bound having the form of a Weibull function (See Equation (5) in BOCPD for description). In addition to X, Y, θ and three parameters α, β, δ related to the shape of threshold, we considered a parameter k which reflected Weber's law (See description under DBM) and scaled the threshold value as $\theta_n = \theta_n + k \times \min\{\bar{X}_n, \bar{Y}_n\}$. We tested the following sub-models of four different parameter sets: P1) X, Y, θ , P2) X, Y, θ, k , P3) $X, Y, \theta, \alpha, \beta, \delta$, and P4) $X, Y, \theta, k, \alpha, \beta, \delta$.

Delta-rule model (DRM)

For completeness, we also evaluated the performance of a change detection algorithm based on a delta-rule process utilized in a number of previous studies of change detection in non-temporal domains (Nassar et al., 2010, 2012; McGuire et al., 2014). By utilizing a simple delta-rule, this model has been reported to achieve a significant reduction in computational complexity while attaining reasonable performance in describing participants' behavior even compared to an optimal Bayesian model. In this model, the probability of a change point (cp) at n th-bin given the observation x_n denoted by Ω_n is computed using Bayes' rule:

$$\Omega_n = p(cp|x_n) = \frac{p(x_n|cp)p(cp)}{p(x_n|cp)p(cp) + p(x_n|\sim cp)p(\sim cp)} = \frac{0.3h}{0.3h + \mu_n(1 - h)},$$

where h denotes the unconditional change probability at each bin, and μ_n denotes the current estimate of Bernoulli probability of sushi-piece production in each bin. $p(x_n|cp) = 0.3$ for two probable forms of prior distribution of Bernoulli probability: a true prior which has equal point masses at $\varepsilon = 0.1, 0.15, 0.45, 0.5$, and a uniform prior over $[0, 0.6]$. The Bernoulli probability μ_n is learned by delta-rule as

$$\mu_n = \mu_{n-1} + \alpha_n(x_n - \mu_{n-1}),$$

where α_n denotes the learning rate given by

$$\alpha_n = \frac{1 + \Omega_n \hat{r}_n}{1 + \hat{r}_n}.$$

Here, \hat{r}_n is the expected run length. Instead of updating the full probability distribution of run length as in BOCPD (See the description of BOCPD), the model tracks its expected value:

$$\hat{r}_{n+1} = (\hat{r}_n + 1)(1 - \Omega_n) + \Omega_n.$$

The model detects a change if the decision value $DV_n = \Omega_n$ is greater than the threshold value of θ_n . We tested a collapsing bound having the form of a Weibull function (See Equation (5) in BOCPD for description). Thus, the model has five parameters: h, θ , and three parameters α, β, δ related to the shape of threshold.

Although the performance of this model was significantly better than the random model, it fell short of the other three models on the sequence of binary stimuli in our task (Figure S10). For this reason, we do not consider this model further. A likely explanation for why the delta-rule model did not perform especially well for our task is that our temporal detection task differs substantially from the change detection tasks utilized in previous studies that have reported good performance for the delta-rule model. This is because the delta rule depends on generating a reliable estimate of the generative distributions from a relatively small number of observations. In the case of the present studies the observations are binary (i.e., drawn from a Bernoulli distribution), depicting whether the event has occurred or not, as opposed to being continuous where the event is a continuous variable drawn from a one-dimensional Gaussian distribution with a reasonably small variance. The individual binary observations utilized in the present tasks are thus less informative and noisier for the updating of the delta rule than the individual observations obtained in previous tasks in which this model variant has performed successfully. We think this is the primary reason this model does not perform adequately in our task. However, we acknowledge that it is possible that this model

variant could potentially be extended and elaborated on in subsequent work, which could potentially lead to it achieving better performance on the present task.

Random model

We also tested the performance of a random model as a benchmark. In this model, we used the response of the random agent described earlier who made responses randomly in each bin using the rate of button presses observed in individual participant's behavior. We simulated the random model for each participant 5000 times and evaluated the performance.

Model fitting to participants' behavior

We calibrated each of the three computational models to predict the timing of button presses in participants' behavior. For each participant, we obtained the parameter values of each model by minimizing the f measure between the participant's button presses and the button presses simulated by the model. The constituents of the f measure, precision and sensitivity, were defined as $TP/(TP+FP)$ and $TP/(TP+FN)$. TP corresponds to the total number of instances where a model-produced button press was within a 10-second window around the button press made by an actual participant. FP corresponds to the total number of instances when a model-produced button press did not fall within a 10-second window of the participant's button presses or when a model produced button press occurred within a 10-second time window of a participant's actual button press but when there was also another button press in the same window already categorized as being a TP. FN corresponds to the total number of instances when there was no model button press within a 10-second time window. Therefore, precision described the fraction of the model's button presses that conformed to the participant's button while sensitivity captured the fraction of the participant's button presses that were correctly captured by the model. Finally, precision and sensitivity were combined to compute the overall f measure using Equation (3).

In the model fitting, we pooled participant's behavior across runs because there were no significant differences in the behavior across runs (Figure S5). This confirmed that participants could not learn to better detect a change point without feedback regarding their performance. We adopted hold-out validation by using odd-numbered trials as the training set and even-numbered trials as the validation set. For each sub-model and participant, we first obtained parameter values that maximized the f measure defined above in the training set and then computed the f measure in the validation set by simulating the sub-model using the parameter values obtained in the previous step. We refer to the f measure resulting from hold-out validation as the "predictive f measure". We assessed the performance of each sub-model by averaging the predictive f measure across participants. Finally, we compared and selected the best sub-model for each of the three models. Once the best sub-model for each model was selected, we calibrated the model for each participant to all the trials across runs. We refer to the resulting f measure as the "fitted f measure".

We checked how dependent the result of model fitting is to the size of time window chosen to calculate the f measure. To this end, we fit the best sub-model for each of the three model classes using 6-second, and 14-second time windows alongside the 10-second window used in the main analyses. Consistent with the result found using a 10-second time window, we found no significant differences between the goodness of fit of the different model class (6-second time window: $t(20)=0.07$, $p=0.94$ between BOCPD and DBM, $t(20)=0.44$, $p=0.66$ between DBM and TAM, and $t(20)=0.43$, $p=0.67$ between TAM and BOCPD, paired-sample t test; 14-second time window: $t(20)=1.06$, $p=0.30$ between BOCPD and DBM, $t(20)=1.61$, $p=0.12$ between DBM and TAM, and $t(20)=0.59$, $p=0.56$ between TAM and BOCPD, paired-sample t test). Thus, our model fitting results are robust to arbitrary choices of time window and not an artifact of a specific time window.

Note that for model fitting we used the method described above instead of a likelihood maximization method. The maximum likelihood method is inappropriate to use here because it would test for the accuracy of the model predictions for single responses emitted at each unit of time (i.e., at a single-bin level of temporal precision), which is not a reasonable test for any of the three models, as no model could be expected to predict response accuracy to that arbitrary level of precision.

DATA AVAILABILITY

The behavioral data has been deposited in GitHub: <https://github.com/RyoAdachi/tcd>.

CODE AVAILABILITY

Key analysis codes written in MATLAB and behavioral analysis results have been deposited in GitHub: <https://github.com/RyoAdachi/tcd>.

RESULTS

Temporal change detection task: “The sushi chef task”

In the experiment, participants observed a chef making sushi pieces represented by a sequence of images while being scanned with fMRI (main task; Figure 6). Participants were instructed that there are two chefs; one is diligent and productive while the other is lazy and unproductive. Participants were also told that the diligent chef has a higher average rate of production than the lazy one although the exact number of production by each chef fluctuates from time to time. Participants were also informed that while both of these two chefs work in the kitchen only one of them is working at a particular time, and that the chef who is making sushi pieces might or might not change from one to the other during a given trial. The participants were instructed to press a button

on the keypad using their right finger anytime they thought that a changeover had occurred between the two chefs. We instructed the participants that there could be either zero, one or multiple switches between the chefs in each trial and no information was provided to participants regarding the timing of changes. The specific image used to denote the sushi pieces and the location of the sushi pieces on the screen was kept fixed throughout the trials for all the participants and thus the only relevant information in the main task was the timing of image presentations. Each trial began after the presentation of a white fixation cross (2-6 seconds, mean 4 seconds). Each trial lasted for 30 seconds.

In the main task, the timing of production of sushi pieces was determined as follows: in a 30-second trial, we set up 120 bins each having the duration of 0.25 second and a sushi piece was presented in each bin with probability p for 0.125 second around the middle of the time bin. We manipulated the value of p to achieve two levels of difficulty; p jumped between 0.1 and 0.5 and between 0.15 and 0.45 at a change point in easy and hard conditions, respectively. The numbers of trials from easy and hard conditions were balanced and whether a sequence started with the low or high value of p was randomized within each condition. The change points were prescribed as follows: there was no change point during the first 20 bins (i.e., 5 seconds) at the start of the trial and then a change point location was drawn from a geometric distribution with a mean of 60 bins (i.e., 15 seconds). Once a change occurred, there was no change point during the following 20 bins (i.e., 5 seconds) and the next change point was drawn from the same geometric distribution. We omitted change points during the last 20 bins (i.e., 5 seconds) of a trial because the participants would not be able to detect these change points on the basis of a relatively small number of data points available thereafter until the end of the trial.

We also included a control condition designed to account for the effects of visual and motor properties that were not specifically related to temporal change detection. This involved trials in which participants were tasked with detecting subtle visual oddballs in the image sequence. The control task started with the presentation of a green fixation cross (2-6 seconds, mean 4 seconds). The participants observed the production of sushi pieces similar to the main task and they were instructed to press the button using their right index finger as soon as they observed an oddball. The oddball, which was described to the participants as a sushi piece that was not fresh, was slightly different in contrast from other sushi pieces, and the identity of the oddball remained the same throughout the experiment for all the participants. We calibrated the contrast of the oddball image using the data from a pilot experiment such that the average difficulty of the main and the control task across participants was comparable (Figure S6). The timing of production of sushi pieces was determined by the same process as the main task with the parameter p fixed to 0.3 throughout the trials. This parameter value was chosen to match the mean production rate in the main task, which was the probability of sushi piece production in each bin. Oddball locations were determined using the same protocol that specified the change points in the main task. After oddball locations were specified, the sushi piece in the sequence that was closest in time to each of these locations was replaced by the oddball image. We explicitly instructed the participants that trials from the control task contained no switches between the chefs and there was no need to track the rate of production of sushi pieces.

After practice trials (3 trials each for the two difficulties of the main task and 2 trials of the control task in a randomized order within participants), the participants engaged in a total of four runs, each containing 18 trials of the main task (9 trials for each difficulty) and 6 trials of the control task randomly intermixed. We used two pseudo-random sets of sequences: 50% of the participants underwent one sequence set, and the other participants underwent the second sequence set. These

two sets of sequences were selected so that the distribution of the number of change points and the timing of change points in each set were close to its expected value (See Figure S3 for their properties). This procedure reduced the possibility that the behavioral and/or neural effects we observed were driven by peculiar properties existent in a set of sequences and thus lacked generality. The order of sequences used in the trials was randomized within participants.

The participants were incentivized to maximize the number of correct detections of change points while minimizing the number of incorrect detections as follows (See Figure S4 for the description used in the experiment): at the end of the experiment, we randomly selected 25% of the trials. For each of the selected trials, a correct detection and an incorrect detection were rewarded and punished by \$2, respectively. The first button press within 5 seconds after a change point and 1 second after an oddball presentation counted as a correct detection. If there were multiple button presses within 5-second and 1-second time window for the main and control tasks, respectively, the button presses other than the first one in time within the window were deemed incorrect. All the button presses outside these time windows were incorrect detections. Importantly, there was no feedback given to participants about their performance during the trials and thus no feedback-based learning could take place during the experiment.

Task performance

The average number of participants' button presses per trial was significantly greater than the actual number of changes per trial in both easy and hard conditions (Figure 7A; $t(20)=2.23$, $p=0.037$, and $t(20)=4.12$, $p=5.3\times 10^{-4}$ for the easy and hard conditions, respectively, one-sample t test). The participants showed a significant increase in the rate of button press (Figure 7A; $t(20)=4.57$, $p=1.9\times 10^{-4}$, paired-sample t test) and a significant decrease in the performance f measure (Figure 7B; $t(20)=4.95$, $p=7.7\times 10^{-5}$, paired-sample t test) for the hard condition compared to the easy

condition. The performance f measure is our metric of participants' performance based on the distance between task-determined actual change points and participants' button presses (See Statistical Analysis section in Materials and Methods). Although the human participants performed considerably short of the theoretical optimum (Figure S7 shows the comparison of performance between the participants and an optimal Bayesian agent), they performed significantly better than the random agent in both conditions (Figure 7B; $t(20)=11.6, p=2.7\times 10^{-10}$ and $t(20)=12.0, p=1.3\times 10^{-10}$ for easy and hard conditions, respectively, paired-sample t test), indicating that the participants were capable of detecting temporal change points considerably beyond that expected by chance. The participants were not increasing the number of button presses randomly in the hard condition because if they were random, the sensitivity (i.e., the proportion of actual change points correctly detected by the participants) would have increased. This was not the case, and in fact we observed a significant decrease in sensitivity for the hard compared to easy condition (Figure S6; $t(20)=3.19, p=0.0046$, paired-sample t test). In the control task, we successfully matched the average difficulty of the two tasks in terms of the performance f measure (Figure S6; non-significant as $t(20)=1.36, p=0.19$, paired-sample t test), although there was a significant increase and decrease in precision and sensitivity, respectively (Figure S6; $t(20)=5.11, p=5.3\times 10^{-5}$ and $t(20)=6.83, p=1.2\times 10^{-6}$, paired-sample t test), compared to the main task.

Model fitting to participants' behavior

To describe participants' behavior quantitatively, we calibrated three computational models of various complexity: two Bayesian models inspired by the Bayesian Online Change Point Detection model (BOCPD; Adams and MacKay, 2007) and the Dynamic Belief Model (DBM; Yu and Cohen, 2009), alongside a computationally frugal model which we named the Temporal Averaging Model (TAM). We fit each model to predict the timing of individual button presses within each trial (See Statistical Analysis section in Materials and Methods for the description of each model). For each

model, we compared the performance of sub-models (i.e., different set of parameters used within a class of model) by means of predictive f measure averaged across participants. Predictive f measures for each participant were obtained by calculating the f measure (i.e., the distance between participants' button presses and button presses prescribed by the computational model) on the even-numbered trials using the best fitting parameter values estimated on the odd-numbered trials (See Methods; Table S4 reports the detailed result of this procedure). All the models performed significantly better than the random model (Figure 7C; $t(20)=17.1$, $p=2.1\times 10^{-13}$ for BOCPD, $t(20)=17.5$, $p=1.3\times 10^{-13}$ for DBM, and $t(20)=17.6$, $p=1.2\times 10^{-13}$ for TAM, paired-sample t test). However, in a model comparison, when comparing the predictive f measures between each of the models we found no significant differences among the models in terms of their capacity to account for human performance (Figure 7C and Figure S8A; $t(20)=0.95$, $p=0.35$ between BOCPD and DBM, $t(20)=1.80$, $p=0.088$ between DBM and TAM, and $t(20)=0.75$, $p=0.46$ between TAM and BOCPD, paired-sample t test). Note that because we used different behavioral data (odd trials) to fit the parameters for each model than we used to validate each model (even trials), it is not necessary to adjust our model comparison process to take into account differences in model complexity.

After identifying the best sub-model for each model, we fit each model again using the data from all trials for each participant (Figure S8B; Figure S9 reports fitted parameter values). Although we did not differentiate model's button presses by the distance to the participant's button presses as long as they were in the 10-second window around a participant's button press, all the models fitted to replicate the timing of participants' button presses reasonably well (Figure 7D). The models also captured trial-by-trial variability in the number of button presses in participants' behavior. As the number of participants' button presses per trial increased, the average number of button presses prescribed by the model increased (Figure 7E).

Taken together our behavioral results demonstrate that all three candidate models do equally well in capturing participant's behavioral data. Even the very simple TAM performed as well as the most sophisticated Bayesian model in explaining participant's change detection behavior. Given this situation of equivocation on the model predictions against the behavioral data, in order to understand better which model provides a best account of the neural computations underlying change detection, we next turned to the fMRI data analysis in Chapter 4.

DISCUSSION

This study investigated the computational basis of temporal change point detection. We tested three different computational models of varying degree of sophistication against behavioral data to examine the capacity of each model in describing participants' behavior. All the models captured important aspects of participants' behavior, and in fact there were no significant differences among the models in their predictive performance using the validation set despite the difference in their complexities.

We could successfully modify and apply computational models of change detection aimed to solve non-temporal domain tasks to our novel temporal change detection task. This is because our task is conceptually similar to tasks from previous reports where participants inferred changes in non-temporal domain and utilized the inference to their decision making (Behrens et al., 2007; Nassar et al., 2010, 2012; Summerfield et al., 2011; Payzan-LeNestour et al., 2013; McGuire et al., 2014). For instance in one study, participants were tested in a predictive inference task where they predicted the spatial location of stimulus appearance on the following trial located on a 1-dimensional line. Here, the position of stimulus in each trial was drawn from a Gaussian distribution containing occasional jumps in its mean value and these jumps had to be inferred for a successful performance

in the task (McGuire et al., 2014). Our task can be interpreted in a similar manner such that the participants inferred jumps in the mean of Bernoulli probability that governed the production of sushi pieces by observing existence and non-existence of the production in each time bin of 0.25-second. Alternatively, our task can be regarded as a modified-version of the one-armed bandit task (Behrens et al., 2007). The participants observe only one bandit and reports jumps in reward probability instead of selecting one of two bandits, which is thought to be associated with higher reward probability.

Despite of these conceptual resemblances, there are two important aspects that make our task subjectively very distinct from previous ones in non-temporal domains. First, we explicitly asked the participants to detect and report change points rather than utilizing the knowledge of inferred change points to other decision making tasks (e.g., prediction regarding the stimulus in the following trial in the aforementioned predictive-inference task). Second, participants were required to update their beliefs about the statistics of the environment much more frequently than previous tasks required. In our task, the participants observed a sequence of existence or non-existence of stimulus updated every 0.25-second and the variables relevant to the computational models we tested were calculated at this time resolution. Because previous studies adopted a traditional trial-by-trial design, observing a stimulus and updating the belief about the task environment occurred at a much lower frequency of around every 10 seconds.

These differences could be helpful in understanding why simple heuristic model performed as equally well as complex Bayesian models in our task. The superiority of a Bayesian approach comes from encoding explicit probability distributions of the task environment and updating the belief in an optimal manner. However, Bayesian models added little benefit over the heuristic model in our task. Since we informed the participants explicitly that a temporal change point

involved a switch between binary states either from high to low or from low to high probability of sushi piece production in each bin, detecting a change point in our task necessitated only a less precise estimate of the Bernoulli probability compared to a case where states were continuous (e.g., location on a 1-dimensional line in a predictive-inference task). Also, due to the high frequency nature of the observations of existence and non-existence of sushi piece productions, a simple arithmetic mean of image frequency provided a good approximation to the Bernoulli probability of the generative process in a short period of time (e.g., 5 seconds).

This finding could be compatible with previous reports that favor approximate or non-fully Bayesian strategies when task demands are high (i.e., require storing large amount of information and/or complex computation of posterior distribution) and/or when the task environment is volatile (Summerfield et al., 2011; O'Reilly et al., 2012). In the present study, during each step of Bayesian updating, the BOCPD model required the updating of joint probability distributions of observed data and run length while the DBM only computed the posterior distribution of one parameter: the Bernoulli probability of the appearance of a sushi piece in each bin. Although one step of updating was less complex compared to models in other studies that updated joint distribution of three parameters (Behrens et al., 2007), the high frequency of updating which was a feature of our task added considerable computational demand. Thus in our task, implementing the Bayesian models was in fact computationally very taxing.

In conclusion, we investigated the computational account of temporal change detection in the human behavior in this chapter. A simple heuristic model relying on arithmetic mean and subtraction performed as equally well as Bayesian models in describing participants' behavior. Thus, the behavioral data alone could not differentiate computational accounts underlying temporal

change detection. To further elucidate this, we tested the computational models against the fMRI data in the next chapter.

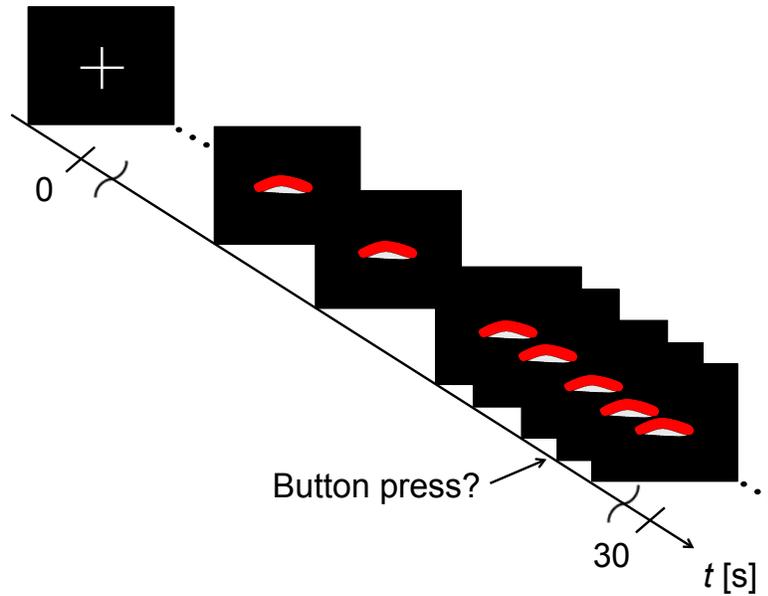
FIGURES AND TABLES

Figure 6. Illustration of task structure.

In the main task, participants observed sushi pieces produced by either one of the two chefs differing in their productivity in a trial of 30 seconds. The chef in production switched unexpectedly during a trial and participants were tasked to report changes between the chefs by the button press. In the illustrated example, participants observed a sudden increase in the rate of production. Some participants would press the button by reasoning that the perceived increase was due to the switch of chefs while others would not press the button by speculating that the observed increase was due to randomness and the chef in production did not change. In the control task, participants observed sushi pieces being produced as in the main task but the sequence contained sushi pieces that were not fresh represented by an image that was slightly different in contrast from other sushi pieces. Participants were tasked to detect these visual oddballs by pressing the button as soon as they appeared on the screen.

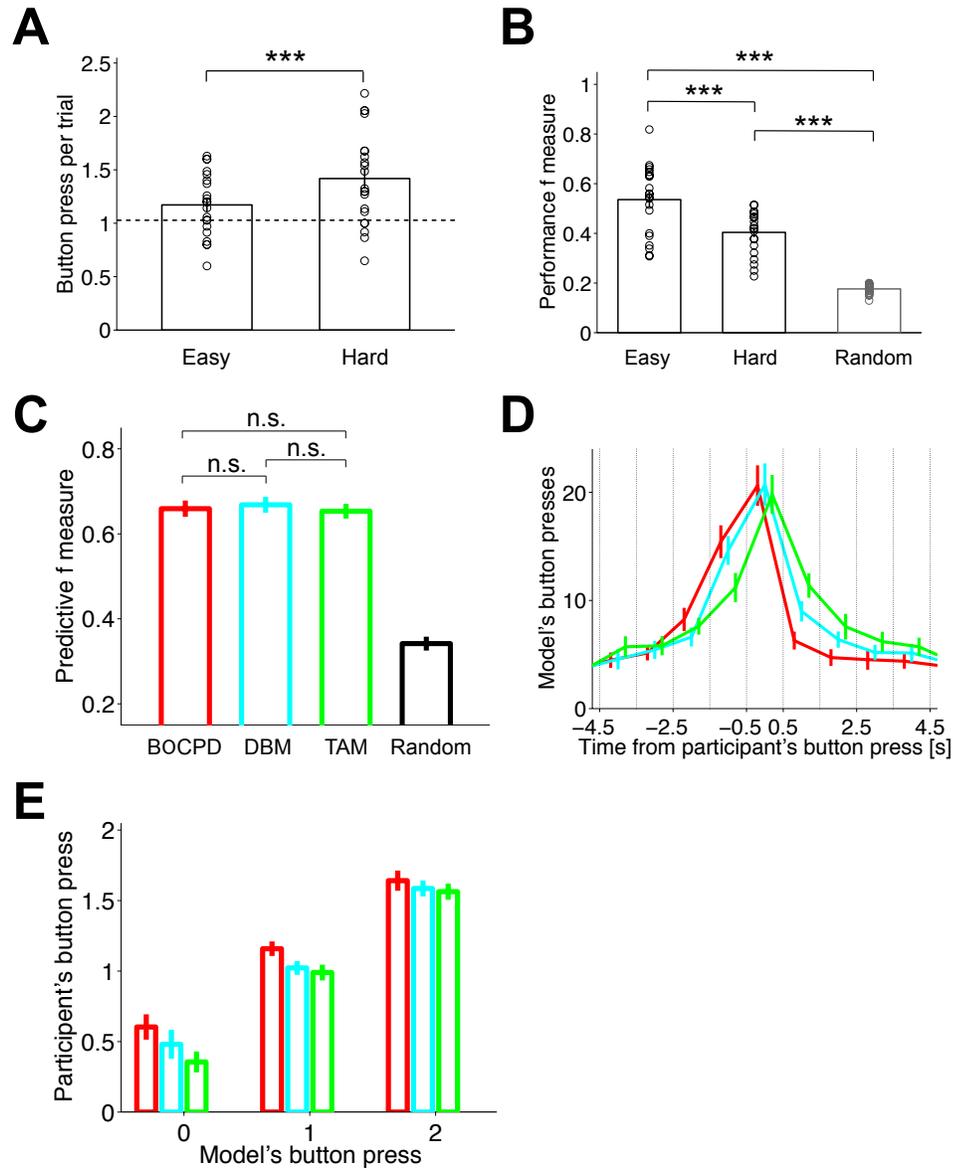


Figure 7. Behavioral results.

A. The average number of button presses across participants was significantly greater than the number of actual change points represented by a dashed line for both conditions ($t(20)=2.23$, $p=0.037$, and $t(20)=4.12$, $p=5.3 \times 10^{-4}$ for the easy and hard conditions, respectively, one-sample t test). The participants also showed a significant increase in the rate of button press in the hard condition compared to the easy condition (***: $t(20)=4.57$, $p=1.9 \times 10^{-4}$, paired-sample t test). Circles denote individual participant. Bars denote SEM across participants.

B. The participants showed a significant decrease in their performance f measure in the hard condition compared to the easy conditions (***: $t(20)=4.95$, $p=7.7\times 10^{-5}$, paired-sample t test). The performance f measures of the participants was significantly better than that of the random agent in both conditions (***: $t(20)=11.6$, $p=2.7\times 10^{-10}$ and $t(20)=12.0$, $p=1.3\times 10^{-10}$ for easy and hard conditions, respectively, paired-sample t test). Circles denote individual participant. Bars denote SEM across participants.

C. Predictive f measure from hold-out validation procedure of model fitting. For each model and participant, the model's predictive performance on the even-numbered trials was calculated using the best fitting parameters obtained from the calibration to the odd-numbered trials. There were no significant differences among the performance of BOCPD, DBM, and TAM (n.s., non-significant as $t(20)=0.95$, $p=0.35$ between BOCPD and DBM, $t(20)=1.80$, $p=0.088$ between DBM and TAM, $t(20)=0.75$, $p=0.46$ between TAM and BOCPD, paired-sample t test). All the three models performed significantly better than the random model ($t(20)=17.1$, $p=2.1\times 10^{-13}$ for BOCPD, $t(20)=17.5$, $p=1.3\times 10^{-13}$ for DBM, $t(20)=17.6$, $p=1.2\times 10^{-13}$ for TAM, paired-sample t test). Error bars represent SEM. Red: BOCPD, Cyan: DBM, Green: TAM, Black: random model.

D. Histogram of the average number of button presses predicted by each model compared to the timing of participant's button presses. We created 1-second bins from -4.5 to 4.5 seconds around participant's button presses and counted the number of model button presses that fell into each bin. Bars denote SEM across participants. Red: BOCPD, Cyan: DBM, Green: TAM.

E. Comparison of the trial-by-trial number of button presses of the actual participants vs. each model prediction. We categorized the trials by the number of button presses prescribed by the model and computed the average number of participant's button presses for trials in each category. Bars denote SEM across participants. Red: BOCPD, Cyan: DBM, Green: TAM.

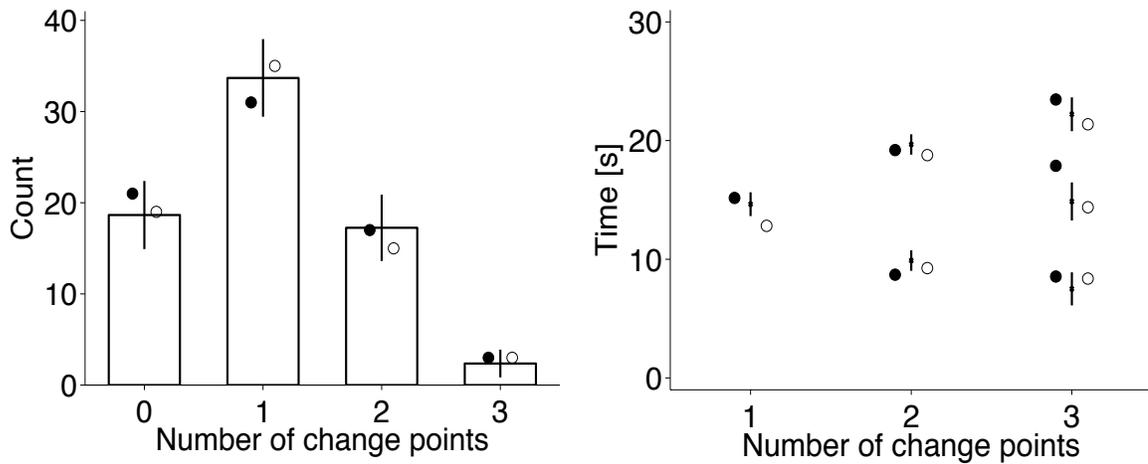


Figure S3. Related to Figure 6. Experimental design.

Characteristic of the two sets of sequences used in the main task. For each of the 72 trials in the main task, we generated temporal change points (See Results) and counted the number and recorded the timing of change points. We iterated this process for 1000 times and averaged the number of change points contained in each set (left). We also averaged the timing of change points for trials in which there were one, two or three change points, separately (right). Bars denote the standard deviation of the simulation result. Open and filled circles denote the two sets of sequences used in the main task of the experiment.

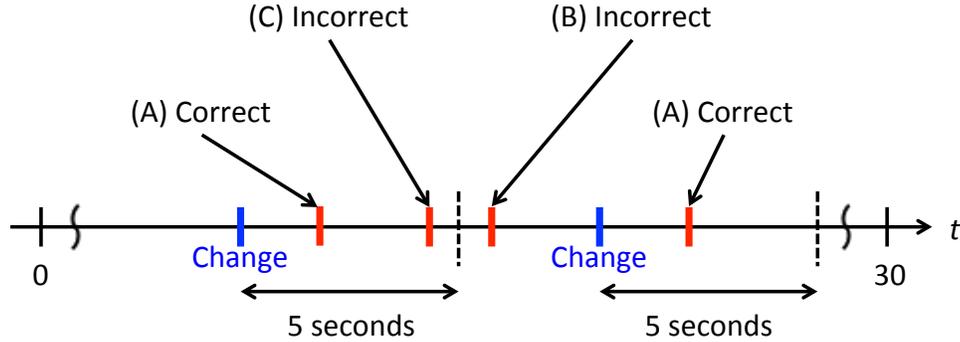


Figure S4. Related to Figure 7. Performance metric.

Illustration of the incentive scheme of the main task used in the instruction for the participants. A correct button press was when a change occurred and a button press was made within 5 seconds after the change point (A in Figure). Each correct button press was rewarded by \$2. There were two instances of an incorrect button press. One was that there was a button press but no change occurred within 5 seconds before the button press (B in Figure). The other was that there was a change point within 5 seconds before the button press, but after the change point and before the button press there was another button press (C in Figure). An incorrect button press was punished by \$2. For the control task, the same incentive scheme was used with 1-second time window.

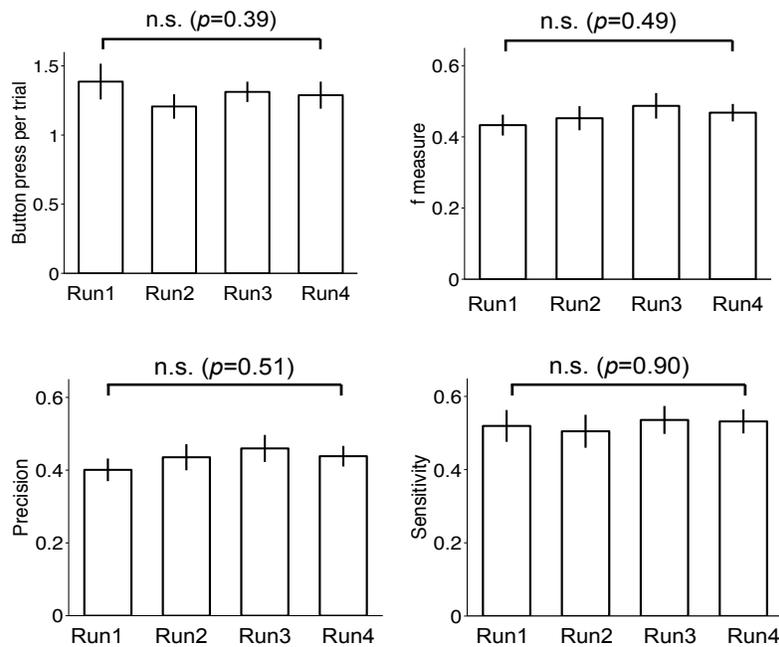


Figure S5. Related to Figure 7. Participants' behavior across runs.

Participants' behavior in the main task across runs. No significant differences were observed across runs in terms of the number of button presses per trial, performance f measure, precision, and sensitivity of participants' performance (n.s., non-significant with associated p value denoted in the parentheses, one-way repeated measure ANOVA). Bars represent SEM.

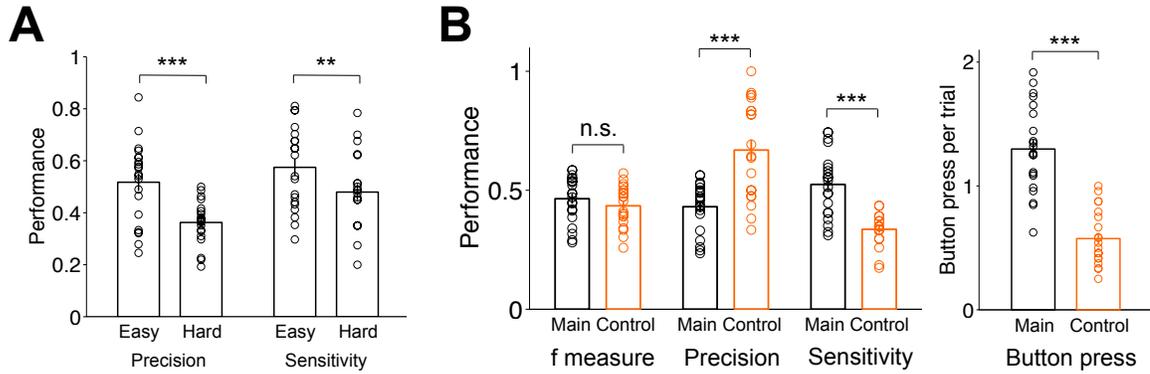


Figure S6. Related to Figure 7. Participants' behavior across conditions.

A. Precision and sensitivity of participants' behavior in easy and hard conditions. Participants showed a significant decrease in both precision and sensitivity in the hard condition compared to the easy condition (***: $t(20)=5.64$, $p=1.6 \times 10^{-5}$, **: $t(20)=3.19$, $p=0.0046$, paired-sample t test). Open circles denote each participant. Bars denote SEM.

B. Comparison of participants' behavior in the main and the control task. We matched the difficulty of the two tasks in term of performance f measure (n.s., non-significant as $t(20)=1.36$, $p=0.19$, paired-sample t test). There were significant differences between the participants' behavior in the two tasks with regard to precision (***: $t(20)=5.11$, $p=5.3 \times 10^{-5}$, paired-sample t test), sensitivity (***: $t(20)=6.83$, $p=1.2 \times 10^{-6}$, paired-sample t test), and the number of button presses per trial (***: $t(20)=8.33$, $p=6.2 \times 10^{-8}$, paired-sample t test). Black: participants' behavior in the main task, Orange: participants' behavior in the control task. Open circles denote each participant. Bars represent SEM.

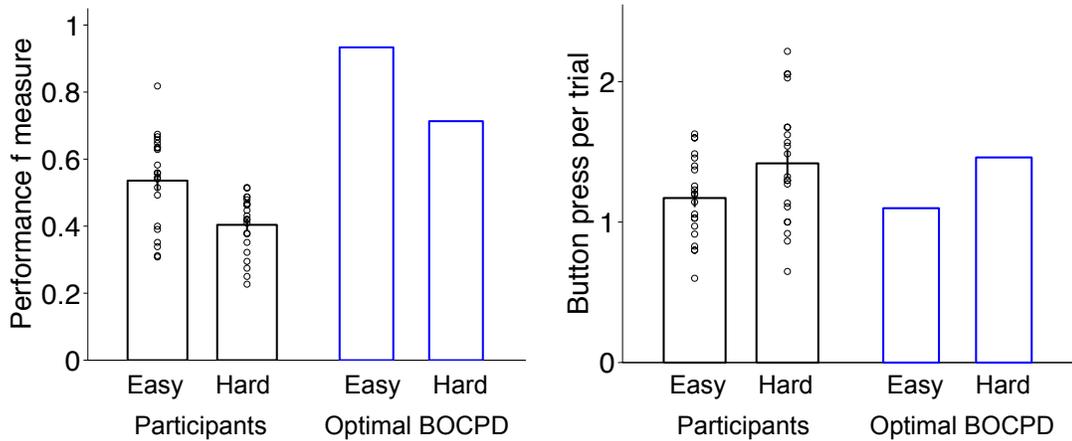


Figure S7. Related to Figure 7. Performance of the optimal Bayesian model.

Comparison of the performance between the participants and an optimal Bayesian agent. The optimal Bayesian agent employed the BOCPD with the optimal parameter values (i.e., $h=1/80$, $\theta=0.5$, $T=5[s]$). The task was difficult for the participants as seen from the discrepancy in the performance f measure (left). On aggregate, the number of button presses of the participants was comparable to that of the optimal Bayesian agent (right). Black: participants' behavior, Blue: Optimal Bayesian agent. Open circles denote each participant. Bars represent SEM.

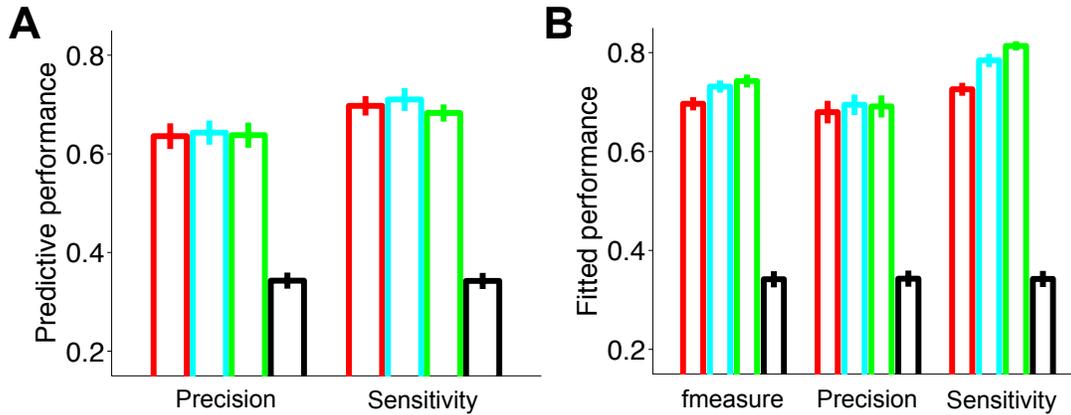


Figure S8. Related to Figure 7. Performance of computational models.

A. Predictive performance in terms of precision and sensitivity. Best parameter values calibrated on the odd-numbered trials were used to obtain predictive precision and sensitivity on the even-numbered trials. Bars denote SEM across participants. Red: BOCPD, Cyan: DBM, Green: TAM, Black: random model.

B. The performance of each model in terms of fitted f measure, precision, and sensitivity resulting from model fitting using data from all the trials. For each participant, we first identified the best sub-model for each model using hold-out validation, and then fitted the model using all data points. Bars denote SEM across participants. Red: BOCPD, Cyan: DBM, Green: TAM, Black: random model.

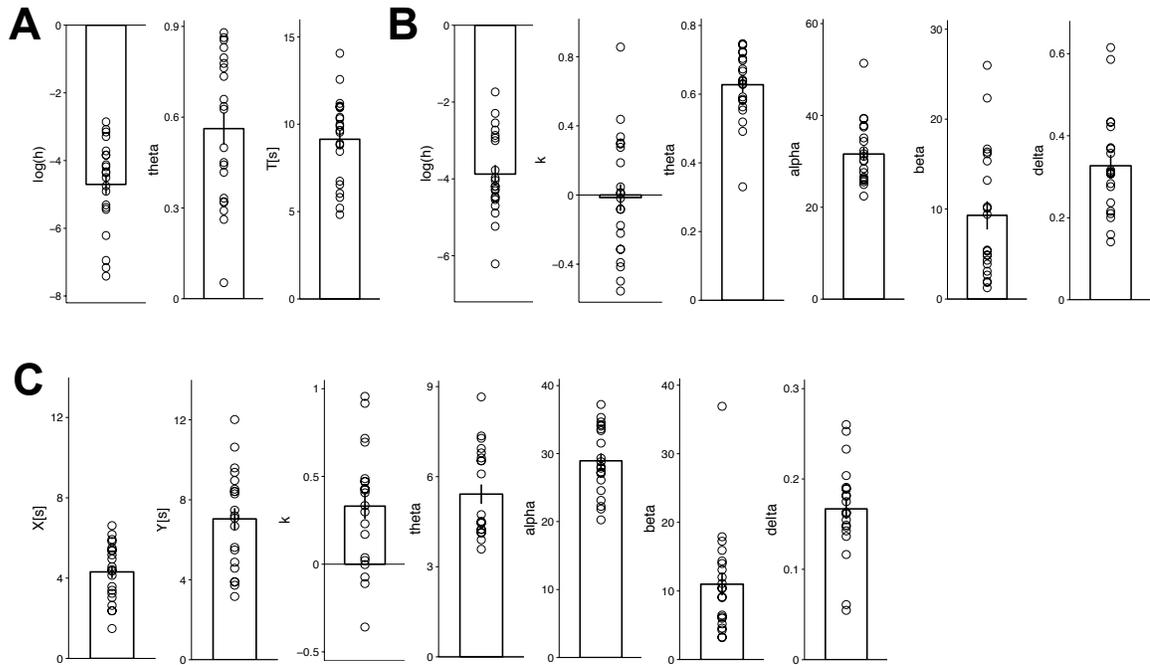


Figure S9. Related to Figure S8B and Figure S11. Best fitting parameter values for the best sub-model of each computational model of participants' behavior.

A. BOCPD. The estimated hazard rate was close to the true hazard rate of $\log h = -4.382$. The threshold value was close to its optimal value of 0.5 on average although there was a great heterogeneity across participants. Most participants employed greater than optimal size of time window ($T = 5[s]$) to compute change point probability. Open circles denote each participant. Bars represent SEM.

B. DBM. On average, the estimated hazard rate was greater than the true hazard rate. Although the average estimated parameter value associated with Weber's law was zero, there was a great variability across participants.

C. TAM. Participants were using larger number of data points before than after the change point as seen from $X < Y$ on average to determine the average rate of image presentations. This was sensible because the larger number of data points a participant used, the better estimate of

underlying rate of image presentations he/she could obtain under the stable period before the change point. Also, the duration of the time window X used to calculate the average rate of image presentations after perceived change points was close to 5 seconds. This was plausible given that the participants were informed that they would be rewarded for a correct button press within 5 seconds from a change point. Furthermore, the value of the parameter k was greater than zero on average. Therefore, the greater the current rate of image presentation was, the greater the threshold to detect a change point became. This effect conformed to Weber's law.

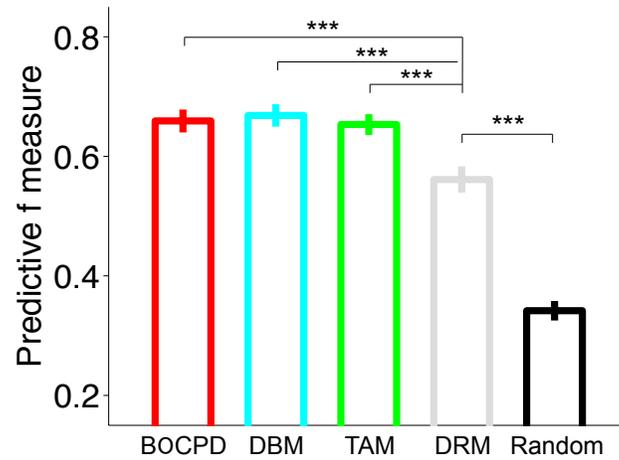


Figure S10. Related to Figure 7. Performance of the computational model based on delta-rule.

Comparison of the performance of the delta-rule model against the three models analyzed in the main text (See Figure 7C). The performance of DRM fell short of the other three computational models (***: $t(20)=5.18$, $p=4.6\times 10^{-5}$ between BOCPD and DRM, $t(20)=5.23$, $p=4.0\times 10^{-5}$ between DBM and DRM, $t(20)=4.69$, $p=1.4\times 10^{-4}$ between TAM and DRM, paired-sample t test). DRM performed significantly better than the random model (***: $t(20)=12.0$, $p=1.3\times 10^{-10}$, paired-sample t test). Error bars represent SEM. Red: BOCPD, Cyan: DBM, Green: TAM, Gray: DRM (Delta-rule model), Black: random model.

BOCPD (All sub-models included free parameters θ and h .)				
True prior				
Parameter	-	T	α, β, δ	T, α, β, δ
Reset to 0	0.636	0.6593	0.658	0.6591
Posterior mean	0.607	0.637	0.636	0.634
Posterior mode	0.593	0.609	0.631	0.630
Uniform prior				
Parameter	-	T	α, β, δ	T, α, β, δ
Reset to 0	0.605	0.634	0.625	0.634
Posterior mean	0.585	0.644	0.602	0.623
Posterior mode	0.568	0.592	0.611	0.626

DBM (All sub-models included free parameters θ and h .)								
True prior								
Parameter	-	T	k	T, k	α, β, δ	T, α, β, δ	k, α, β, δ	$T, k, \alpha, \beta, \delta$
Post. mean	0.590	0.620	0.597	0.617	0.661	0.667	0.669	0.656
Post. mode	0.454	0.477	0.442	0.463	0.648	0.652	0.641	0.654
Uniform prior								
Parameter	-	T	k	T, k	α, β, δ	T, α, β, δ	k, α, β, δ	$T, k, \alpha, \beta, \delta$
Post. mean	0.567	0.588	0.554	0.585	0.641	0.646	0.640	0.655
Post. mode	0.445	0.468	0.451	0.472	0.659	0.639	0.645	0.630

TAM (All sub-models included free parameters X, Y and θ .)				
Parameter	-	k	α, β, δ	k, α, β, δ
	0.621	0.626	0.626	0.653

Table S4. Related to Figure 7C. Details of model fitting and comparison.

Model comparison using hold-out validation for Bayesian Online Change Point Detection model (BOCPD), Dynamic Belief Model (DBM), and Temporal Averaging Model (TAM). The values reported are the predictive f measures obtained by simulating the model on even-numbered trials using the best fitting parameter values to odd-numbered trials. For each model, we tested sub-models that varied in the number of free parameters. The best sub-model within each class of models is written in bold and italic. See Materials and Methods for the description of each model and parameter.

NEURAL COMPUTATIONS MEDIATING TEMPORAL CHANGE DETECTION IN THE HUMAN BRAIN

Adachi R, Suzuki S, O’Doherty JP. Neural computations mediating temporal change detection in the human brain.

R.A. and J.P.O. designed the experiment. R.A. and S.S. collected the data. R.A. performed data analysis with feedback from S.S. and J.P.O. R.A. and J.P.O. wrote the manuscript. J.P.O. supervised the research.

INTRODUCTION

In spite of the importance of understanding how the brain solves change detection problems where inference about the change point has to be inferred as opposed to being explicitly observable, the literature on this topic is surprisingly sparse to date, and has focused on only a small subset of domains. In value-based learning, some studies have identified neural systems involved when inferences about changes in reward contingencies in the absence of explicit signaling of change point result in changes in behavior (Cools et al., 2002; O’Doherty et al., 2003; Gläscher et al., 2009a). Computational fMRI studies have characterized the neural computations underlying detection of changes in reward contingencies (Hampton et al., 2006; Behrens et al., 2007; Payzan-LeNestour et al., 2013). Related work has also been conducted in the domain of perceptual category learning (Summerfield et al., 2011), while others have characterized the mechanisms underlying change detection in the occurrence of stimuli at different spatial locations (McGuire et al., 2014). In spite of existing work on change detection in perceptual, spatial, and value-domains, the neural mechanisms underlying change detection in the temporal domain (beyond the trivial case of oddball

paradigms in which the change is directly signaled and need not be inferred) is largely unexplored.

The goal of the present study was two-fold: first of all, we aimed to describe the neural mechanisms of change detection in the temporal domain. Secondly, we aimed to conduct a comprehensive assessment of which computational mechanism could best describe the neuronal processes involved in this form of change detection, by formally comparing optimal Bayesian models of change detection to simpler yet computationally more tractable approximate strategies.

At the neuroanatomical level, we hypothesized that regions of the frontoparietal network would be involved in temporal change detection. This is because this network has been implicated not only in oddball detection (Kim, 2014), but also in change detection in non-temporal domains (McGuire et al., 2014), as well as in perceptual decision-making more generally (Heekeren et al., 2004, 2006; Gold and Shadlen, 2007).

MATERIALS AND METHODS

FMRI DATA PREPROCESSING

fMRI data acquisition

The task was conducted in a Siemens Trio 3T scanner using a 32-channel radio frequency head coil. The imaging data was acquired at a 20-degree angle from the anterior commissure-posterior commissure line in order to maximize orbital sensitivity (Deichmann et al., 2003), using a multi-band echo planar imaging sequence with the acceleration factor of 4. 56 axial slices were acquired with no slice gap with the following scan parameters: echo time 30 ms, repetition time 1000 ms, field of view 200×200 mm, and voxel size 2.5 mm isotropic. High-resolution T1 images were collected at the beginning of each participant's session.

fMRI data preprocessing

SPM12 (Wellcome Department of Imaging Neuroscience, London, UK; www.fil.ion.ucl.ac.uk/spm) was used to preprocess and analyze the data. The functional images for each participant were slice time corrected to the beginning of each TR (0 ms) and then spatially realigned to the first volume using a 6-parameter rigid body spatial transformation. The high-resolution structural image was then co-registered to the mean functional image generated from the realignment procedure. The realigned functional images and co-registered structural image were spatially normalized to the MNI template by using warping parameters obtained through segmentation process. Finally, the normalized functional images were spatially smoothed using a Gaussian kernel of 8 mm isotropic.

STATISTICAL ANALYSIS

Creating the decision variable regressor

To examine the brain regions implementing the decision variable to detect temporal change points, we estimated the time course of the decision variable derived from each model during each trial at 250 ms time resolution. For each participant and model, we simulated the model using the best fitting parameters to all the trials and timing of the image presentations (production of sushi pieces) actually experienced by the participant in each trial (Figure S11, Table S5). This time course was used in the first-level participant-specific general linear models (GLMs) after being convolved with a canonical hemodynamic response function (HRF) provided by SPM12.

fMRI statistical analysis

We estimated several participant-specific GLMs to test our hypotheses. All the regressors of interest in the GLM were concatenated across runs and convolved with a canonical hemodynamic response function provided by SPM12. The contrasts obtained from the estimated first-level GLM of each participant were used for second-level random effect analyses. For the whole-brain analyses, we

applied correction for multiple comparisons at the cluster level $p < 0.05$ with the underlying voxel-wise threshold of $p < 0.001$ and a minimum spatial extent of 97 voxels, calculated by using a version of 3dClustSim in AFNI that does not assume Gaussian spatial correlations, and which is therefore not susceptible to inflated false positives (Cox et al., 2017). For display purposes, all the statistical maps presented are thresholded at $p < 0.001$ uncorrected. In the tables, the coordinates reported are in MNI coordinates of the peak voxel in each of the activation clusters. All the GLMs included the following regressors of no interest: 6 movement regressors for each of the 4 runs generated from the realignment procedure in the image preprocessing accounting for head motion, a constant term throughout the runs, and a constant term for each run except the last one. SPM's default serial orthogonalization was turned off and each regressor in the GLMs captured the unique variance in the acquired signal.

GLM1

This GLM was estimated to test for regions showing greater activation with the button press in the main task than in the control task. It included (1) image presentations, (2) button presses in the main task, (3) button presses in the control task. The contrast of interest was (2)-(3) (Figure 8A), and (3)-(2). See Table S7 for all the activations.

GLM2

Neural correlates of the variables relevant for the BOCPD were tested. The model included (1) image presentations, (2) button presses, (3) the decision period in the control task, (4) a stick function set to occur every 250 ms from the start to the end of a trial, parametrically modulated by (5) change point probability (i.e., the DV in Equation (4)), and (6) the variance of the posterior run length distribution. The contrast of interest was (5) (Figure 9A). See Table S8 for all the activations and for other contrasts.

GLM3

This GLM was estimated to investigate the regions representing the variables derived from the DBM. The model included (1) image presentations, (2) button presses, (3) the decision period in the control task, (4) a stick function set to occur every 250 ms from the start to the end of a trial, parametrically modulated by (5) posterior mean of Bernoulli probability in the current bin (i.e., $post_n$ in Equation (6)), (6) the posterior mean of the Bernoulli probability within the last T seconds that showed maximum absolute difference from the value in (5) (i.e., $post_{n-i}$ in Equation (6)), (7) the absolute difference between (5) and (6) (i.e., DV in Equation (6)), and (8) the variance of posterior Bernoulli probability distribution. The contrast of interest was (7) (Figure 9A). See Table S8 for all the activations and for other contrasts.

GLM4

Areas correlating with the variables in the TAM were tested here. The model included (1) image presentations, (2) button presses, (3) the decision period in the control task, (4) a stick function set to occur every 250 ms from the start to the end of a trial, parametrically modulated by (5) the mean image presentation rate per second at $-X-Y \sim -X$ second (i.e., \bar{Y}_n in Equation (7)), (6) the mean image presentation rate per second at $-X \sim 0$ second (i.e., \bar{X}_n in Equation (7)), and (7) the absolute difference between (5) and (6) (i.e., DV in Equation (7)). The contrast of interest was (7) (Figure 9A). See Table S8 for all the activations.

GLM5

After identifying areas significantly correlated with the DV from the TAM in GLM4, in order to determine where in the brain the TAM model uniquely explained variance over and above the predictions of the other models, we ran a further GLM in which we included in the same design

matrix all three model time series (with orthogonalization disabled). This meant that all three model variables could compete for variance, and according to the sum of squares principle we could therefore report regions responding to one of the models over and above the variance explained by any of the other models. The model included (1) image presentations, (2) button presses, (3) the decision period in the control task, (4) a stick function set to occur every 250 ms from the start to the end of a trial, parametrically modulated by (5) DV of the BOCPD ((5) in GLM2), (6) DV of the DBM ((7) in GLM3), and (7) DV of the TAM ((7) in GLM4). The decision variables from the three computational models were z-scored before the convolution with HRF to make the regression coefficients comparable across the models. The contrast of interest was (7) (Figure 10; see Table S6 for the correlation between regressors; see Table S9 for all the activations).

ROI analysis

To illustrate the relationship between the task-related variable and the BOLD signal, we conducted two ROI analyses. First, we tested how much the activity in the right dlPFC changed in response to button presses in the main and the control task using GLM1 (Figure 8B). Second, we explored using GLM4 if the activity in right dlPFC and right IPS scaled with the value of decision variable derived from the TAM. We categorized the value of the decision variable at each bin into low, medium, and high using tertiles and then fit this variable to the BOLD signal to plot the average effect size for each category (Figure 9B). We specified ROIs using a leave-one-subject-out cross-validation procedure to ensure statistical independence. For each ROI analysis, we ran a second-level analysis of the associated GLM excluding one participant and identified peak voxel of activation nearest to the peak of activation from second-level analysis using all the participants. For the participant excluded from this procedure, we computed the effect size of the variable of interest in the identified peak voxel. We repeated this analysis for the number of participants excluding one different participant for each analysis. When we selected the peak activation coordinates in right

dIPFC and right IPS from the second-level analysis of GLM4 using all the participants, we restricted our search to the voxels within independent ROIs in these areas because the activation clusters extended to areas outside right dIPFC and right IPS. We defined our independent ROI as a sphere of 10 mm radius around average peak coordinates in right dIPFC and right IPS found in relevant previous studies to be involved in perceptual and value-based decision making: $x=43$, $y=18$, $z=38$ for right dIPFC (Gluth et al., 2012; Suzuki et al., 2012, 2016; McGuire et al., 2014) and $x=37$, $y=-41$, $z=42$ for right IPS (Daw et al., 2006; Hare et al., 2011; Hunt et al., 2014; McGuire et al., 2014). We used rfxplot toolbox for this analysis (Gläscher, 2009b).

CODE AVAILABILITY

Key analysis codes written in MATLAB, and fMRI analysis results have been deposited in GitHub: <https://github.com/RyoAdachi/tcd>.

RESULTS

Brain regions showing different magnitude of activity associated with button presses in the main and control task

We first investigated which areas in the brain showed differential activations associated with the button presses in the main compared to the control task (See GLM1 in Statistical analysis in Materials and Methods). This analysis assessed the regions exhibiting an involvement in temporal change detection relative to a visual oddball detection task, with very similar visuomotor properties and matched difficulty levels. We found that activity in the regions of the frontoparietal network including right dorsolateral prefrontal cortex (dIPFC) and right intraparietal sulcus (IPS) showed significantly greater correlation with the button press in the main task than in the control task (Figure 8, $p<0.05$ whole brain corrected at cluster level; see Table S7 for other activated areas). We

also found that activity in various areas including ventromedial prefrontal cortex (vmPFC) and cingulate cortex exhibited significantly greater correlations with button presses in the control task than in the main task ($p < 0.05$ whole brain corrected at cluster level; see Table S7 for other activated areas).

Neural correlates of decision variables derived from three computational models of temporal change detection

We tested for brain regions representing the time course of the decision variables derived from each of the BOCPD, DBM, and TAM models. Using the best fitting parameter values for each model, we generated an estimated time course for the decision variable at the resolution of 250 ms for each of the three computational models. We directly correlated each model-predicted time series against the BOLD signal after convolving each time series with a canonical HRF in separate GLMs (See GLM2-4 in Statistical analysis in Materials and Methods). We found that the time course of the decision variable of the TAM showed a significantly positive correlation with activity in the frontoparietal network of right dlPFC and bilateral IPS (Figure 9, $p < 0.05$ whole brain corrected at cluster level; see Table S8 for other activated areas). We did not find areas showing significant positive correlations with the time course of the decision variables of either the BOCPD or DBM (Figure 9, $p < 0.05$ whole brain corrected at cluster level; see Table S8 for activated areas in other contrasts of BOCPD and DBM).

We further tested whether the correlation with the decision variable of the TAM in areas of the frontoparietal network survived when we included the decision variable from the other two computational models simultaneously so that each model time series could compete for variance (See GLM5 in Statistical analysis in Materials and Methods). We found that right dlPFC and right intraparietal lobule (IPL) encompassing IPS remained significantly correlated with the variable even

after the inclusion of the other model time series (Figure 10, $p < 0.05$ whole brain corrected at cluster level; see Table S9 for other activated areas). This result therefore indicates frontoparietal voxels in which activity can be explained as being consistent with the TAM model over and above any variance accounted for by the other two models. In other words, the findings show that activity in both dorsolateral prefrontal cortex and IPL incorporating IPS are significantly better explained by the TAM compared to the other two models.

DISCUSSION

Here we investigated the computational and neural basis of temporal change detection. We tested three different computational models of varying degrees of sophistication against behavioral and fMRI data to examine the neural mechanisms behind each model. All of these models captured important aspects of participants' behavior, and in fact there were no significant differences among the models in their predictive performance about behavior using the validation set despite the difference in model complexities. Yet, participants' behavior did not enable us to determine which model can best account for temporal change detection because each of the models made equally good predictions about behavioral performance on the task. Thus, we could not determine which model best accounts for human performance, based purely on behavioral analyses alone. However, when we examined correlates of the model variables in the brain using fMRI, there was a clear difference in the degree to which the different models could adequately explain variance in the BOLD signal. The simple heuristic model robustly captured brain activity in a frontoparietal network, whereas we found very little evidence throughout the brain for regions positively correlating with the output of the other models. Thus, when taking the fMRI data together with the behavioral data, our findings support the superiority of the simple heuristic model as a means of

describing the computations being implemented in the human brain during temporal change detection.

This finding could be compatible with previous reports that favor approximate or non-fully Bayesian strategies when task demands are high (i.e., require storing large amount of information and/or complex computation of posterior distribution) and/or when the task environment is volatile (Summerfield et al., 2011; O'Reilly et al., 2012). In the present study, during each step of Bayesian updating, the BOCPD model required the updating of joint probability distributions of observed data and run length while the DBM only computed the posterior distribution of one parameter: the Bernoulli probability of the appearance of a sushi piece in each bin. Although one step of updating was less complex compared to models in other studies that updated joint distribution of three parameters (Behrens et al., 2007), the high frequency of updating which was a feature of our task added considerable computational demand. Thus in our task, implementing the Bayesian models were in fact computationally very taxing.

The decision variable from the TAM was represented in dlPFC and IPS. In the TAM, the mean image presentation rates in two time windows (current and past time windows) are compared to detect change points. As the number of samples from the new regime increased in the current window following a change point, the mean image rate in the current window gradually increased or decreased. Since the samples in the past time windows were from the old regime and the mean image rate within the past time window was stable, it resulted in a gradual increase in the decision variable with random fluctuations due to the stochasticity of the data generating process. Thus, the time series of the decision variable of TAM resembled an evidence accumulation process toward a threshold, even though TAM was not in itself an evidence accumulation model per se such as the drift diffusion model (Ratcliff, 1978; Krajbich et al., 2010; Hare et al., 2011). As a consequence, the

present findings are broadly compatible with previous reports that dlPFC and IPS are involved in an evidence accumulation process to reach a decision in perceptual and value-based decision making tasks (Heekeren et al., 2004, 2006; Hare et al., 2011).

An important avenue to pursue in future research will be to elucidate the more general functional contributions of the dlPFC and IPS to change detection beyond the specific characteristics of the present task and temporal domain. Although the results of our simple subtraction analysis suggest that these regions are more involved in temporal change detection than in detecting explicitly observable changes in stimulus properties as in an oddball task, a deeper question concerns whether these regions are recruited in detecting hidden and non-directly observable changes in task relevant variables across other domains beyond the temporal sequences measured here. Furthermore, even though the present results pertain specifically to temporal change detection, it is feasible that the computational strategy for change detection embodied by the temporal averaging model could also be utilized for other forms of change detection beyond the temporal domain. However, work will be needed to test for this possibility.

The present study provides an example of how fMRI data can be used as an important complement to behavioral measurements in order to elucidate the computational mechanisms underlying human cognition. In the present case, behavioral data alone was insufficient to enable us to determine which out of several candidate models could best account for how temporal change detection is implemented. It was only by examining the degree to which activity in the brain could be accounted for by internal decision variables produced by each model that we could discriminate clearly between the models, ultimately favoring our simple heuristic model over the more normative Bayesian approaches.

One important limitation of the use of brain activity to arbitrate between models is that it is possible that the relationship between the numerical time series produced by the algorithm that describes the internal decision variable and the neural implementation of the relevant decision variable might not be as straightforward as is assumed. Future work will be needed to bridge the gap between the algorithmic level descriptions of the models being used here, and biologically plausible implementation schemes. It is possible that once armed with more detailed neurobiologically plausible implementation schemes, the predictions about the BOLD responses as described by each model might turn out to be very different. This notwithstanding, the finding of the superiority of the heuristic model in capturing brain activity does seem reasonable from the point of view of the principle of model parsimony in which the simplest model would tend to be favored, all else being equal. Furthermore, the heuristic model is also the more plausible one given likely constraints in cognitive capacity. It also important to note that although in our hands with the present task, the behavioral data alone could not discriminate between the models, it is also possible that novel behavioral tasks could be designed that would be more sensitive to differences in the model predictions, so as to adequately discriminate the model predictions solely based on behavioral data.

In conclusion, our study provides the first behavioral and neural investigation of temporal change detection, which is an important capacity necessary to facilitate adaptive behavior. A frontoparietal network including dlPFC and IPS was found to manifest activity consistent with the decision variable produced by a simple heuristic model, suggesting that the human brain relies on a computationally frugal strategy to detect changes in the environmental statistics of temporal sequences.

FIGURES AND TABLES

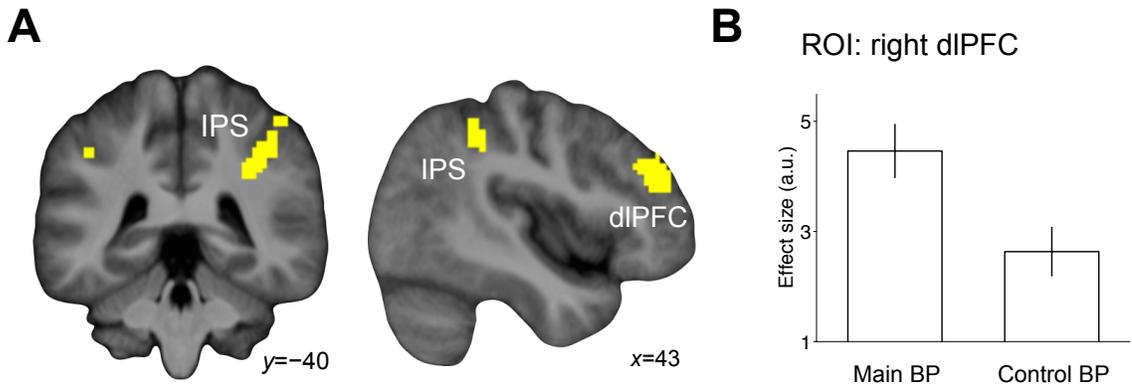


Figure 8. Brain areas showing greater activity associated with button presses in the main task than in the control task.

A. Activity in right dIPFC and right IPS showed significantly greater correlation with the button press in the main task than in the control task ($p < 0.05$ whole brain corrected at cluster level; see Table S7 for other activated areas). Activation map is thresholded at $p < 0.001$ for display purposes.

B. Plot of the button press effects in the main task and the control task in right dIPFC ROI identified by leave-one-subject-out cross-validation procedure. Error bars represent SEM. BP: button press, a.u.: arbitrary unit.

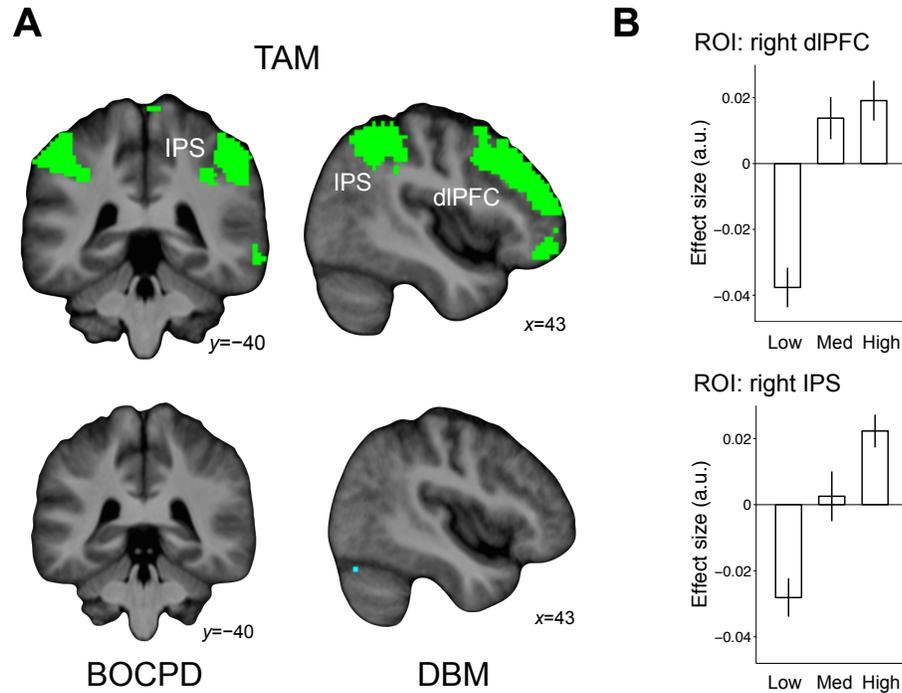


Figure 9. Neural correlates of the time course of the decision variables derived from each of the three computational models of temporal change detection.

A. Activity in right dIPFC and bilateral IPS showed a significant correlation with the model-predicted time course of the decision variable from the TAM ($p < 0.05$ whole brain corrected at cluster level; see Table S8 for other activated areas). The activation map is shown thresholded at $p < 0.001$. The model-predicted time course of the decision variable from the BOCPD or the DBM showed no significant positive correlations with any of the brain regions (See Table S8 for activated areas in other model-related contrasts).

B. Effect size of decision variable of TAM in right dIPFC and right IPS ROI identified by leave-one-subject-out cross-validation procedure. To obtain this plot, we categorized the decision variable estimated by using the fitted model at every 250 ms bin into low, medium, and high value for each participant and computed the average neural effect size in each category. Error bars represent SEM. a.u.: arbitrary unit.

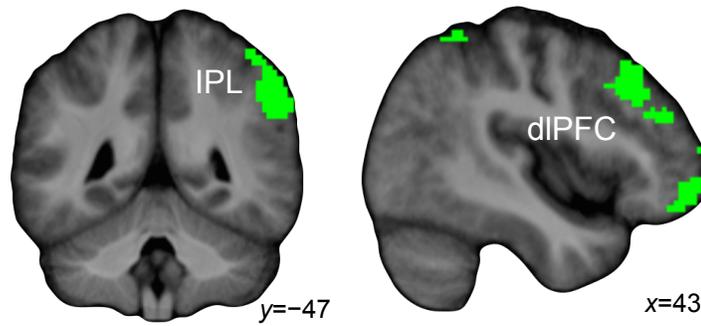


Figure 10. Brain regions that uniquely correlated with the time course of the decision variables derived from the TAM.

Activity in right dlPFC and right IPL encompassing IPS showed a significant correlation with the model-predicted time course of the decision variable from the TAM when including the decision variable from the other two models as confounds ($p < 0.05$ whole brain corrected at cluster level; see Table S9 for other activated areas). The activation map is shown thresholded at $p < 0.001$.

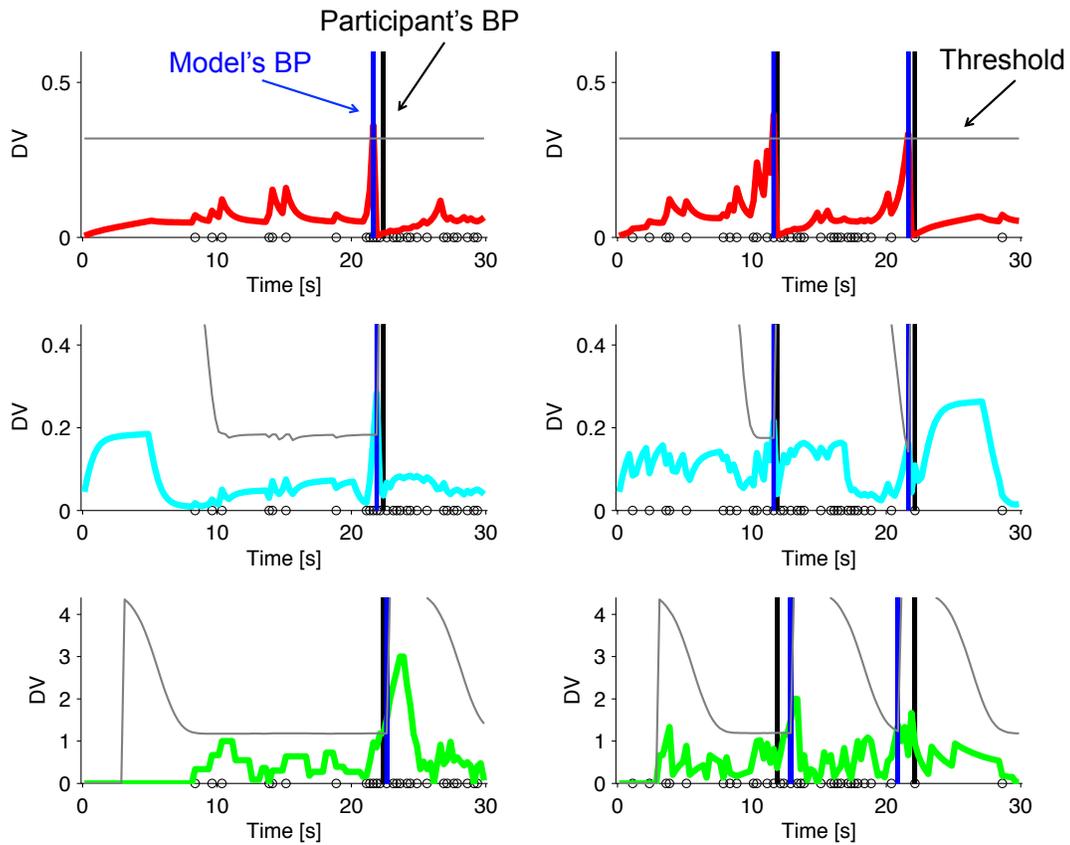


Figure S11. Related to Figure S9, Figure 9, and Figure 10. Illustration of the decision variable regressor.

We show the time course of the decision variable derived from each model for two example trials in one participant. We simulated the model using the set of best fitting parameters to all the trials across runs and the timing of image presentations experienced by the participant. Each row and column corresponds to each model and trial, respectively. Red: BOCPD, Cyan: DBM, Green: TAM. Thick colored lines denote the time course of decision variable. Thin black lines represent threshold value. Blue and black vertical lines are the timing of model's and participant's button presses, respectively. Open black circles on the x -axis are the timing of image presentations.

	Min	25th	Median	75th	Max
BOCPD	0.005	0.049	0.058	0.075	0.586
DBM	7.84×10^{-4}	0.032	0.083	0.165	0.367
TAM	0.00	0.133	0.364	0.714	4.00

Table S5. Related to Figure S11. Properties of the decision variables.

We report the minimum, 25th percentile, median, 75th percentile, and maximum of the decision variable of each of the three computational models.

	BOCPD vs DBM	DBM vs TAM	TAM vs BOCPD
DV time course	0.172	0.127	0.091
Convolved DV time course	0.195	0.186	0.231

Table S6. Related to Figure 10 and Figure S11. Average correlation between the time courses of the decision variables.

We computed the correlation between the time courses of the decision variable for each pair of computational models for each participant separately and calculated the average of the absolute value of the correlation across participants. DV: decision variable.

Contrast	Region	Hemi	Peak MNI coordinates			<i>t</i> score	Cluster size (voxels)
			<i>x</i>	<i>y</i>	<i>z</i>		
BP(main) - BP(control)	Superior frontal gyrus	R	23	8	55	7.33	334
	Precuneus	R	16	-67	60	6.72	220
	Middle frontal gyrus	R	43	40	22	5.18	102
	Intraparietal sulcus	R	33	-40	32	4.94	102
BP(control) - BP(main)	Medial frontal gyrus	R	8	53	-2	8.11	5943
	Cingulate gyrus	L	-4	-37	40	7.98	
	Cingulate gyrus	L	-2	6	28	7.69	
	Hippocampus	R	30	-22	-12	6.60	571
	Angular gyrus	L	-42	-70	40	5.58	312
	Postcentral gyrus	R	53	-14	52	5.20	212
	Fusiform gyrus	R	36	-44	-22	4.93	116

Table S7. Related to Figure 8A. MNI coordinates of peak activation voxels from the contrasts in GLM1.

Reported activation clusters survived $p < 0.05$ whole-brain correction at cluster level with the height threshold of $p < 0.001$ and the minimum spatial extent of 97 voxels. BP: button press.

Contrast	Region	Hemi	Peak MNI coordinates			<i>t</i> score	Cluster size (voxels)
BOCPD							
DV	-	-	-	-	-	-	-
Variance of	Lingual gyrus	L	-14	-47	-5	9.33	15393
	Superior temporal gyrus	R	60	-42	10	9.31	
	Superior temporal gyrus	R	48	-50	8	8.80	
	Postcentral gyrus	L	-30	-32	50	5.92	235
DBM							
DV	-	-	-	-	-	-	-
Variance of posterior distribution on Bernoulli probability	Insula	R	30	18	0	6.70	213
	Cingulate gyrus	R	8	20	42	6.20	359
	Occipital gyrus	R	26	-97	5	6.04	159
	Precentral gyrus	R	43	8	28	5.77	470
	Insula	L	-27	26	-2	5.66	102
	Occipital gyrus	L	-20	-97	5	5.36	196
TAM							
DV	Precuneus (incl. bilateral IPS)	R	10	-64	65	8.46	3025
	Precentral gyrus (incl. dlPFC)	R	20	16	60	6.84	2321
	Cerebellum	L	-10	-77	-30	5.69	137
	Middle temporal gyrus	R	60	-50	-12	5.63	163
	Middle frontal gyrus	L	-24	10	58	5.28	134
	Inferior frontal gyrus	L	-44	46	-18	4.97	102

Table S8. Related to Figure 9A. MNI coordinates of peak activation voxels from the contrasts in GLM2-GLM4.

Reported activation clusters survived $p < 0.05$ whole-brain correction at cluster level with the height threshold of $p < 0.001$ and the minimum spatial extent of 97 voxels. DV: decision variable.

Contrast	Region	Hemi	Peak MNI coordinates			<i>t</i> score	Cluster size (voxels)
TAM-DV	Precentral gyrus (incl. dlPFC)	R	18	18	65	6.48	1040
	Inferior parietal lobule	R	63	-52	18	4.34	172

Table S9. Related to Figure 10. MNI coordinates of peak activation voxels from the contrasts in GLM5.

Reported activation clusters survived $p < 0.05$ whole-brain correction at cluster level with the height threshold of $p < 0.001$ and the minimum spatial extent of 97 voxels. DV: decision variable.

BIBLIOGRAPHY

- Adams RP, MacKay DJ (2007) Bayesian online changepoint detection. Cambridge, UK: University of Cambridge Technical Report.
- Armel KC, Beaumel A, Rangel A (2008) Biasing simple choices by manipulating relative visual attention. *Judgment and Decision Making* 3:396-403.
- Basten U, Biele G, Heekeren HR, Fiebach CJ (2010) How the brain integrates costs and benefits during decision making. *Proc Natl Acad Sci U S A* 107:21767-21772.
- Becker GM, Degroot MH, Marschak J (1964) Measuring utility by a single-response sequential method. *Behav Sci* 9:226-232.
- Behrens TE, Woolrich MW, Walton ME, Rushworth MF (2007) Learning the value of information in an uncertain world. *Nat Neurosci* 10:1214-1221.
- Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD (2006) The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol Rev* 113:700-765.
- Bogacz R (2007) Optimal decision-making theories: linking neurobiology with behaviour. *Trends Cogn Sci* 11:118-125.
- Boorman ED, Behrens TE, Woolrich MW, Rushworth MF (2009) How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron* 62:733-743.
- Brett M, Anton JL, Valabregue R, Poline JB (2002) Region of interest analysis using an SPM toolbox. In: Eighth International Conference on Functional Mapping of the Human Brain, Sendai, Japan, June 2-6.
- Chib V, Rangel A, Shimojo S, O'Doherty JP (2009) Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J Neurosci* 29:12315-12320.
- Clithero JA, Rangel A (2013) Informatic parcellation of the network involved in the computation of subjective value. *Social Cogn Affect Neurosci* 9:1289-1302.
- Cools R, Clark L, Owen AM, Robbins TW (2002) Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *J Neurosci* 22:4563-4567.
- Cox RW, Chen G, Glen DR, Reynolds RC, Taylor PA (2017) FMRI clustering in AFNI: false positive rates redux. *Brain connect* 7:152-171.

- Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* 8:1704-1711.
- Daw ND, O'Doherty JP, Dayan P, Dolan RJ, Seymour B (2006) Cortical substrates for exploratory decisions in humans. *Nature* 441:876-879.
- Deichmann R, Gottfried J, Hutton C, Turner R (2003) Optimized EPI for fMRI studies of the orbitofrontal cortex. *Neuroimage* 19:430-441.
- De Martino B, Fleming SM, Garrett N, Dolan RJ (2013a) Confidence in value-based choice. *Nat Neurosci* 16:105-110.
- De Martino B, O'Doherty JP, Ray D, Bossaerts P, Camerer C (2013b) In the mind of the market: Theory of mind biases value computation during financial bubbles. *Neuron* 79:1222-1231.
- Dunne S, O'Doherty JP (2013) Insights from the application of computational neuroimaging to social neuroscience. *Curr Opin Neurobiol* 23:387-392.
- Forstmann BU, Ratcliff R, Wagenmakers EJ (2016) Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annu Rev Psychol* 67:641-666.
- Gläscher J, Hampton AN, O'Doherty JP (2009a) Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cereb Cortex* 19:483-495.
- Gläscher J (2009b) Visualization of group inference data in functional neuroimaging. *Neuroinformatics* 7:73-82.
- Gluth S, Rieskamp J, Büchel C (2012) Deciding when to decide: time-variant sequential sampling models explain the emergence of value-based decisions in the human brain. *J Neurosci* 32:10686-10698.
- Gold JI, Shadlen MN (2007) The neural basis of decision making. *Annu Rev Neurosci* 30:535-574.
- Hampton AN, Bossaerts P, O'Doherty JP (2006) The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J Neurosci* 26:8360-8367.
- Hampton AN, Bossaerts P, O'Doherty JP (2008) Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc Natl Acad Sci USA* 105:6741-6746.
- Hare T, Schultz W, Camerer CF, O'Doherty JP, Rangel A (2011) Transformation of stimulus value signals into motor commands during simple choice. *Proc Natl Acad Sci USA* 108:18120-18125.

- Heekeren HR, Marrett S, Bandettini PA, Ungerleider LG (2004) A general mechanism for perceptual decision-making in the human brain. *Nature* 431:859-862.
- Heekeren HR, Marrett S, Ruff DA, Bandettini PA, Ungerleider LG (2006) Involvement of human left dorsolateral prefrontal cortex in perceptual decision making is independent of response modality. *Proc Natl Acad Sci USA* 103:10023-10028.
- Heekeren HR, Marrett S, Ungerleider LG (2008) The neural systems that mediate human perceptual decision making. *Nat Rev Neurosci* 9:467-479.
- Hunt LT, Kolling N, Soltani A, Woolrich MW, Rushworth MFS, Behrens TEJ (2012) Mechanisms underlying cortical activity during value-guided choice. *Nat Neurosci* 15:470-476.
- Hunt LT, Dolan RJ, Behrens TE (2014) Hierarchical competitions subserving multi-attribute choice. *Nat Neurosci* 17:1613-1622.
- Jazayeri M, Shadlen MN (2010) Temporal context calibrates interval timing. *Nat Neurosci* 13:1020-1026.
- Jocham G, Hunt LT, Near J, Behrens TE (2012) A mechanism for value-guided choice based on the excitation-inhibition balance in prefrontal cortex. *Nat Neurosci* 15:960-961.
- Kim H (2014) Involvement of the dorsal and ventral attention networks in oddball stimulus processing: a meta-analysis. *Hum Brain Mapp* 35:2265-2284.
- Krajbich I, Armel C, Rangel A (2010) Visual fixations and the computation and comparison of value in simple choice. *Nat Neurosci* 13:1292-1298.
- Krajbich I, Rangel A (2011) Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proc Natl Acad Sci U S A* 108:13852-13857.
- Krajbich I, Lu D, Camerer C, Rangel A (2012) The attentional drift-diffusion model extends to simple purchasing decisions. *Front Cogn Sci* 3:193.
- Krugel L, Biele G, Mohr P, Li S, Heekeren HR (2009) Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proc Natl Acad Sci USA* 106:17591-17596.
- Larsen T, O'Doherty JP (2014) Uncovering the spatio-temporal dynamics of value-based decision-making in the human brain: a combined fMRI-EEG study. *Philos Trans R Soc Lond B: Biol Sci* 369(1655), 20130473.
- Lee SW, Shimojo S, O'Doherty JP (2014) Neural computations underlying arbitration between model-based and model-free learning. *Neuron* 81:687-699.

- Lim SL, O'Doherty JP, Rangel A (2011) The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *J Neurosci* 31:13214-13223.
- Lopez-Persem A, Domenech P, Pessiglione M (2016) How prior preferences determine decision-making frames and biases in the human brain. *Elife* 5:e20317.
- McGuire JT, Nassar MR, Gold JI, Kable JW (2014) Functionally dissociable influences on learning rate in a dynamic environment. *Neuron* 84:870-881.
- McNamee D, Rangel A, O'Doherty JP (2013) Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nat Neurosci* 16:479-485.
- Mulder MJ, Van Maanen L, Forstmann BU (2014) Perceptual decision neurosciences—a model-based review. *Neuroscience* 277:872-884.
- Nassar MR, Wilson RC, Heasly B, Gold JI (2010) An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *J Neurosci* 30:12366-12378.
- Nassar MR, Rumsey KM, Wilson RC, Parikh K, Heasly B, Gold JI (2012) Rational regulation of learning dynamics by pupil-linked arousal systems. *Nat Neurosci* 15:1040-1046.
- O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. *Neuron* 38:329-337.
- O'Doherty JP, Hampton A, Kim H (2007) Model-based fMRI and its application to reward learning and decision making. *Ann N Y Acad Sci* 1104:35-53.
- O'Doherty JP, Cockburn J, Pauli WM (2017) Learning, reward, and decision making. *Annu Rev Psychol* 68:73-100.
- O'Reilly JX, Jbabdi S, Behrens TE (2012) How can a Bayesian approach inform neuroscience? *Eur J Neurosci* 35:1169-1179.
- Payzan-LeNestour E, Dunne S, Bossaerts P, O'Doherty JP (2013) The neural representation of unexpected uncertainty during value-based decision making. *Neuron* 79:191-201.
- Pisauro A, Fouragnan E, Retzler C, Philiastides M (2017) Neural correlates of evidence accumulation during value-based decisions revealed via simultaneous EEG-fMRI. *Nat Commun* 8:1-9.
- Plassmann H, O'Doherty JP, Rangel A (2007) Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J Neurosci* 27:9984-9988.
- Polanía R, Krajbich I, Grueschow M, Ruff CC (2014) Neural oscillations and synchronization differentially support evidence accumulation in perceptual and value-based decision making. *Neuron* 82:709-720.

- Polanía R, Moisa M, Opitz A, Grueschow M, Ruff CC (2015) The precision of value-based choices depends causally on fronto-parietal phase coupling. *Nat Commun* 6:8090.
- Rangel A, Hare T (2010) Neural computations associated with goal-directed choice. *Curr Opin Neurobiol* 20:262-270.
- Ratcliff R (1978) A theory of memory retrieval. *Psychol Rev* 85:59-108.
- Ratcliff R, McKoon G (2008) The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput* 20:873-922.
- Ratcliff R, Smith PL, Brown SD, McKoon G (2016) Diffusion decision model: current issues and history. *Trends Cogn Sci* 20:260-281.
- Rodriguez CA, Turner BM, Van Zandt T, McClure SM (2015) The neural basis of value accumulation in intertemporal choice. *Eur J Neurosci* 42:2179-2189.
- Roitman JD, Shadlen MN (2002) Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J Neurosci* 22:9475-9489.
- Rushworth MF, Behrens TE (2008) Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat Neurosci* 11:389-397.
- Shadlen MN, Newsome WT (2001) Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J Neurophysiol* 86:1916-1936.
- Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ (2009) Bayesian model selection for group studies. *Neuroimage* 46:1004-1017.
- Summerfield C, Behrens TE, Koechlin E (2011) Perceptual classification in a rapidly changing environment. *Neuron* 71:725-736.
- Summerfield C, Tsetsos K (2012) Building bridges between perceptual and economic decision-making: Neural and computational mechanisms. *Front Neurosci* 6:1-20.
- Suzuki S, Harasawa N, Ueno K, Gardner JL, Ichinohe N, Haruno M, Cheng K, Nakahara H (2012) Learning to simulate others' decisions. *Neuron* 74:1125-1137.
- Suzuki S, Adachi R, Dunne S, Bossaerts P, O'Doherty JP (2015) Neural mechanisms underlying human consensus decision-making. *Neuron* 86:591-602.
- Suzuki S, Jensen ELS, Bossaerts P, O'Doherty JP (2016) Behavioral contagion during learning about another agent's risk-preferences acts on the neural representation of decision-risk. *Proc Natl Acad Sci USA* 113:3755-3760.
- Tavares G, Perona P, Rangel A (2017) The attentional drift diffusion model of simple perceptual decision-making. *Front Neurosci* 11:468.

- Towal RB, Mormann M, Koch C (2013) Simultaneous modeling of visual saliency and value computation improves predictions of economic choice. *Proc Natl Acad Sci U S A* 110:E3858-E3867.
- Usher M, McClelland JL (2001) The time course of perceptual choice: The leaky, competing accumulator model. *Psychol Rev* 108:550-592.
- Wästlund E, Shams P, Otterbring T (2018) Unsold is unseen... or is it? Examining the role of peripheral vision in the consumer choice process using eye-tracking methodology. *Appetite* 120:49-56.
- Wilson, RC, Nassar MR, Gold JI (2010) Bayesian online learning of the hazard rate in change-point problems. *Neural Comput* 22:2452-2476.
- Wunderlich K, Rangel A, O'Doherty JP (2009) Neural computations underlying action-based decision making. *Proc Natl Acad Sci U S A* 106:17199-17204.
- Yoshida W, Dolan RJ, Friston KJ (2008) Game theory of mind. *PLoS Comput Biol* 4:e1000254.
- Yoshida W, Seymour B, Friston KJ, Dolan RJ (2010) Neural mechanisms of belief inference during cooperative games. *J Neurosci* 30:10744-10751.
- Yu AJ, Dayan P (2005) Uncertainty, neuromodulation, and attention. *Neuron* 46:681-692.
- Yu AJ, Cohen J (2009) Sequential effects: superstition or rational behavior? In: *NIPS 2008, Advances in Neural Information Processing Systems*, Ed 21 (Koller D, Schuurmans D, Bengio Y, Bottou L, eds), pp1873-1880. Vancouver: MIT.