

RESEARCH ARTICLE

# Primary Visual Cortex as a Saliency Map: A Parameter-Free Prediction and Its Test by Behavioral Data

Li Zhaoping<sup>1\*</sup>, Li Zhe<sup>2</sup>

**1** University College London, London, United Kingdom, **2** Tsinghua University, Beijing, China

\* [z.li@ucl.ac.uk](mailto:z.li@ucl.ac.uk)



 OPEN ACCESS

**Citation:** Zhaoping L, Zhe L (2015) Primary Visual Cortex as a Saliency Map: A Parameter-Free Prediction and Its Test by Behavioral Data. *PLoS Comput Biol* 11(10): e1004375. doi:10.1371/journal.pcbi.1004375

**Editor:** Olaf Sporns, Indiana University UNITED STATES

**Received:** August 14, 2014

**Accepted:** June 1, 2015

**Published:** October 6, 2015

**Copyright:** © 2015 Zhaoping, Zhe. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The data are available at [www.cs.ucl.ac.uk/staff/Zhaoping.Li/](http://www.cs.ucl.ac.uk/staff/Zhaoping.Li/)

**Funding:** This work was supported by the Gatsby Charitable Foundation and by 985 fund of Tsinghua University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

It has been hypothesized that neural activities in the primary visual cortex (V1) represent a saliency map of the visual field to exogenously guide attention. This hypothesis has so far provided only qualitative predictions and their confirmations. We report this hypothesis' first quantitative prediction, derived without free parameters, and its confirmation by human behavioral data. The hypothesis provides a direct link between V1 neural responses to a visual location and the saliency of that location to guide attention exogenously. In a visual input containing many bars, one of them saliently different from all the other bars which are identical to each other, saliency at the singleton's location can be measured by the shortness of the reaction time in a visual search for singletons. The hypothesis predicts quantitatively the whole distribution of the reaction times to find a singleton unique in color, orientation, and motion direction from the reaction times to find other types of singletons. The prediction matches human reaction time data. A requirement for this successful prediction is a data-motivated assumption that V1 lacks neurons tuned simultaneously to color, orientation, and motion direction of visual inputs. Since evidence suggests that extrastriate cortices do have such neurons, we discuss the possibility that the extrastriate cortices play no role in guiding exogenous attention so that they can be devoted to other functions like visual decoding and endogenous attention.

## Author Summary

It has been hypothesized that neural activities in the primary visual cortex represent a saliency map of the visual field to exogenously guide attention. This hypothesis has so far provided only qualitative predictions and their confirmations. We report this hypothesis' first quantitative prediction, derived without free parameters, and its confirmation by human behavioral data. Using the shortness of reaction times in visual search tasks to measure saliency of the search target's location, the hypothesis predicts the quantitative distribution of the reaction times to find a salient bar unique in color, orientation, and motion direction in a background of bars that are identical to each other. The prediction matches experimental observations in human observers. Since the prediction would be

invalid without a particular neural property of the primary visual cortex, the extrastriate cortices may give little contribution to exogenous attentional guidance since they lack this neural property. Implications of this prospect on the framework of attentional network and the computational role of the higher brain areas are also discussed.

## Introduction

### Attentional selection and saliency

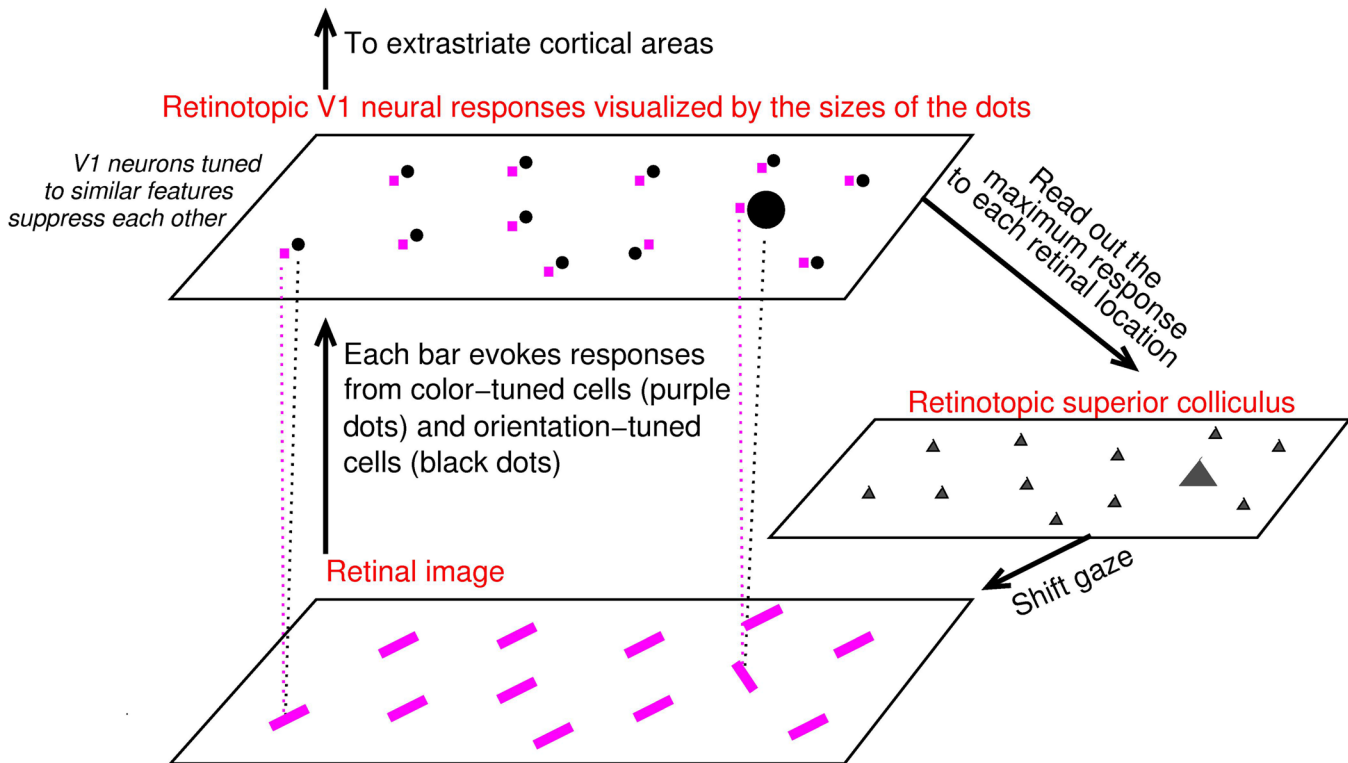
Spatial visual selection, often called spatial attentional selection, enables vision to select a visual location for detailed processing using limited cognitive resources [1]. Metaphorically, the selected location is said to be in the attentional spotlight, typically centered on the gaze position. An object outside the spotlight is difficult to recognize. Therefore, the reaction time (RT) to find a particular word on this page depends on how long it takes the spotlight to arrive at the word location. The spotlight is guided by goal-dependent (or top-down, endogenous) mechanisms, such as to direct our gaze to the right words while reading, and/or by goal-independent (or bottom-up, exogenous) mechanisms such as when our reading is distracted by a sudden drastic change in visual periphery.

In this paper, an input is said to be salient when it strongly attracts attention by bottom-up mechanisms, and the degree of this attraction is defined as saliency. For example, an orientation singleton such as a vertical bar in a background of horizontal bars is salient, so is a color singleton such as a red dot among many green ones; and the location of such a singleton has a high saliency value. Therefore, saliency of a visual location can often be measured by the shortness of the reaction time in a visual search to find a target at this location [2], provided that saliency, rather than top-down attention, dictates the variabilities of attentional guidance and reaction time. It can also be measured by attentional (exogenous) cueing effect, the degree in which a salient location speeds up and/or improves visual discrimination of a probe presented at this location immediately after a brief salient cue at the same location [3, 4].

Traditional views presume that higher brain areas, such as those in the parietal and frontal brain areas, are responsible for guiding attention exogenously [2, 5, 6, 1]. This belief was partly inspired by noting that saliency is a general property that could arise from visual inputs with any feature values (e.g., vertical or red) in any feature dimension (e.g., color, orientation, and motion) whereas neurons in lower visual areas like the primary visual cortex are (more likely) tuned to specific feature values (e.g., a vertical orientation) rather than being feature untuned.

### V1 saliency hypothesis: Its feature-blind nature, neural mechanisms, and qualitative experimental support

It was proposed a decade ago [7, 8] that V1 computes a saliency map, such that the saliency of a location is represented by the maximum response from V1 neurons to this location relative to the maximum responses to the other locations. It is only the V1 response vigor that matters for saliency, and not the preferred features of the responding neurons. For example, the image in Fig 1 contains many colored bars, each activates some V1 neurons tuned to its color and/or orientation. The maximum response to each bar signals the saliency of its location regardless of whether the V1 neuron giving this response is tuned to the color or orientation (or both color and orientation) of the bar. In another example, Fig 2A and 2B contain an orientation and color singleton, respectively, in the same background of uniformly feature bars. If the two images evoke the same background V1 responses to all the background locations, then the two

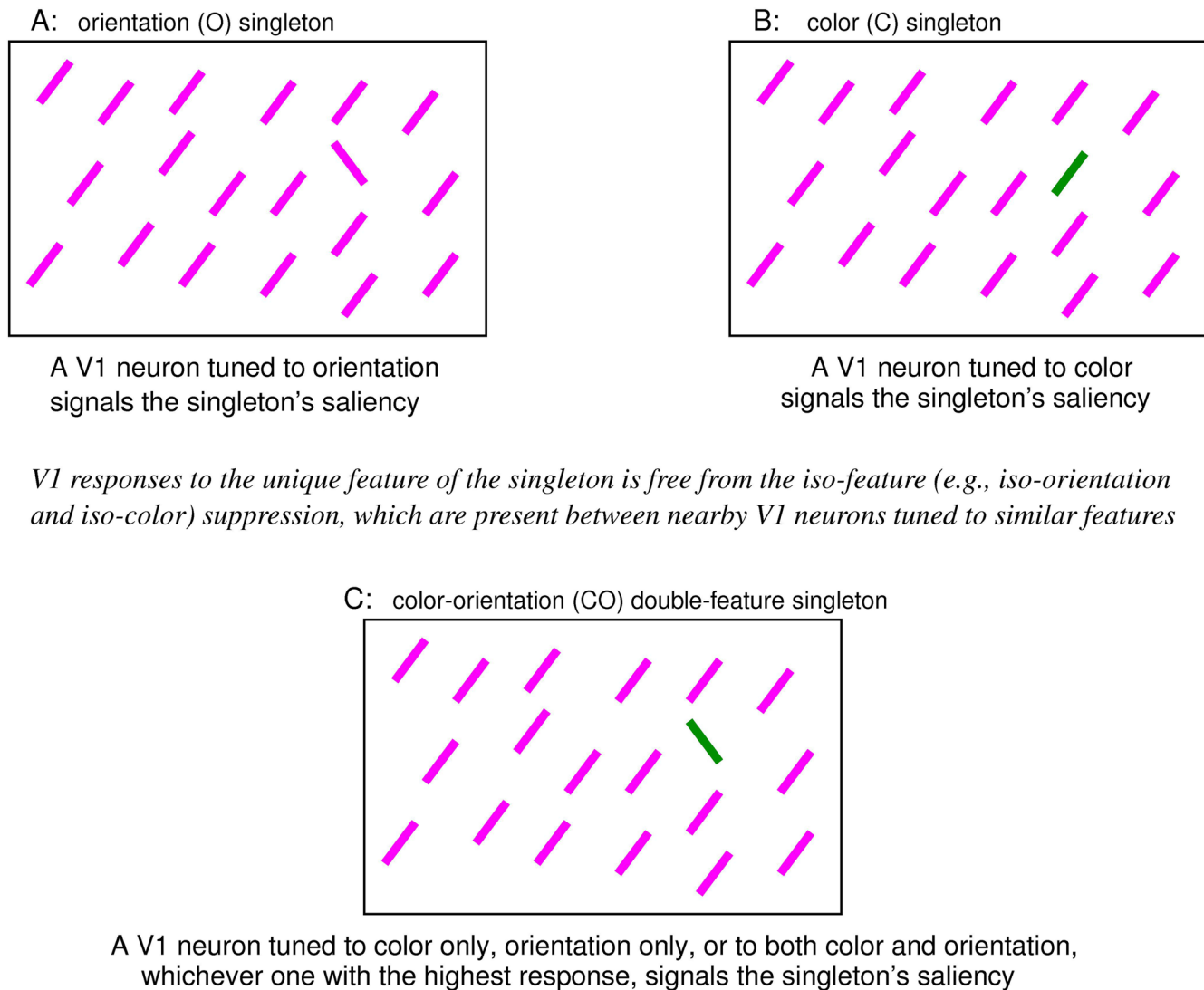


**Fig 1. V1 saliency hypothesis states that the bottom-up saliency of a location is represented by the maximum V1 response to this location.** In this schematic, V1 is simplified to contain only two kinds of neurons, one tuned to color (their responses are visualized by the purple dots) and the other tuned to orientation (black dots). Each input bar evokes responses in a cell tuned to its color and another cell tuned to its orientation (indicated for two input bars by linking each bar to its two evoked responses by dotted lines), and the receptive fields of these two cells cover the same bar location even though (for better visualization) the dots representing these cells are not overlapping in the cortical map. Iso-feature suppression makes nearby V1 neurons tuned to similar features (e.g., similar color or similar orientation) suppress each other. The orientation singleton in this image evokes the highest V1 response to this image because the orientation-tuned neuron responding to it escapes iso-orientation suppression. The color tuned neuron tuned and responding to the singleton's color is under iso-color suppression. The saliency map is likely read out by the superior colliculus to execute gaze shifts to salient locations [9].

doi:10.1371/journal.pcbi.1004375.g001

singletons are equally salient if they evoke the same level of maximum response even if the maximum response is evoked in an orientation-tuned cell in one image and a color-tuned cell in the other; conversely, if the two singletons differ by their respective maximally evoked responses, then the singleton evoking the higher response is more salient regardless of the preferred features of the responding neurons.

The feature-blind nature of this saliency representation in V1 enables the brain to have a bottom-up saliency map in V1 in terms of the various maximum V1 responses for various locations, despite the feature tuning of V1 neurons, without resorting to higher cortical areas such as the frontal eye field or lateral-intraparietal cortex [10, 1]. This saliency map may potentially be read out by the superior colliculus, which receives monosynaptic input from V1 and controls eye movement to execute the attentional selection [11]. If an observer searches for a uniquely oriented bar in the retinal image in Fig 1, the reaction time to find this target bar, associated with the saliency of the target location, should thus be associated with the maximum V1 response to this location. In particular, a shorter reaction time should result from a larger



**Fig 2. Schematics of visual stimuli for singleton searches.** Due to iso-feature suppression, the highest response to each image is from a neuron responding to the singleton bar. This most activated neuron is tuned to orientation for image A, tuned to color for image B, and to color, orientation, or both features of the singleton for image C. The maximum V1 response to the singleton signals the saliency of its location.

doi:10.1371/journal.pcbi.1004375.g002

value of the maximum response to the target location (when the maximum responses to various non-target locations are fixed).

The neural mechanisms in V1 to compute saliency is intracortical interactions that cause contextual influences, making a neuron's response to inputs within its receptive field dependent on contextual inputs [12, 13, 14]. One particular form of contextual influences is iso-orientation suppression between nearby neurons tuned to same or similar orientations. It makes orientation-tuned neurons responding to neighboring background bars in Fig 1 suppress each other because they are tuned to the same orientation of these bars, whereas the neuron responding to the orientation singleton escapes such suppression because it is tuned to a very different orientation of the singleton. Hence, the orientation singleton in Fig 1 is the most

salient because a V1 neuron, with its receptive field covering the bar, responds more vigorously than any neuron responding to the background bars. Throughout the paper, 'a neuron responding to a bar' means the most responsive neuron among a local population of neurons with similar input selectivities responding to this bar regardless of the number of neurons in this local population.

In addition to the orientation feature, V1 neurons are also tuned to other input feature dimensions including color, motion direction, and eye of origin [15, 16]. Hence, each colored bar in the retinal image of Fig 1 evokes not only a response in a cell tuned to its orientation but also another response in another cell tuned to its color (omitting other input features for simplicity), this is indicated by the dotted lines linking the two example input bars and their respective evoked V1 responses. In general, there are many V1 neurons whose receptive fields cover the location of each visual input item (including neurons whose preferred orientations or colors do not match the visual input feature), and only the highest response from these neurons represents the saliency of this location according to the V1 saliency hypothesis (note that this highest response is unlikely to be from a neuron whose preferred feature is not in the input item). In the example of Fig 1, responses from the color-tuned neurons to all bars suffer from iso-color suppression [17], which is analogous to iso-orientation suppression, since all bars have the same color. Focusing on V1 neurons tuned to color only and neurons tuned to orientation only for simplicity, the highest response evoked by the orientation singleton is in the orientation-tuned rather than the color-tuned cell, and this response alone (relative to the maximum responses to the background bars) determines the saliency of the orientation singleton. Later in the paper, the notion that many V1 neurons respond to a single input location or item will be generalized to include neurons tuned to motion direction and neurons jointly tuned to multiple feature dimensions. Determining the highest V1 response to each input location will involve determining which of the many neurons whose receptive fields cover this location has the highest response.

Analogous to iso-orientation suppression and iso-color suppression, iso-motion-direction and iso-ocular-origin suppressions are also present in V1 [12, 13, 14, 18, 19, 20], and we call them iso-feature suppression in general [7]. Accordingly, an input singleton in any of these feature dimensions should be salient (see Fig 2B for a color singleton), since the neuron responding to the unique feature of the singleton escapes the iso-feature suppression from the neurons responding to the uniformly featured background items. This is consistent with known behavioral saliency and has led to the successful prediction of the salient singleton in eye-of-origin [21]. Iso-feature suppression is the dominant form of contextual influences, and it is believed to be mediated by intracortical neural connections [22, 23] linking neurons whose receptive fields are spatially nearby but not necessarily overlapping. A neural circuit model of V1 [24, 25, 7, 26, 27] with its intracortical interactions has successfully explained many prototypical visual search and segmentation examples by using the model responses to predict a saliency map which in turn predicts the relative degrees of ease in the visual behavior associated with the saliencies of the task relevant locations.

Although the V1 saliency hypothesis is a significant departure from traditional psychological theories, it has received substantial experimental support [28, 29, 30, 31, 21, 32, 33], detailed in [9]. In particular, behavioral data confirmed the surprising prediction from this hypothesis that an eye-of-origin singleton (e.g., an item uniquely shown to the left eye among other items shown to the right eye) that is hardly distinctive from other visual inputs can attract attention and gaze qualitatively just like, or quantitatively more strongly than, a salient and highly distinctive orientation singleton does [21, 33]. This finding provides a hallmark of the saliency map in V1 because, cortical neurons, except many in V1, are not tuned to eye-of-origin feature [34, 35], making this feature non-distinctive to perception. Furthermore, behavioral data

confirmed that saliency is represented by the maximum rather than the weighted summation or the average of responses to a visual location [30, 29]. Functional magnetic resonance imaging and event related potential measurements also confirmed that, when top-down confounds are avoided or minimized, a salient location evokes brain activations in V1 but not in the parietal and frontal regions [32].

## The current study

So far, predictions and experimental tests of the V1 saliency hypothesis have been *qualitative*. Here, we report its first *quantitative* prediction derived without free parameters. The predicted quantity is the distribution of the reaction times in a visual search for a singleton bar unique simultaneously in color, orientation, and motion direction among uniformly featured background bars. We will derive a precise mathematical relationship between this quantity and the distributions of the reaction times to search for other types of singleton bars, thus enabling us to predict this quantity from the observed reaction times for the other singletons. This mathematical relationship requires, other than the V1 saliency hypothesis, only the following two *qualitative* features in neural physiology: (1) the feature-tuned neural interaction, in particular iso-feature suppression that depends on whether the preferred features of the interacting neurons are similar and causes higher responses to feature singletons, and (2) an assumption motivated by data that V1 does not have neurons tuned simultaneously to color, orientation, and motion direction. It does not depend on other details, e.g., colinear facilitation [36, 37] between V1 neurons and its contrast dependence [38, 39, 20]; otherwise, currently imprecise knowledge of V1 physiology (e.g., its intracortical interactions), which may vary with adaptation state and experience of observers, would have prevented the prediction to be parameter-free.

Furthermore, we show that this prediction quantitatively matches previously collected behavioral data [29]. We develop data analysis methods to obtain the predicted distribution of the reaction times for one type of feature singletons from the observed reaction times for the other types of feature singletons, and compare the predicted quantity with its behavioral counterpart using a custom designed statistical test. We further show that our data have a sufficient statistical power to falsify some spurious predictions that are likely incorrect based on V1 physiology. Since the same set of behavioral data and analysis methods are used to test the spurious predictions and our (non-spurious) prediction, we conclude that our (non-spurious) prediction is confirmed within the resolution provided by the statistical power in our data.

In addition, this paper explores the implications of the experimental confirmation of our quantitative (non-spurious) prediction. We will discuss experimental evidence on whether the extrastriate cortical areas also possess the two required physiological features for the prediction and thus whether they can be excluded from playing a role in saliency. Parts of this work have been presented in abstract form elsewhere [40, 41].

## Results

In this section, we show a direct link between the reaction time to find a visual feature singleton in a homogeneous background (like that in Fig 1) and the highest V1 response to this singleton. From this link, we derive the quantitative prediction and present its experimental test. In this process, we also present some related but spurious predictions that should be violated unless certain conditions on the V1 neural mechanisms hold. These spurious predictions and their tests (falsification) by behavioral data not only provide further insights in the underlying neural mechanisms but also verify that our methods can use our behavioral data to falsify a prediction.



## Linking V1 responses with reaction times

When the effect of top-down attentional guidance is negligible or held constant in a visual search task, a higher saliency at the target location should lead to a shorter reaction time to find the target, by the definition of saliency. In stimuli like those in Fig 2, the feature singletons are assumed as salient enough to dictate immediate attention shifts. The latency of the attentional shift to the singleton is shorter for a more salient singleton. Assuming a fixed additional latency from this attention shift to an observer's response to report the singleton, then the reaction time for the visual search task, e.g., for the reporting response, is determined by the singleton's saliency.

Let a visual scene have visual input items at  $n$  locations  $i = 1, 2, \dots, n$ , and let  $r_i$  be the maximum V1 response evoked by location  $i$ . Then the saliency of location  $i$  is determined by  $r_i$  relative to the other  $r_j$  for  $j \neq i$ . This is because, according to the V1 saliency hypothesis, saliency read-out process is like an auction for attention, with  $r_i$  the bidding price for attention by location  $i$ , such that the location giving the highest bid is the most likely to win attention [42]. Let us order  $i$  such that

$$r_1 \geq r_2 \geq r_3 \geq \dots \geq r_n, \tag{1}$$

then, we can use a function  $g(\cdot)$  to formally describe

$$\text{saliency at the most salient (when } i = 1 \text{) location} = g(r_1|r_2, r_3, \dots, r_n). \tag{2}$$

This paper is only concerned with scenes like those in Fig 2, and calls each such scene a feature singleton scene. Such a scene has one feature singleton in a background of many items that are identical to each other, and the singleton is far more salient than any other input item. Then,  $r_1$  is the maximum response to the singleton and is substantially and significantly larger than any  $r_i$  for  $i > 1$  (e.g.,  $r_1 > 20$  spikes/second and  $r_i < 10$  spikes/second for  $i > 1$ ). When  $n$  is very large (e.g., 660 in the visual stimuli we will use later), we can reasonably expect that  $g(r_1|r_2, r_3, \dots)$  depends on  $(r_2, r_3, \dots)$  mainly through the statistical properties across the  $r_i$ 's (for  $i > 1$ ) rather than the exact value of each  $r_i$ . Let the statistical properties be partly characterized by the average  $\bar{r}$  and standard deviation  $\sigma$  across  $(r_2, r_3, \dots, r_n)$ ; then a singleton with a larger  $(r_1 - \bar{r})/\sigma$ , and perhaps also a larger  $r_1/\bar{r}$ , tends to be more salient [7]. More strictly, the function  $g(r_1|r_2, r_3, \dots, r_n)$  may also depend on the locations of visual inputs for all  $i$ . However, we assume that this dependence is negligible in this paper since we are only concerned with singleton scenes satisfying the following condition: (1) the eccentricity of the singleton from the center of the visual field is the same across all singleton scenes, (2) different non-singleton items evoke sufficiently similar maximum responses  $r_i$  for  $i > 1$ , and (3) the distribution of the locations of the non-singleton items is approximately the same across all singleton scenes.

If two scenes are identical to each other in terms of the number  $n$  of visual input locations and the distribution of the responses  $r_2, r_3, \dots, r_n$ , we say that they share an *invariant background response distribution*. The three singleton scenes in Fig 2 are approximately sharing an invariant background response distribution, even though the highest response  $r_1$  to the singleton may be larger in Fig 2C than Fig 2A 2B. This is because the response  $r_i$  to each background bar  $i > 1$  is determined by the bar itself and by its surrounding neighbors which exert contextual influence (mainly iso-feature suppression) on the response, the singleton can at best be the least influential neighbor since its most activated neuron exerts a limited or negligible iso-feature suppression on neurons most responsive to the background bar and preferring a very different feature. Hence the singleton has a negligible influence on the statistical properties of the background responses, which are determined by such characteristics as the contrast, density, and the degree of regularities in the locations of the background bars.

Assuming an invariant background response distribution shared by a set of feature singleton scenes, we can omit the explicit expression of  $(r_2, r_3, \dots)$  in Eq (2) and write (still using the same notation  $g(\cdot)$  for convenience)

the saliency of the singleton location =  $g(r_1)$ ,  $r_1$  is the highest response to the singleton. (3)

The  $g(r)$  monotonically increases with  $r$  in a way that is determined by the properties of the invariant background response distribution. Since a larger saliency at the singleton location gives a shorter reaction time to find it (assuming again negligible or constant top-down factors), we can write

the reaction time for a feature singleton =  $f(r_1)$ ,  $r_1$  is the highest response to the singleton, (4)

and  $f(r_1)$  is a monotonically decreasing function of  $r_1$ . The exact form of  $f(r)$  should depend on the invariant background response distribution, the saliency read-out system, and the observer (e.g., some observers can respond faster than others). We will see that the details of  $f(r)$  do not matter as long as  $f(r)$  monotonically decreases with  $r$ . With  $f(r)$ , the reaction time for a feature singleton is directly linked to its maximum evoked V1 responses.

### A previously known race model in reaction times can be derived from a toy V1

Let us call the singletons in Fig 2A, Fig 2B, and Fig 2C (which share an invariant background response distribution) O, C, and CO singletons, respectively, by the feature dimension(s) in which the singleton has a unique feature. The C and O singletons are single-feature singletons and the CO singleton is a double-feature singleton. Let a toy V1 have only two kinds of neurons, one tuned to color only and one tuned to orientation only, and assume that V1 responses are deterministic rather than stochastic given a visual input. The toy V1 and the deterministic nature of V1 responses are both temporary simplifications to illustrate the method, and these simplifications will be removed later. Let  $r^O$  or  $r^C$ , respectively, be the response of the orientation-tuned neuron or the color-tuned neuron to the singleton in Fig 2A or Fig 2B, respectively. They are also the highest responses to the respective singletons due to iso-feature suppression. Then, according to Eq (4), the reaction times  $RT_O$  and  $RT_C$  to find the O and C singletons, respectively, are

$$RT_O = f(r^O) \quad \text{and} \quad RT_C = f(r^C). \quad (5)$$

The CO singleton in Fig 2C should evoke higher responses in both the neuron tuned to its unique orientation and the neuron tuned to its unique color than the responses to the background bars, again due to iso-feature suppression. Furthermore, we assume that the response property of the orientation-tuned neuron and the contextual influences on it are not affected by the color of the visual input, so that  $r^O$  is the same to the O and CO singletons. Analogously, the response  $r^C$  of the color-tuned neuron is assumed the same to the C and CO singletons. The maximum V1 response to the CO singleton is  $\max(r^C, r^O)$  (where  $\max(\cdot)$  means the maximum value among the arguments). Hence, the reaction time  $RT_{CO}$  to find the CO singleton is

$$RT_{CO} = f[\max(r^C, r^O)] = \min[f(r^C), f(r^O)] = \min(RT_C, RT_O), \quad (6)$$

when we combine Eqs (4) and (5) and note that  $f(\cdot)$  is a monotonically decreasing function ( $\min(\cdot)$  means the minimum value of the arguments). The equation

$$RT_{CO} = \min(RT_C, RT_O) \quad (7)$$



describes the deterministic version of a race model, often used to model a behavioral reaction time as the shorter reaction time of two or more underlying processes [43], as if (e.g.,) the reaction time for the CO singleton is the winning reaction time in a race between two racers, C and O singletons, with their respective reaction times. Here we see (see also [29]) that this model can arise from the neural substrates, given the V1 saliency hypothesis, if V1 has only neurons tuned to orientation only and neurons tuned to color only but no neurons tuned to both. This is because by such a V1 the double-feature singleton is as salient as the more salient of the two single-feature singletons. We note that this race model arises regardless of the details of  $f(r)$  as long as it is a monotonically decreasing function.

V1 responses are actually stochastic, each a random sample from a specific distribution. To proceed, we assume the following two conditions. First, there are sufficiently many background items that the statistical properties of the invariant background response distribution (e.g., the mean and standard deviation across the responses to the background items) are not stochastic despite the stochasticity of the individual responses. Second, the singletons are salient enough that their evoked responses  $r^C$  and  $r^O$  are always larger than any responses to the background. By Eq (5), the stochastic  $r^C$  and  $r^O$  make  $RT_C$  and  $RT_O$  also stochastic. For example, if  $P_{r^O}(r^O)$  is the probability density of  $r^O$ , then the probability density of  $RT_O$  is

$$P_{RT_O}(RT_O) = \left[ P_{r^O}(r^O) \left( \frac{df(r^O)}{dr^O} \right)^{-1} \right], \quad \text{at } r^O = f^{-1}(RT_O). \tag{8}$$

In any case,  $RT_{CO} = f[\max(r^C, r^O)] = \min[f(r^C), f(r^O)]$  still holds. If the trial to trial fluctuations of  $r^C$  and  $r^O$  are regardless of the visual input in the feature dimension in which the neuron is untuned, and if they fluctuate independently of each other in the responses to the CO singleton, then the deterministic equation  $RT_{CO} = \min(RT_C, RT_O)$  becomes

$$\text{Distribution of } RT_{CO} = \text{Distribution of } \min(RT_C, RT_O), \tag{9}$$

in which  $RT_C$  and  $RT_O$  are independent random samples from their respective distributions. The average of  $RT_{CO}$  will be shorter than *both* the average of  $RT_C$  and the average of  $RT_O$ , due to statistical facilitation, since each sample of  $RT_{CO}$  is the race winner of the two samples  $RT_C$  and  $RT_O$ . For simplicity, Eq (9) is written by this shorthand

$$RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O), \tag{10}$$

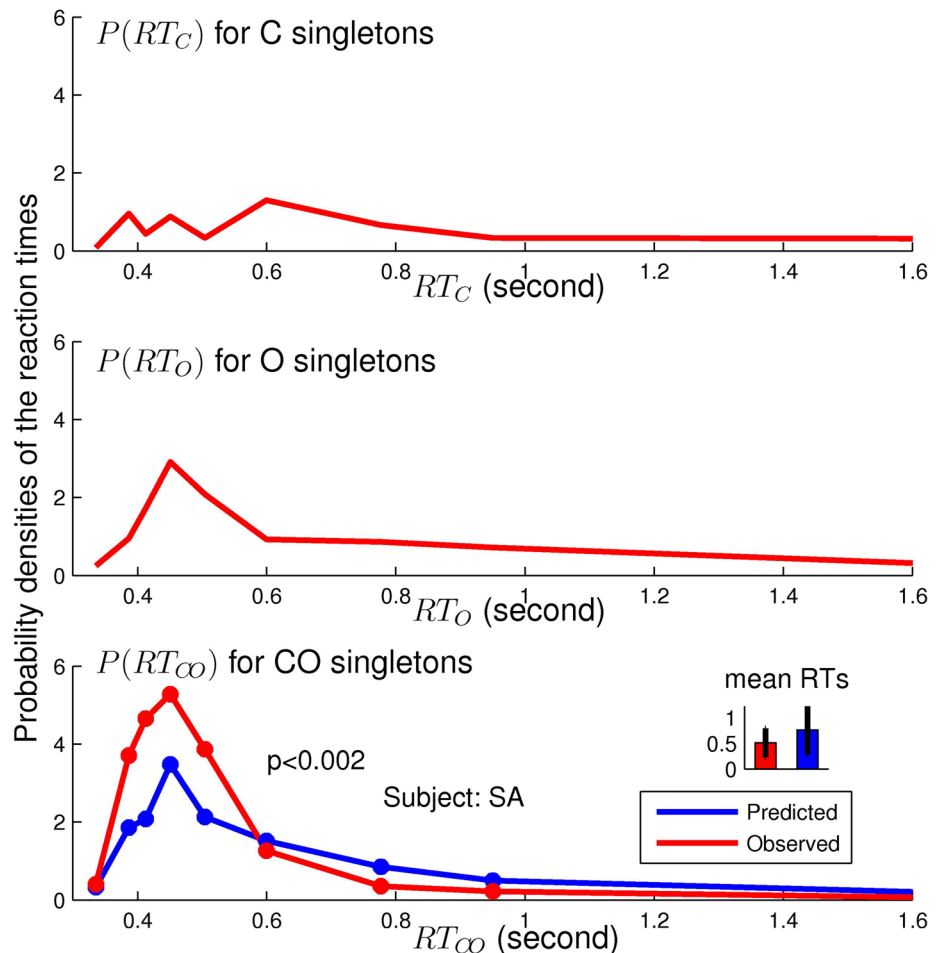
with  $x \stackrel{P}{=} y$  to mean that  $x$  and  $y$  have the same probability distribution.

The race model, or race equality,  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$  is a prediction of the V1 saliency hypothesis if one were hypothetically to assume a toy V1 that has no V1 neuron which can respond more vigorously to the CO singleton than the orientation-only-tuned neuron and the color-only-tuned neuron. This assumption is wrong. Hence  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$  is called a spurious race equality and its predicted distribution of  $RT_{CO}$  from experimentally observed distribution of  $\min(RT_C, RT_O)$  is called a spurious prediction.

### The spurious race equality $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$ is violated

Fig 3 shows that the spurious prediction of the distribution of  $RT_{CO}$  is significantly different from the distribution of the behaviorally observed  $RT_{CO}$ , with a  $p$  value  $p < 0.002$  in the statistical test of the null hypothesis that the predicted and the observed distributions of  $RT_{CO}$  are the same. (See the Methods section for how to obtain the prediction and the  $p$  value). The behavioral  $RT_{CO}$  values are significantly shorter than the predicted ones.

Observed and predicted distributions of various reaction times (RTs)

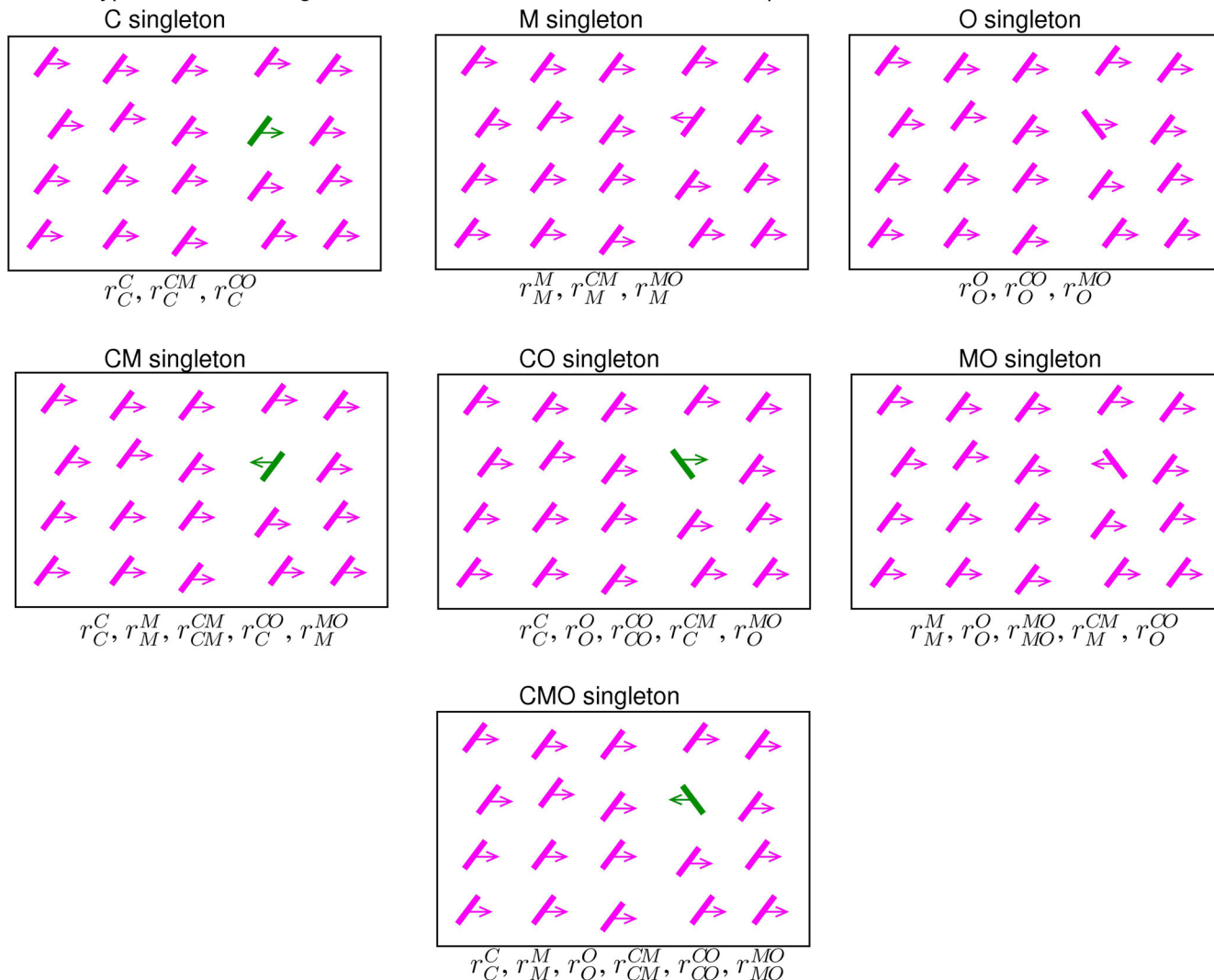


**Fig 3. Behavioral refutation of a spurious prediction  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$  based on the incorrect assumption that V1 lacks neurons tuned simultaneously to both orientation and color.** The graphs show distributions (in discrete time bins) of  $RT_O$ ,  $RT_C$ , and  $RT_{CO}$  (and the average and the standard deviation of  $RT_{CO}$ ) of a particular observer SA in searches of the singletons. Experimental data are shown in red, the prediction is in blue. The predicted and actual distributions of  $RT_{CO}$  are significantly different from each other, indicated by a  $p < 0.002$  in the bottom plot.

doi:10.1371/journal.pcbi.1004375.g003

With motion direction as another feature dimension, a feature singleton in motion direction, an M singleton, is the analogy of a C or O singleton. Analogous to a CO singleton, a double-feature singleton CM or MO is unique in both color and motion direction, or in both motion direction and orientation, respectively. A triple-feature CMO singleton is unique in all the three feature dimensions. Fig 4 shows the schematics of all the seven types of singletons. Let the reaction times to find singletons M, CM, MO, and CMO be  $RT_M$ ,  $RT_{CM}$ ,  $RT_{MO}$ , and  $RT_{CMO}$ , respectively. Then the spurious equality  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$  has the following

Seven types of feature singleton stimuli; the most activated neural responses are indicated under each stimulus



**Fig 4. Schematics of the seven kinds of feature singleton scenes.** Each bar is colored green or purple (of the same luminance in the behavioral experiment), tilted to the left or right from vertical by the same absolute tilt angle, moving to the left or right (indicated by an arrow pointing to left or right) by the same motion speed. Under each schematic, the non-trivial neural responses (e.g., these responses are expected to be substantially higher than the responses to the background bars) evoked by the singleton are listed. Each singleton scene here is called a purple scene (in this paper) to denote that the color of the background bars are purple. Swapping between the green and purple colors changes a purple scene into a green scene. All purple scenes are assumed to share an invariant background response distribution, so are all the green scenes. The behavioral experiment [29] randomly interleaved purple and green scenes between trials.

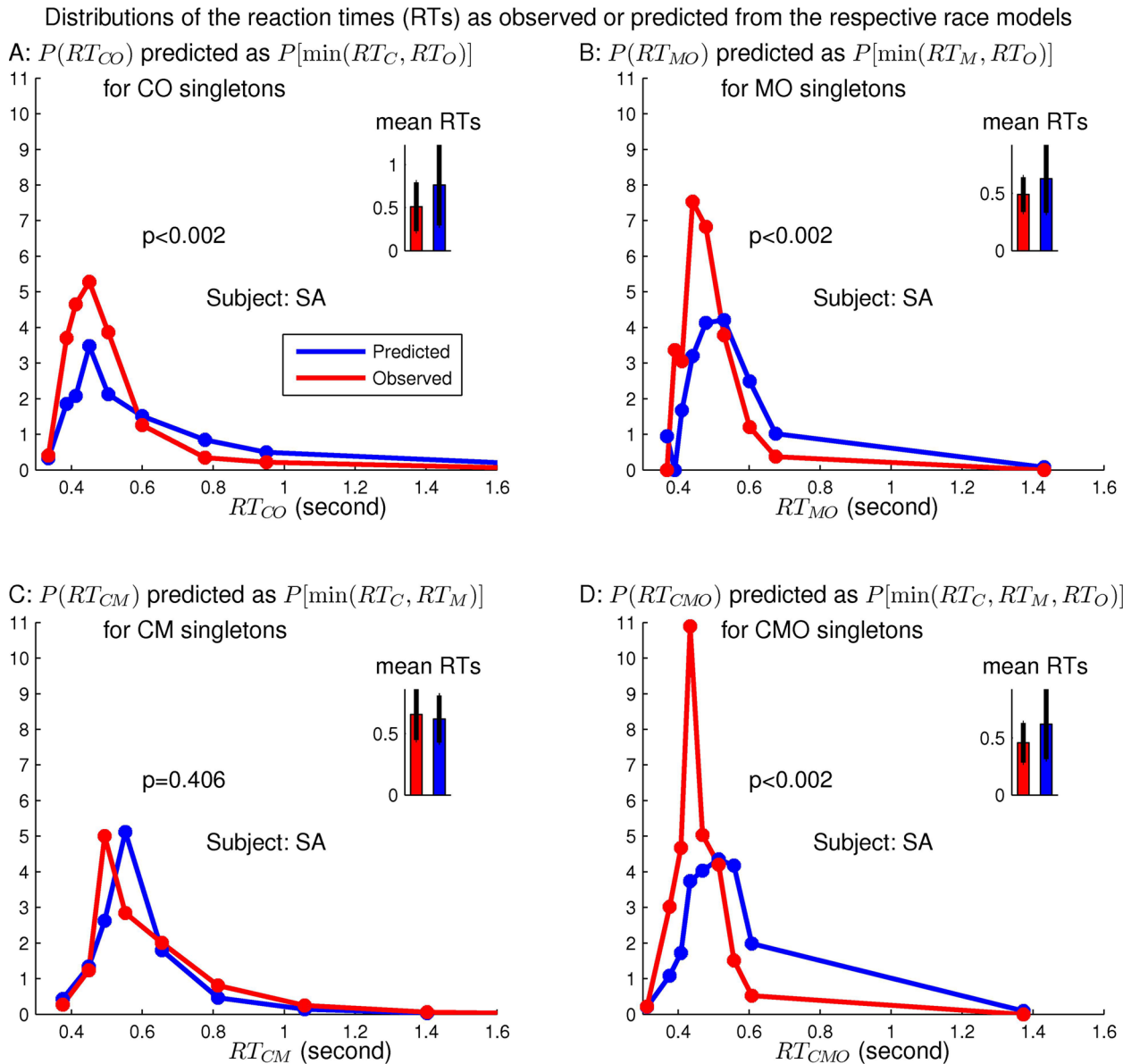
doi:10.1371/journal.pcbi.1004375.g004

generalizations:

$$RT_{CM}^P \stackrel{P}{=} \min(RT_C, RT_M), \tag{11}$$

$$RT_{MO}^P \stackrel{P}{=} \min(RT_M, RT_O), \text{ and} \tag{12}$$

$$RT_{CMO}^P \stackrel{P}{=} \min(RT_C, RT_M, RT_O). \tag{13}$$



**Fig 5. The observed and predicted distributions of reaction times for a double- or triple-feature singleton, using four different race models (race equalities),  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$  (in panel A),  $RT_{MO} \stackrel{P}{=} \min(RT_M, RT_O)$  (in panel B),  $RT_{CM} \stackrel{P}{=} \min(RT_C, RT_M)$  (in panel C), or  $RT_{CMO} \stackrel{P}{=} \min(RT_C, RT_M, RT_O)$  (in panel D), in a race between the corresponding racers whose reaction times are those of the corresponding single-feature singletons. The data are from the same subject SA already shown in Fig 3, panel A shows the same information as that in the bottom panel of Fig 3. The predicted and observed distributions are significantly different from each other except in panel C.**

doi:10.1371/journal.pcbi.1004375.g005

Each equality above holds when V1 is assumed to have no neurons, i.e., the CM, MO, CO, or CMO neurons, which are tuned to more than one feature dimension and can respond more vigorously to the corresponding double-feature (or triple-feature) singleton than it does to the corresponding singleton-feature singletons. Each equality predicts the distribution of the reaction times for a double- or triple-feature singleton from the observed reaction times for the corresponding single-feature singletons. Using data from the same observer as that in Fig 3, Fig 5 shows that other than  $RT_{CM}$ , the predictions disagree with the behavioral observations.

### V1 neurons tuned conjunctively to color and orientation predict that $RT_{CO}$ is likely shorter than predicted by the race model

Here we show that, because real V1 contains neurons (we call CO neurons) that are tuned simultaneously to color and orientation [16], the predicted  $RT_{CO}$  using  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$  can be longer than the observed  $RT_{CO}$ . Neurons tuned to color or orientation only are referred to as C or O neurons. Let  $r^{CO}$  denote the response of the CO neuron to the CO singleton, which thus evokes a maximum response  $\max(r^C, r^O, r^{CO})$ . According to Eq (4),

$$RT_{CO} = f[\max(r^C, r^O, r^{CO})]. \tag{14}$$

A CO neuron also responds to a C or O singleton matching its preferred color and orientation. For example, each of the C, O, and CO singletons in Fig 4 evokes a vigorous response in a CO neuron preferring its color and orientation. We use  $r_\alpha^{CO}$  to denote such a response of a CO neuron to a singleton  $\alpha = C, O,$  or  $CO$ . Then  $r_C^{CO}$  suffers from iso-orientation suppression (since the C singleton has the same orientation as the background bars),  $r_O^{CO}$  suffers from iso-color suppression, and  $r_{CO}^{CO}$  is free from iso-feature suppressions. For completeness,  $r_B^{CO}$  denotes a CO neuron’s response to a background bar matching its preferred features. Since  $r_B^{CO}$  suffers from both iso-color and iso-orientation suppressions it is likely that  $r_B^{CO} < r_\alpha^{CO}$  for  $\alpha = C, O,$  and  $CO$ .

Our notations for the responses ignore the binary tilt direction (clockwise or anticlockwise from vertical), color (isoluminant purple or green), or motion direction (leftward or rightward) of our singletons. This is because, in terms of evoked V1 response levels under contextual influences, reflection symmetry is assumed between the two tilt directions and between the two motion directions in our singleton scenes (all bars in all scenes have the same absolute angle from vertical and the same absolute motion speed). If a symmetry for V1 responses is not assumed between the two isoluminant colors with associated contextual influences, then our notations and derivations are only applicable when all singleton scenes are restricted to those with a given (e.g., purple) color of the background bars. For convenience, we call our singleton scenes with purple or green background bars purple or green scenes, respectively. For example, all the scenes in Fig 4 are purple. Without the color symmetry, the behavioral data from the purple scenes should be analyzed separately from those from the green scenes.

For consistency, we similarly use  $r_\alpha^C$  and  $r_\alpha^O$  to denote C and O neural responses to a singleton bar  $\alpha = C, O,$  and  $CO$  or a background bar  $\alpha = B$ . For example, the responses of the C neuron to the four kinds of bars are  $r_C^C, r_O^C, r_{CO}^C,$  and  $r_B^C$ . We have previously ignored  $r_O^C$  and identified  $r_{CO}^C$  with  $r_C^C$  since we argued that

$$r_B^C \stackrel{P}{=} r_O^C \quad \text{with iso-color suppression, and} \quad r_C^C \stackrel{P}{=} r_{CO}^C \quad \text{without iso-color suppression,} \tag{15}$$

because a C neuron’s response should be regardless of the orientation feature. Similarly, the O neuron’s response should be regardless of the input color (since the green and purple bars have the same high luminance contrast against a dark background) and have the following two types of responses,

$$r_B^O \stackrel{P}{=} r_C^O \quad \text{and} \quad r_O^O \stackrel{P}{=} r_{CO}^O. \tag{16}$$

Neural responses such as  $r_O^C$  ( $\stackrel{P}{=} r_B^C$ ) and  $r_C^O$  ( $\stackrel{P}{=} r_B^O$ ) that can be statistically equated with the same neurons’ responses to a background bar will be called trivial responses.

Note that the meaning of, e.g., C, in our mathematical expression depends on whether it is a superscript or a subscript. As a superscript in, e.g.,  $r^C$  it means that the neuron giving the response is tuned to the color (C) feature; as a subscript in, e.g.,  $r_C^O$  or  $RT_C$  it means the input

bar evoking the response or reaction time is a color (C) singleton. Without loss of validity, responses from neurons preferring feature(s) different from the feature(s) of the bars are ignored, since they are always smaller and do not affect saliency dictated by the maximum response to each location.

Combining Eq (4) with the equations above, we have

$$RT_C = f[\max(r_C^c, r_C^o, r_C^{co})] \stackrel{p}{=} f[\max(r_C^c, r_B^o, r_C^{co})], \tag{17}$$

$$RT_O = f[\max(r_O^c, r_O^o, r_O^{co})] \stackrel{p}{=} f[\max(r_B^c, r_O^o, r_O^{co})], \tag{18}$$

$$RT_{CO} = f[\max(r_{CO}^c, r_{CO}^o, r_{CO}^{co})] \stackrel{p}{=} f[\max(r_C^c, r_O^o, r_{CO}^{co})]. \tag{19}$$

Since a C singleton is more salient than a background bar, by V1 saliency hypothesis, its maximum evoked response  $\max(r_C^c, r_B^o, r_C^{co})$  must be larger than the maximum response  $\max(r_B^c, r_B^o, r_B^{co})$  to a background bar, i.e.,  $\max(r_C^c, r_B^o, r_C^{co}) > \max(r_B^c, r_B^o, r_B^{co})$ . Combining this with  $\max(r_B^c, r_B^o, r_B^{co}) \geq r_B^o$  gives  $\max(r_C^c, r_B^o, r_C^{co}) > r_B^o$ , consequently  $\max(r_C^c, r_B^o, r_C^{co}) = \max(r_C^c, r_C^{co})$ . Similarly  $\max(r_B^c, r_O^o, r_O^{co}) = \max(r_O^o, r_O^{co})$ . Hence, we can ignore  $r_B^c$  and  $r_B^o$  in Eqs (17)–(18) to have

$$RT_C = f[\max(r_C^c, r_C^{co})] \quad \text{and} \quad RT_O = f[\max(r_O^o, r_O^{co})]. \tag{20}$$

The above two equations are just examples of the following equation for our singleton scenes:

$$\text{reaction time to a singleton} = f[\max(\text{list of all the non-trivial responses to this singleton})]. \tag{21}$$

This can be seen by noting that a  $r_C^o$  ( $\stackrel{p}{=} r_B^o$ ) is a trivial response (i.e., statistically the same as the neuron’s response to a background bar) to a C singleton whereas  $r_C^c$  ( $\stackrel{p}{=} r_B^c$ ) is a trivial response to an O singleton. From Eq (20),

$$\begin{aligned} \min(RT_C, RT_O) &= \min\{f[\max(r_C^c, r_C^{co})], f[\max(r_O^o, r_O^{co})]\} \\ &= f\{\max[\max(r_C^c, r_C^{co}), \max(r_O^o, r_O^{co})]\} \\ &= f[\max(r_C^c, r_O^o, r_C^{co}, r_O^{co})], \end{aligned} \tag{22}$$

in which the second line follows from that  $f(\cdot)$  is a monotonically decreasing function, the third line arises from the equality  $\max(\max(a, b), \max(c, d), \dots) = \max(a, b, c, d, \dots)$ . Eq (22) is a special case of

$$\begin{aligned} &\min(\text{list of reaction times for various singletons}) \\ &= f[\max(\text{list of all the non-trivial responses to these singletons})]. \end{aligned} \tag{23}$$

This equation is the extension of Eq (21) to multiple reaction times for multiple singletons, each alone in a singleton scene. It will be used to derive other race equalities.

Since  $\min(RT_C, RT_O) = f[\max(r_C^c, r_O^o, r_C^{co}, r_O^{co})]$  and  $RT_{CO} = f[\max(r_C^c, r_O^o, r_{CO}^{co})]$ , equality  $RT_{CO} \stackrel{p}{=} \min(RT_C, RT_O)$  requires  $\max(r_C^c, r_O^o, r_{CO}^{co}) \stackrel{p}{=} \max(r_C^c, r_O^o, r_C^{co}, r_O^{co})$ . This requirement can be met either when the CO neural responses are relatively negligible such that

$$\max(r_C^{co}, r_O^{co}, r_{CO}^{co}) < \max(r_C^c, r_O^o), \tag{24}$$



so as to reduce both  $\max(r_C^C, r_O^O, r_{CO}^{CO})$  and  $\max(r_C^C, r_O^O, r_C^{CO}, r_O^{CO})$  to  $\max(r_C^C, r_O^O)$ , or

$$r_{CO}^{CO} \stackrel{P}{=} \max(r_C^{CO}, r_O^{CO}). \tag{25}$$

The two conditions, Eqs (24) and (25), can both be satisfied when CO neurons are absent so that  $r_C^{CO} = r_O^{CO} = r_{CO}^{CO} = 0$ . In this paper, a prediction (such as  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$ ) is called spurious if the neural properties (such as the two conditions above) upon which it relies are either known to be violated in V1 or whose presence in V1 is uncertain. Whether the neural properties required for a spurious prediction can be satisfied may depend on individual observers, for example, sensitivities to different colors vary by a few fold between different observers with normal color vision [44] and V1 properties may vary accordingly [45].

Meanwhile, the equality  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$  is likely broken when the CO neurons are present [16]. Iso-feature suppression makes it likely that

$$\langle r_{CO}^{CO} \rangle > \langle \max(r_C^{CO}, r_O^{CO}) \rangle, \tag{26}$$

where  $\langle x \rangle$  means the ensemble average of  $x$ . If so,  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$  is likely replaced by a race inequality

$$\langle RT_{CO} \rangle < \langle \min(RT_C, RT_O) \rangle. \tag{27}$$

Hence, the V1 saliency hypothesis predicts *qualitatively* that  $RT_{CO}$  and  $\min(RT_C, RT_O)$  are likely to be statistically different, in particular it predicts that  $RT_{CO}$  is likely shorter, without predicting the *quantitative* difference between  $RT_{CO}$  and  $\min(RT_C, RT_O)$ .

Similarly, V1 also contains MO neurons that are tuned simultaneously to orientation and motion direction [34]. Hence,  $RT_{MO} \stackrel{P}{=} \min(RT_M, RT_O)$  is likely broken and the following inequality

$$\langle RT_{MO} \rangle < \langle \min(RT_M, RT_O) \rangle, \tag{28}$$

analogous to  $\langle RT_{CO} \rangle < \langle \min(RT_C, RT_O) \rangle$ , is likely. However, V1 is reported to contain few CM neurons that are tuned simultaneously to color and motion direction [46], although conflicting reports [46, 47, 48] make it unclear whether CM neurons are indeed absent or just fewer.

Hence, it is unclear whether  $RT_{CM} \stackrel{P}{=} \min(RT_C, RT_M)$  holds or whether the inequality  $\langle RT_{CM} \rangle < \langle \min(RT_C, RT_M) \rangle$  may occur.

Although V1 has CO and MO cells, we do not know enough about their properties. Hence, our educated guesses such as  $\langle r_{CO}^{CO} \rangle > \langle \max(r_C^{CO}, r_O^{CO}) \rangle$  and the breaking of

$RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$  are merely predicted as likely rather than certain. For observer SA in Fig 5, the behaviorally observed  $\langle RT_{CO} \rangle$  and  $\langle RT_{MO} \rangle$  are indeed shorter than their respective race model predicted values  $\langle \min(RT_C, RT_O) \rangle$  and  $\langle \min(RT_M, RT_O) \rangle$ , respectively. Meanwhile,  $RT_{CM} \stackrel{P}{=} \min(RT_C, RT_M)$  holds for this observer within the resolution provided by our data.

The inequality  $\langle RT_{\alpha\alpha'} \rangle < \langle \min(RT_{\alpha}, RT_{\alpha'}) \rangle$  for  $\alpha$  or  $\alpha' = C, M, \text{ or } O$  and  $\alpha \neq \alpha'$  is called a double-feature advantage or redundancy gain, and has been observed previously. Focusing on the time bins for the shortest reaction times, Krummenacher et al [49] showed that the densities of  $RT_{CO}$  in these bins were more than the summations of the densities of the racers  $RT_C$  and  $RT_O$ . Koene and Zhaoping [29] showed that  $\langle RT_{CO} \rangle < \langle \min(RT_C, RT_O) \rangle$  and  $\langle RT_{MO} \rangle < \langle \min(RT_M, RT_O) \rangle$  hold statistically across eight observers, whereas the average  $\langle RT_{CM} \rangle$  is not significant different from  $\langle \min(RT_C, RT_M) \rangle$ . The current work extends the previous findings by comparing the whole distribution of the observed  $RT_{\alpha\alpha'}$  with that of  $\min(RT_{\alpha}, RT_{\alpha'})$ . The difference between  $RT_{\alpha\alpha'}$  and  $\min(RT_{\alpha}, RT_{\alpha'})$  should reflect the contribution of the double-

feature tuned neurons, CO, MO, or CM, to the saliency of the double-feature singleton (via its response  $r_{CO}^{CO}$ ,  $r_{MO}^{MO}$ , or  $r_{CM}^{CM}$ , respectively, beyond the contribution of these neurons to the saliency of the single-feature singletons), as evaluated by Zhaoping and Zhe [50].

Generalizing our derivations (in Eqs (14)–(27)), the triple-feature race model

$RT_{CMO} \stackrel{P}{=} \min(RT_C, RT_M, RT_O)$  is likely broken when the responses from the CM, CO, and MO neurons are not negligible unless, analogous to Eq (25), the response equality  $\max(r_{CM}^{CM}, r_{CO}^{CO}, r_{MO}^{MO}) \stackrel{P}{=} \max(r_C^{CM}, r_M^{CM}, r_C^{CO}, r_O^{CO}, r_M^{MO}, r_O^{MO})$  holds. Here,  $r_{\alpha}^{CM}$  and  $r_{\alpha}^{MO}$  are responses of the CM and MO neurons, respectively to single- or double-feature singleton  $\alpha$ , and V1 is assumed to have no CMO cells tuned simultaneously in all the three feature dimensions. Additionally, just as  $\langle RT_{CO} \rangle < \langle \min(RT_C, RT_O) \rangle$  can arise from  $\langle r_{CO}^{CO} \rangle > \langle \max(r_C^{CO}, r_O^{CO}) \rangle$ , the inequality  $\langle RT_{CMO} \rangle < \langle \min(RT_C, RT_M, RT_O) \rangle$  can arise from

$$\langle \max(r_{CM}^{CM}, r_{CO}^{CO}, r_{MO}^{MO}) \rangle > \langle \max(r_C^{CM}, r_M^{CM}, r_C^{CO}, r_O^{CO}, r_M^{MO}, r_O^{MO}) \rangle, \quad (29)$$

which can occur when the double-feature tuned neurons respond more vigorously to the double- or triple-feature singletons than to the single-feature singletons due to iso-feature suppression.

The above inequality is a composite of the three component inequalities  $\langle r_{CO}^{CO} \rangle > \langle \max(r_C^{CO}, r_O^{CO}) \rangle$ ,  $\langle r_{MO}^{MO} \rangle > \langle \max(r_M^{MO}, r_O^{MO}) \rangle$ , and  $\langle r_{CM}^{CM} \rangle > \langle \max(r_C^{CM}, r_M^{CM}) \rangle$ . Hence, it is likely to hold when two out of the three component inequalities hold. According to analysis around Eqs (25)–(27),  $\langle r_{\alpha\alpha'}^{\alpha\alpha'} \rangle > \langle \max(r_{\alpha}^{\alpha\alpha'}, r_{\alpha'}^{\alpha\alpha'}) \rangle$  is implied by race inequality  $\langle RT_{\alpha\alpha'} \rangle < \langle \min(RT_{\alpha}, RT_{\alpha'}) \rangle$  for  $\alpha\alpha' = CO, MO, \text{ or } CM$ . Therefore, the triple-racer inequality  $\langle RT_{CMO} \rangle < \langle \min(RT_C, RT_M, RT_O) \rangle$  is quite likely when two out of the three double-racer inequalities  $\langle RT_{\alpha\alpha'} \rangle < \langle \min(RT_{\alpha}, RT_{\alpha'}) \rangle$  hold. This is the case in Fig 5. Meanwhile, the composite equality  $\max(r_{CM}^{CM}, r_{CO}^{CO}, r_{MO}^{MO}) \stackrel{P}{=} \max(r_C^{CM}, r_M^{CM}, r_C^{CO}, r_O^{CO}, r_M^{MO}, r_O^{MO})$  may still hold when  $r_{\alpha\alpha'}^{\alpha\alpha'} \stackrel{P}{=} \max(r_{\alpha}^{\alpha\alpha'}, r_{\alpha'}^{\alpha\alpha'})$  is broken for each component  $\alpha\alpha' = CO, MO, \text{ and } CM$ .

### A quantitative prediction of the reaction time for a triple-feature singleton from another race equality

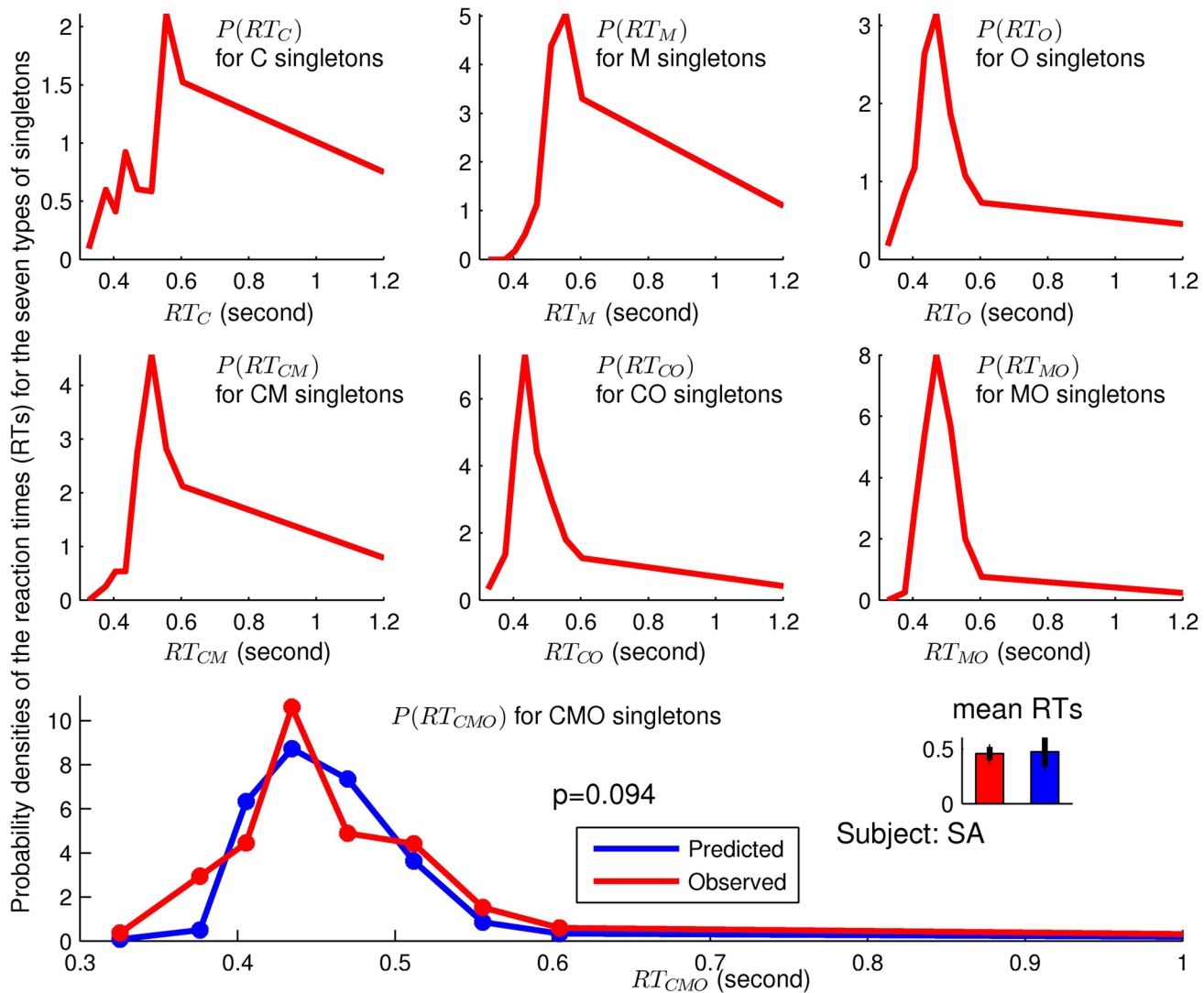
To make a quantitative prediction, we can confidently assume that V1 has no CMO neurons tuned simultaneously to all the three features, C, M, and O, given the existing paucity of neurons tuned simultaneously to C and M [46] (since a CMO neuron should at least be tuned to C and M simultaneously). Just as the absence of CO neurons gives  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$ , the absence of the CMO neurons gives (see proof in the Method section)

$$\min(RT_{CMO}, RT_C, RT_M, RT_O) \stackrel{P}{=} \min(RT_{CM}, RT_{CO}, RT_{MO}). \quad (30)$$

The left side above is the race outcome from four racers with their respective reaction times as  $RT_{CMO}, RT_C, RT_M,$  and  $RT_O$ , and the right side is the race outcome of another three racers with their respective reaction times. Since we are quite confident about the condition (that V1 lacks CMO cells) behind this equality, we call this a non-spurious race equality. It can quantitatively predict the distribution of  $RT_{CMO}$  from the distributions of the other six types of reaction times in the equality. Both the equality and its predicted  $RT_{CMO}$  distribution are also called non-spurious predictions.

Our derivation made clear that this equality does not depend on the details of the contextual influences in V1 other than its most prominent and essential aspects: iso-feature suppression that makes a feature singleton the most salient in our singleton scenes. Although important details such as colinear facilitation do play a role when asking other questions on saliency, as

Observed reaction times (RTs) for various singletons and the predicted  $RT_{CMO}$  by the non-spurious race equality



**Fig 6.** The observed distributions of  $RT_C$ ,  $RT_M$ ,  $RT_O$ ,  $RT_{CM}$ ,  $RT_{CO}$ , and  $RT_{MO}$  for an observer are used to predict the distribution of  $RT_{CMO}$  for the same observer (SA who was also in Figs 3 and 5) by the non-spurious race equality  $\min(RT_{CMO}, RT_C, RT_M, RT_O) \stackrel{p}{=} \min(RT_{CM}, RT_{CO}, RT_{MO})$ . The predicted and observed distributions of  $RT_{CMO}$  are statistically indistinguishable from each other ( $p = 0.094$ ). This figure has the same format as Fig 3.

doi:10.1371/journal.pcbi.1004375.g006

have been shown in model simulations and behavioral data [7, 30], the freedom of our non-spurious equality from such details makes our quantitative prediction possible. This is especially so since we do not yet have accurate information on these details [12, 13, 14, 17, 18, 19, 20, 22, 23] which may also depend on the observers (e.g., on their visual experience and adaptation states).

### The non-spurious prediction agrees with experimental data

Fig 6 shows that the observed distribution of  $RT_{CMO}$  for our example observer SA is statistically indistinguishable from the non-spurious prediction using the other types of reaction times of

this observer. Fig 7 shows that this agreement between the predicted and the observed  $RT_{CMO}$  holds for all six naive adult observers.

Is our non-spurious equality harder to falsify because it has a more complex structure than our spurious race models  $RT_{xx'} \stackrel{P}{=} \min(RT_x, RT_{x'})$  and  $RT_{CMO} \stackrel{P}{=} \min(RT_C, RT_M, RT_O)$ ? To answer this question, we create three new spurious equalities that are as complex as our non-spurious equality but can be falsified by the same data. Listing our non-spurious equality with these three newly created spurious equalities together,

$$\text{non-spurious: } \min(RT_{CMO}, RT_C, RT_M, RT_O) \stackrel{P}{=} \min(RT_{CM}, RT_{CO}, RT_{MO}), \quad (31)$$

$$\text{spurious : } \min(RT_{CMO}, RT_M, RT_{CO}) \stackrel{P}{=} \min(RT_C, RT_O, RT_{CM}, RT_{MO}), \quad (32)$$

$$\text{spurious : } \min(RT_{CMO}, RT_C, RT_{MO}) \stackrel{P}{=} \min(RT_M, RT_O, RT_{CM}, RT_{CO}), \quad (33)$$

$$\text{spurious : } \min(RT_{CMO}, RT_O, RT_{CM}) \stackrel{P}{=} \min(RT_C, RT_M, RT_{CO}, RT_{MO}), \quad (34)$$

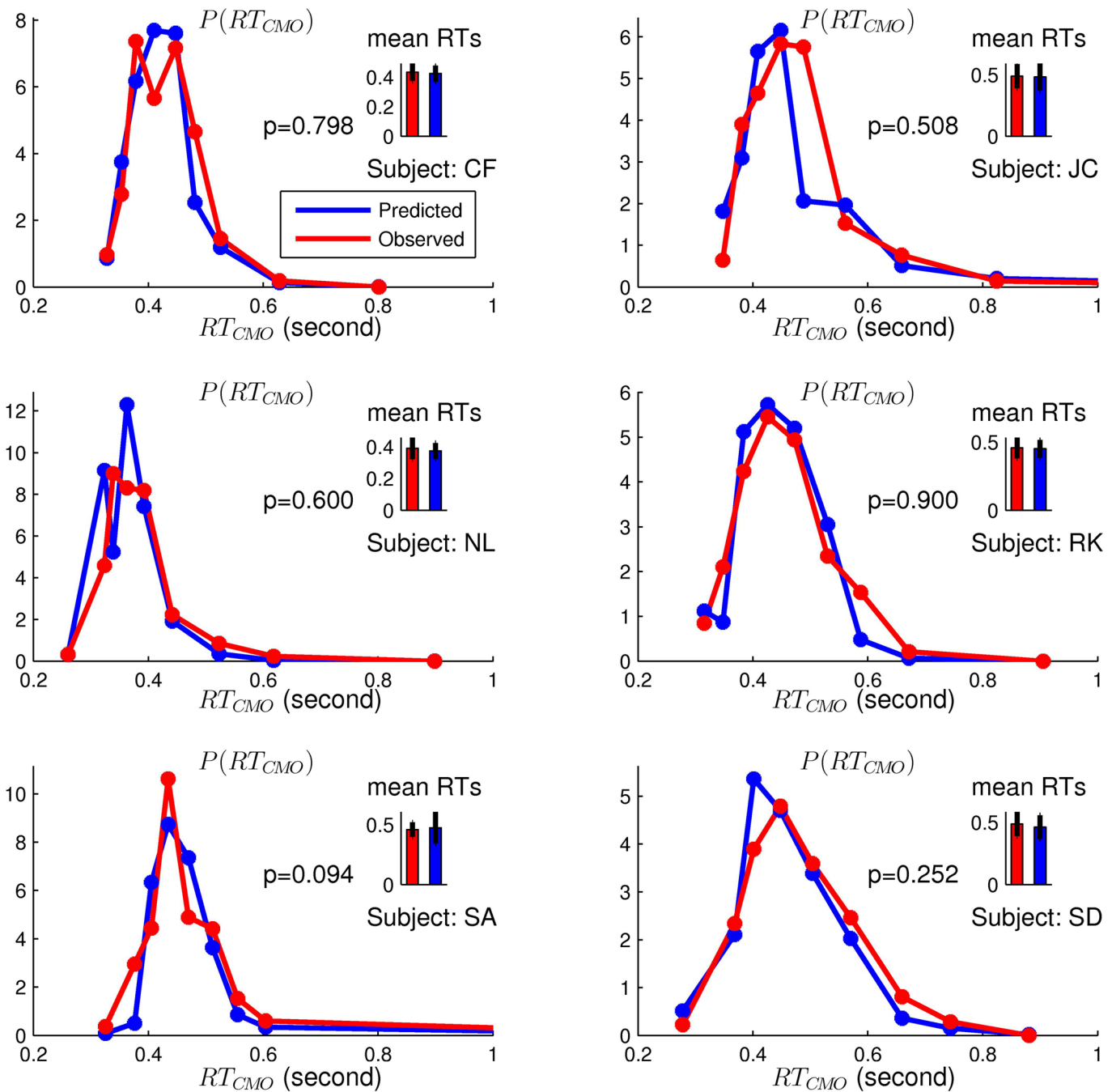
we examine their similarities and relationships. For example, the left side of Eq (31) and that of Eq (32) are identical to each other if  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$  holds, so are the right sides of the equations. Hence, Eq (32) is spurious when  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$  is spurious, unless  $RT_C$ ,  $RT_O$ , and  $RT_{CO}$  do not matter for the outcomes of their respective races (by being losers in the races),  $\min(RT_{CMO}, RT_C, RT_M, RT_O)$  and  $\min(RT_{CM}, RT_{CO}, RT_{MO})$ , in the non-spurious equality. Similarly, the Eq (33) or Eq (34) is spurious when  $RT_{MO} \stackrel{P}{=} \min(RT_M, RT_O)$  or  $RT_{CM} \stackrel{P}{=} \min(RT_C, RT_M)$ , respectively, is spurious, unless the corresponding racers are likely losers in the two races of the non-spurious equality. In other words, each of the three spurious equalities above is a corollary of a corresponding spurious (double-feature) race model  $RT_{xx'} \stackrel{P}{=} \min(RT_x, RT_{x'})$ , which we refer to as the original spurious equality. Violation of the original spurious equality is necessary but not sufficient to violate its corollary equality (subject to random fluctuations in data samples).

Each of Eqs (31)–(34) (one non-spurious) can predict the distribution of  $RT_{CMO}$  using the same set of six types of reaction times  $RT_\alpha$  for  $\alpha = C, M, O, CM, CO, MO$ . Fig 8B 8C 8D show that, in our example observer SA, the first two but not the last one of the spurious, corollary, equalities are falsified, mirroring the falsification of the original spurious equalities in Fig 5A 5B 5C. Hence, complexity in a race equality is insufficient to prevent its falsification.

## Qualitative conclusions across variations in the methods of data analysis

So far, we only illustrated the tests of the spurious equalities using data from one observer, and all the tests have so far been illustrated using a particular set of parameters characterizing the technical details in our procedures (see Methods) to test the race equalities. These technical details do not affect the qualitative conclusions. They can be parameterized by: (1) the number  $N$  of time bins to discretize the reaction time data samples for each singleton type of each observer, (2) the way to determine the boundaries between the time bins given  $N$ , (3) the metric to measure the distance  $D$  between the predicted and the observed distributions of the reaction times to judge whether a race equality holds, and (4) (only applicable to the four more complex equalities in Eqs (31)–(34)), the objective metric, i.e., the distance between the distributions on the two sides of a race equality, to be minimized in the optimization procedure to predict the  $RT_{CMO}$  distribution. The results presented so far in various figures are obtained using this set of parameters: (1)  $N = 9$  (from one of five choices  $N = 8, 9, 10, 11, 12$ ), (2) reaction time bins are

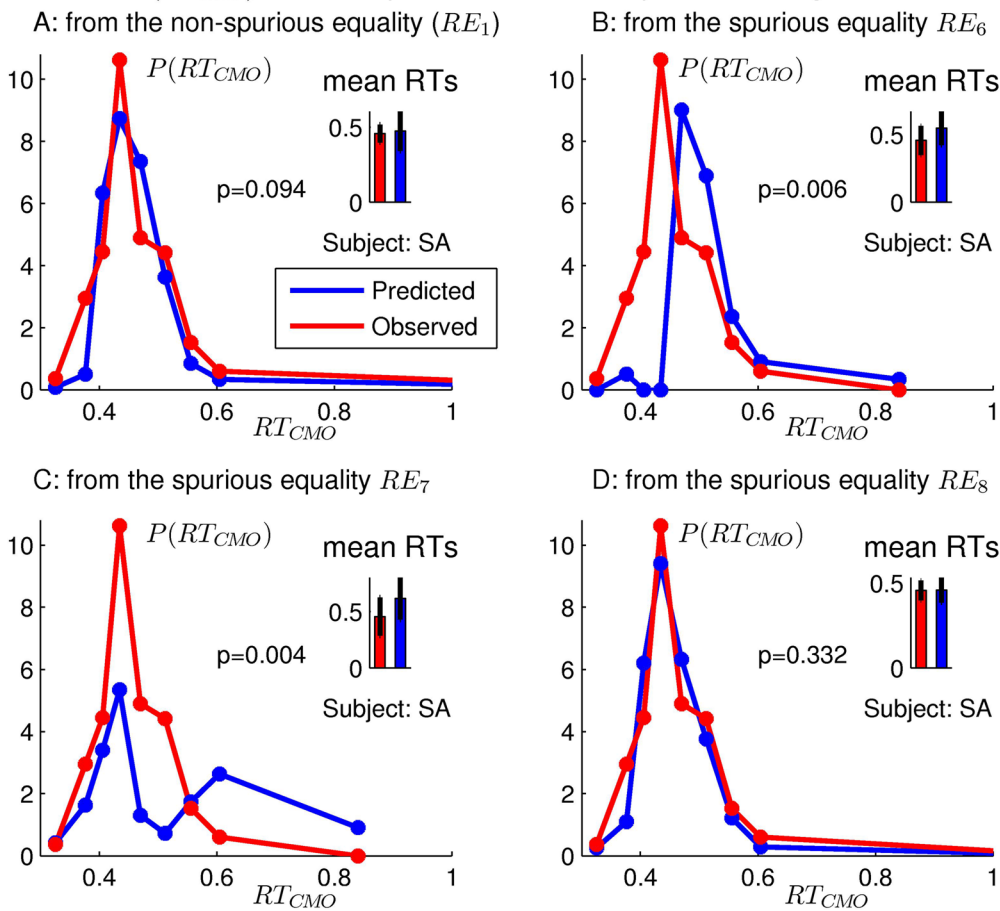
Observed  $P(RT_{CMO})$  and those predicted from the non-spurious race equality for six observers



**Fig 7. Observed and predicted distributions of  $RT_{CMO}$  using the non-spurious race equality for six observers, including observer SA whose details are shown in Fig 6. The predictions agree with data (indicated by  $p > 0.05$ ) for all observers.**

doi:10.1371/journal.pcbi.1004375.g007

Observed  $P(RT_{CMO})$  and those predicted from four equalities sharing a similar structure



**Fig 8. The predicted and observed  $P(RT_{CMO})$  from the non-spurious equality and the three spurious ones, listed in Eqs (31)–(34), are plotted in A, B, C, and D, respectively.** These four equalities share a similar complexity and are also denoted as  $RE_1$ ,  $RE_6$ ,  $RE_7$ , and  $RE_8$ , respectively, in Table 1.

doi:10.1371/journal.pcbi.1004375.g008

chosen using Eq (45) with  $x = 1.35$  (from four different choices listed around Eq (45)), (3) the  $D$  metric and (4) the objective metric are both the KL-like distance (the fourth of the four metric choices, see Eq (43)). This section presents some general statistics of our findings across  $5 \times 4 \times 4 = 80$  (or  $5 \times 4 \times 4 \times 4 = 320$  for the more complex equalities) different sets of the parameters for the method.

Table 1 lists all the (spurious or non-spurious) race equalities, each in the format of  $RT1 \stackrel{P}{=} RT2$  with definitions of  $RT1$  and  $RT2$ . For example, the equality  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$  has  $RT1 \equiv RT_{CO}$  and  $RT2 \equiv \min(RT_C, RT_O)$ . Each race equality (RE) is indexed and referred to as  $RE_1, RE_2, \dots$  or  $RE_8$ , for convenience. The  $RE_1$  is our (only) non-spurious equality  $\min(RT_{CMO}, RT_C, RT_M, RT_O) \stackrel{P}{=} \min(RT_{CM}, RT_{CO}, RT_{MO})$ . The  $RE_i$  for  $i = 2-4$  are the double-racer models  $RT_{xx'} \stackrel{P}{=} \min(RT_x, RT_{x'})$  and the  $RE_i$  for  $i = 6-8$  are their respective corollary (complex) equalities. The  $RE_5$  is the triple-racer model  $RT_{CMO} \stackrel{P}{=} \min(RT_C, RT_M, RT_O)$ . In each equality, the reaction time for the singleton with the largest number of unique features is designated (and denoted as  $RT_{goal}$  in Table 1) as the one whose distribution is predicted from those of the other reaction times.  $RT_{CMO}$  is the  $RT_{goal}$  for all race equalities except  $RE_i$  with  $i = 2-4$ , whose



**Table 1. race equalities  $RT_1 \stackrel{p}{=} RT_2$  considered in this paper.**

Equality Type/label	$RT_1$	$RT_2$	$RT_{goal}$ designated for prediction
Non-spurious			
$RE_1$	$\min(RT_{CMO}, RT_C, RT_M, RT_O)$	$\min(RT_{CM}, RT_{CO}, RT_{MO})$	$RT_{CMO}$
Spurious			
$RE_2$	$RT_{CO}$	$\min(RT_C, RT_O)$	$RT_{CO}$
$RE_3$	$RT_{MO}$	$\min(RT_M, RT_O)$	$RT_{MO}$
$RE_4$	$RT_{CM}$	$\min(RT_C, RT_M)$	$RT_{CM}$
$RE_5$	$RT_{CMO}$	$\min(RT_C, RT_M, RT_O)$	$RT_{CMO}$
$RE_6$	$\min(RT_{CMO}, RT_M, RT_{CO})$	$\min(RT_C, RT_O, RT_{CM}, RT_{MO})$	$RT_{CMO}$
$RE_7$	$\min(RT_{CMO}, RT_C, RT_{MO})$	$\min(RT_M, RT_O, RT_{CM}, RT_{CO})$	$RT_{CMO}$
$RE_8$	$\min(RT_{CMO}, RT_O, RT_{CM})$	$\min(RT_C, RT_M, RT_{CO}, RT_{MO})$	$RT_{CMO}$

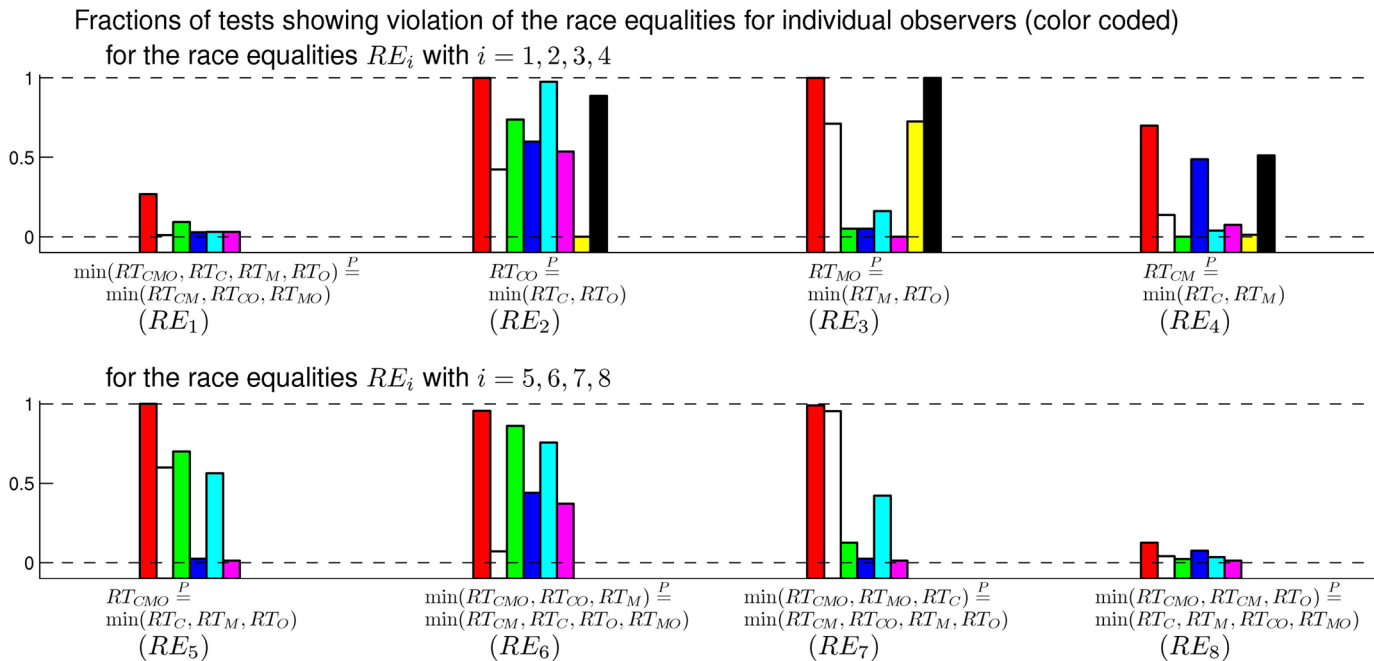
doi:10.1371/journal.pcbi.1004375.t001

$RT_{goal}$  are  $RT_{CO}$ ,  $RT_{MO}$ , and  $RT_{CM}$ , respectively.  $RT_{goal}$  tends to be the shortest reaction time in each equality, thus is more precisely determined, by the nature of the race(s), from the other reaction times.

Koene and Zhaoping [29] collected reaction times for all the single- and double-feature singletons from eight observers, but collected  $RT_{CMO}$  data from only six of these observers. Hence,  $RE_i$  with  $i = 2-4$  can be tested by eight observers while the other equalities by only six observers.

Whether a race equality can be falsified by data from a particular observer depends on several factors. First, as mentioned before, it may depend on the observer, as there may be inter-observer difference in terms of the V1 properties and visual sensitivities [44, 45]. Second, even when a race equality is truly false for a particular observer, it may appear to hold when there are insufficient samples of reaction time data, and thus insufficient statistical power in the data, to reveal a difference (particularly a small difference) between the prediction and its behavioral counterpart. Conversely, even when a race equality is fundamentally true, there is a 5% chance to find it accidentally broken by behavioral data. This is because, by definition (see Methods), a null hypothesis proclaiming the race equality is declared false when the distance  $D$  between the predicted (by the race equality) and observed distributions of reaction times is larger than 95% of the random samples of the distances  $D$  when the null hypothesis strictly holds. Third, empirically, the technical parameters (particularly the metric used to measure the difference between the predicted and observed distributions of reaction times) in our procedure can sometimes affect whether a race equality is falsified by data.

Fig 9 plots the fraction of all the (80 or 320) tests in which an equality is found broken in each observer and each race equality. In more than half of the cases, this fraction is either larger than 90% or smaller than 10%, indicating that the variations in the parameters of our method do not substantially affect whether the race equality holds. For some observers in some race equalities, e.g., observers marked by white, blue, and magenta color for  $RE_2$ , a race equality is consistently broken using one metric and consistently maintained using another metric, (almost) regardless of the variations of the other parameters for the tests. For our non-spurious race equality, no test parameter value of any type consistently break the equality in any observer regardless of the other parameters.



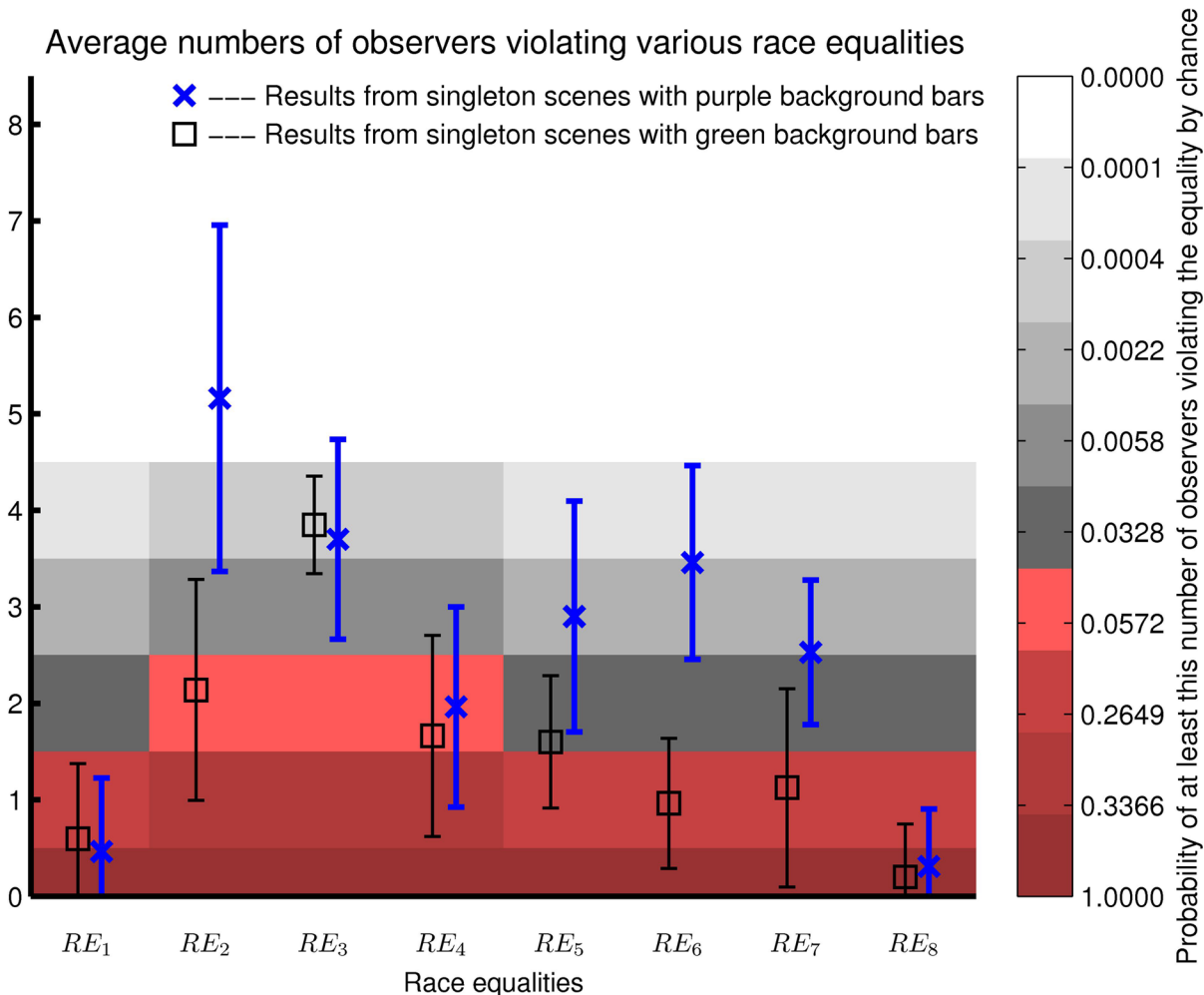
**Fig 9. The fraction of the tests of each race equality that falsify the equality for each observer.** Each observer is color coded by: red, white, green, blue, cyan, magenta, yellow, or black (red for our example observer SA). Different tests of an equality use different sets of parameters in the testing method to include all possible combinations of the parameter values. Each race equality is tested on six or eight observers as indicated. Results for  $RE_i$  for  $i = 2-4$  are placed above that of its corollary equality  $RE_{i+4}$  for easy of comparison.

doi:10.1371/journal.pcbi.1004375.g009

Individual differences in neural response properties and a lack of statistical power in data are likely to partly explain why even the most obviously spurious equality  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$  is not broken by data from all observers. For example, the observer coded by yellow color in Fig 9 appears to show race equality  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$ ; this may either be caused by a lack of vigorously responding CO cells in this observer, or it may be because the difference between  $RT_{CO}$  and  $\min(RT_C, RT_O)$  is too small to be detected by the limited number of random samples of each type of reaction times  $RT_{CO}$ ,  $RT_C$ , and  $RT_O$ .

Given a 5% chance to break a true race equality accidentally, there is a chance of  $\frac{N!}{n!(N-n)!} 0.05^n 0.95^{N-n}$  that  $n$  out of  $N$  observers will break a true equality accidentally. Hence, if more than one or two out of six or eight observers, respectively, break a race equality, we say that the equality is broken or incorrect since such a high tendency of equality breaking can happen only by a chance of less than 0.05 for a truly correct race equality.

Fig 10 plots the number of our observers to break each race equality, averaged over all the tests (each applied to all individual observers) which differ by the parameters in the testing method. Data points on gray or white background are those with more observers breaking an equality than expected by a probability of 0.05 if the equality truly holds. Blue crosses or black squares are, respectively, results from using  $RT_\alpha$  data collected from purple or green scenes, respectively. Our results in Figs 3-9 are all based on data from the purple scenes. Focusing first on blue crosses (from purple scenes) in Fig 10, we have the following qualitative conclusions which are relatively immune to the sensitivities to the details in the testing method. First, the non-spurious race equality ( $RE_1$ ) is confirmed since it is only broken by an average of 0.5



**Fig 10. Average numbers of observers to break various race equalities, as shown in blue or black data points whose error bars denote standard deviations.** The non-spurious race equality is  $RE_1$ . Data from 6 observers were tested for race equalities  $RE_1$  and  $RE_i$  for  $i \geq 5$  and data from 8 observers were tested for  $RE_2$ ,  $RE_3$ , and  $RE_4$ . Applying a test of a given race equality to all the observers gives a number of observers breaking this equality, and the average of this number over 80 (for  $RE_i$  with  $i = 2-5$ ) or 320 (for  $RE_1$  and  $RE_i$  with  $i > 5$ ) tests, each characterized by a unique set of parameters in the testing method, gives a data point (blue cross or black square). The background shadings visualize the probabilities of at least a certain number of observers breaking a true race equality accidentally, shadings in red hue indicate probabilities larger than 0.05. Note that the number of observers in this probability representation is an integer number, whereas the data points are generally non-integers since they are averages of integer numbers.

doi:10.1371/journal.pcbi.1004375.g010

observers, within the range expected for chance breaking of a true equality. Second, two spurious predictions,  $RE_2$  and  $RE_3$  (for equalities  $RT_{CO} \stackrel{p}{=} \min(RT_C, RT_O)$  and  $RT_{MO} \stackrel{p}{=} \min(RT_M, RT_O)$ , respectively), are broken since data from more than about 3.5 observers break each of them, consistent with the presence of CO and MO neurons in V1 [34, 16]. Third, the spurious  $RE_4$  for equality  $RT_{CM} \stackrel{p}{=} \min(RT_C, RT_M)$  is barely broken, or not as seriously broken as  $RE_2$  and  $RE_3$ , since only around 2 out of eight observers have data violating it. This is consistent with the idea that V1 has fewer CM than CO or MO neurons, and is consistent with the controversy in experimental reports [46, 47, 48] regarding whether CM cells exist in V1. Fourth, the spurious prediction  $RE_5$  for equality  $RT_{CMO} \stackrel{p}{=} \min(RT_C, RT_M, RT_O)$  is broken since around three out of six observers violate it. This is consistent with the fact that V1

contains a substantial number of conjunctively tuned cells, in particular the CO and MO cells, and corroborates the finding that its component race equalities  $RE_2$  and  $RE_3$  are clearly broken. Fifth, the complex spurious equalities  $RE_i$  for  $i = 6-8$ , each a variation of the non-spurious  $RE_1$  and can be potentially undermined (when certain conditions hold, as discussed in the text around Eqs (31)–(34)) by the violation of the corresponding original  $RE_{i-4}$ , are broken for  $RE_6$  and  $RE_7$  but maintained for  $RE_8$ . This corroborates our findings for the original spurious  $RE_{2-4}$ . The corollary equalities are less seriously broken than their original counterparts, lending further support to our non-spurious  $RE_1$  as it sustains the corollary equalities against the undermining factors from the violated original equalities.

However, more spurious predictions survive the test by data from the green scenes, see data points in black squares in Fig 10. In particular, the spurious  $RT_{CO} \stackrel{P}{=} \max(RT_C, RT_O)$  is only marginally broken. Reaction times for singletons unique in color,  $RT_C$ ,  $RT_{CO}$ ,  $RT_{CM}$ , and  $RT_{CMO}$ , tend to be smaller in the green than purple scenes, particularly  $RT_C$  which is about 200–300 ms shorter in the green scenes. When both  $RT_{CO}$  and  $\max(RT_C, RT_O)$  are closer to the minimum possible manual reaction time (around 0.3 second) of each observer, their difference also becomes smaller and is thus more difficult to be detected by the limited statistical power in our data. If we do a gross approximation by ignoring the difference between the green and purple scenes so as to increase the statistical power by pooling data across the two kinds of scenes, the outcomes are qualitatively the same as using data from the purple scenes alone except for  $RE_6$  which is marginally (rather than clearly) broken when data are pooled. Importantly, our non-spurious prediction  $RE_1$  agrees with data regardless of whether data come from the green or purple scenes.

The finding that the spurious equality  $RE_8$  agrees with our data is not a problem for the V1 saliency hypothesis. Recall that a prediction (or race equality) is called spurious in this paper if the neural properties upon which it relies are either uncertain or known to be violated in V1. If we were certain that V1 has no CM cells, then  $RE_8$  would be non-spurious since its original equality  $RE_4$  would be non-spurious. Hence, a marginally broken  $RE_4$  makes  $RE_8$  less likely broken, and the lack of serious violation of both  $RE_4$  and  $RE_8$  is consistent with the controversy regarding whether V1 has CM cells. If we were certain that V1 does have substantially responsive CM cells (such that  $\langle r_{CM}^{CM} \rangle > \langle \max(r_C^{CM}, r_M^{CM}) \rangle$  and  $\langle r_{CM}^{CM} \rangle > \langle \max(r_C^C, r_M^M) \rangle$ ) while  $RE_4$  is not substantially violated (given sufficient statistical power in data), then V1 saliency hypothesis would be falsified.

Our non-spurious  $RE_1$  and the spurious  $RE_i$  for  $i = 6-8$  have very similar structures, they use the same technical procedure to predict  $RT_{CMO}$  from the same set of reaction times to the other singleton types. Hence, violations of equalities  $RE_6$  and  $RE_7$  suggest that our data have a sufficient statistical power in the purple scenes to reject our non-spurious equality  $RE_1$  if it were just as clearly incorrect as  $RE_6$  and  $RE_7$ . Therefore, our non-spurious V1 prediction is confirmed within the resolution provided by the statistical power of our data. This resolution is manifested in Fig 8 in which it can clearly distinguish between the two reaction time distributions depicted in red and blue curves in Fig 8B or Fig 8C but not in Fig 8A or Fig 8D.

## Discussion

### The main finding

Our non-spurious prediction,  $\min(RT_{CMO}, RT_C, RT_M, RT_O) \stackrel{P}{=} \min(RT_{CM}, RT_{CO}, RT_{MO})$ , agrees with behavioral data such that the distribution of  $RT_{CMO}$  can be quantitatively predicted from those of the other types of reaction times of the same observer without any free parameters. This prediction is derived using the following essential ingredients: (1) the V1 saliency

hypothesis that the highest V1 neural response to a location relative to the highest V1 responses to the other locations signals this location's saliency, (2) the feature-tuned neural interaction, in particular iso-feature suppression, that depends on the preferred features of the interacting neurons to cause higher responses to feature singletons, (3) the data-inspired assumption that V1 does not have CMO neurons tuned simultaneously to color, motion direction, and orientation, and (4) the monotonic link (within the definition of saliency) between a higher saliency of a location and a shorter saliency-dictated reaction time to find a target at this location. Hence, our finding supports the direct functional link between saliency of a visual location and the maximum (rather than, e.g., a summation) of V1 neural responses to this location, as prescribed by the V1 saliency hypothesis. It also suggests that saliency computation (at least for our singleton scenes) essentially employs only the mechanisms with the following two properties: feature-tuned interaction between neighboring neurons (in particular iso-feature suppression) and a lack of CMO neurons, both available in V1, and neural mechanisms which are absent in V1 are not needed.

### The supporting findings

In addition, the following qualitative findings are obtained. First, two spurious predictions,  $RT_{CO} \stackrel{P}{=} \min(RT_O, RT_O)$  and  $RT_{MO} \stackrel{P}{=} \min(RT_M, RT_O)$ , about which we have good confidence to be incorrect based on the V1 saliency hypothesis and the known presence of the CO and MO cells in V1, are falsified by our reaction time data. Second, using the V1 saliency hypothesis and our knowledge about the V1 neural substrates, we predicted relationships between the three predictions just mentioned, one non-spurious and two spurious, and the other five spurious predictions listed in [Table 1](#). These relationships include the relative degrees of spuriousness between predictions and the dependence of some predictions on the non-spuriousness of some other predictions and certain properties of behavioral reaction times. The outcomes of testing the other five predictions are consistent with the predicted relationships, lending further support to the V1 saliency hypothesis.

### Implications for the V1 saliency hypothesis

Previously, the V1 saliency hypothesis provided only qualitative predictions. One example [\[21\]](#) predicts that an ocular singleton is salient and hence that the reaction time to find a visual search target is shorter when this target is also an ocular singleton, but it cannot quantitatively predict how much shorter this reaction time should be. Another example [\[30\]](#) predicts that a very salient border between two textures of oblique bars can be made non-salient (in a way unexpected from traditional saliency models) by superposing the textures with a checkerboard pattern of horizontal and vertical bars, but it cannot predict the quantitative increase in reaction times to locate the texture border by the superposing texture. Although these qualitative predictions are confirmed [\[21, 30\]](#), we cannot consequently conclude whether, in addition to the V1 mechanisms, more complex mechanisms available only in higher brain centers might also contribute to saliency computation. In contrast, if a prediction *quantitatively* specifies that one reaction time should be, say, 20% shorter than another one, and if data reveal instead that the first reaction time is only 10% shorter, then additional mechanisms for saliency computation must be called for. The quantitative agreement between our non-spurious prediction and the reaction time data without any free parameters suggests that saliency computation requires essentially no other neural mechanisms than those with the feature-tuned interactions between neurons and a lack of CMO neurons—both are V1 properties.

Let us articulate some other mechanistic ingredients or assumptions that were omitted in our closing sentence in the last paragraph and have been explicit or implicit in this paper. We

assumed that the fluctuations in the responses of different types of neurons to an input item (e.g., responses of the C, O, and CO neurons to a CO singleton) are independent of each other. Also, fluctuations of the responses to different input items in a scene are assumed to be sufficiently independent of each other, so that we can treat the statistical properties of the responses to the background bars as independent of the responses to the singleton. We also assumed that the response of a neuron to a singleton is independent of whether this singleton is unique in a feature dimension to which this neuron is not tuned. For example, we assumed no statistical difference between  $r_C^C$ ,  $r_{CO}^C$ , and  $r_{CMO}^C$ , between  $r_C^{CO}$  and  $r_{CM}^{CO}$ , or between  $r_B^C$  and  $r_O^C$ . This assumption may only be seen as an approximation given the known activity normalization in cortical responses [51]. Since V1 neurons' responses are insensitive to small differences in luminance contrast when this contrast is very high [52], we also assumed that, when a V1 cell is not tuned to color, its response to our stimulus bar, which has a 100% luminance contrast against a dark background, is independent of whether our bar is green or purple, even though isoluminance between the two colors was not finely calibrated and adjusted to suit individual observers [29]. This assumption was needed to assume no statistical difference (e.g.,) between  $r_B^O$  and  $r_C^O$  and between  $r_O^O$  and  $r_{CO}^O$ . The statistical properties of the population responses to the background bars are also assumed to be regardless of the type, location, and feature values of the singleton in our singleton scenes (provided that we restrict all singleton scenes to purple scenes only or to green scenes only, see Fig 4). This assumption enabled us to write Eq (3). Meanwhile, Eq (3) led to Eq (4) by an implicit assumption that fluctuations in the saliency readout to motor responses are negligible (this might be more likely for bottom-up than top-down responses). Furthermore, observers' perceptual learning to do the visual search is assumed as negligible over the course of the data taking, so that the monotonic function relating V1 responses to reaction times is fixed. The above simplifications or idealizations were made to keep our question focused on the most essential mechanisms. That our prediction agrees quantitatively with data suggests that these simplifications or idealizations are sufficiently good approximations within the resolution that can be discerned by our data.

Future investigations could further test the V1 saliency hypothesis using more complex feature conjunctions. For example, one can test whether behavioral data on a conjunction of two orientations [30] match V1's physiological property regarding whether V1 has sufficiently active cells tuned to such a conjunction [53].

## Implications for the role of extrastriate cortices

An important question is whether extrastriate cortices, i.e., cortical areas beyond V1, might also contribute to compute saliency. We have concluded that the two essential properties of the neural mechanisms for saliency computation are (1) feature-tuned contextual influences (in particular iso-feature suppression) and (2) a lack of CMO tuned cells. If extrastriate mechanisms also possess these properties, they could contribute to computing saliency, and we could extend to them the hypothesized link between the highest neural response to a location and the saliency of this location. After all, extrastriate visual areas also project to superior colliculus and so can influence eye movements.

Extrastriate cortices have been known [12] to exhibit feature-tuned contextual influences, in particular the iso-feature suppression. For example, V4 neurons exhibit iso-color, iso-orientation, and iso-spatial-frequency suppression [54, 55], V2 neurons exhibit iso-orientation suppression [56], and MT neurons exhibit iso-motion-direction suppression [12].

However, extrastriate cortices contain CMO neurons (private communication from Stewart Shipp, 2011). For example, Burkhalter and van Essen [35] observed that, in V2 and VP, many cells were feature selective in multiple feature dimensions, including orientation, color, and



motion direction, and that the probability for a cell to be tuned in one feature dimension is independent of whether the cell is also tuned in another feature dimension. These observations imply that triple-feature tuned CMO cells are present. In fact, since they observed that most neurons are tuned to orientation and most neurons are tuned to color, the probability that a cell can be a CMO cell must be no less than 25% of the probability of this cell being tuned to direction of motion (M). Similar conclusions in V2 are reached by other investigations [57, 58], although different researchers use different criteria to classify feature tuning. In addition, unlike the case in V1 where the presence of CM neurons is controversial, V2 is known to have CM neurons in addition to CO and MO neurons [57, 48, 58]. Some of these CM, CO, and MO neurons (which are defined experimentally as being tuned to the two specified feature dimensions simultaneously without restrictions on the neuron’s selectivity in the other feature dimensions) in V2 can well be CMO neurons, especially when the chance for a neuron to be tuned to a feature dimension is independent of whether it is already tuned to any other dimensions. Selectivity to conjunctions of more than two types of features in extrastriate cortices is consistent with general observations that neurons in cortical areas beyond V1 tend to have more complex and specialized visual receptive fields.

According to our analysis in the [Methods](#) section, if a cortex containing the saliency map had CMO neurons, then, statistically,  $RT_{CMO}$  would be likely smaller than predicted by our non-spurious race equality  $\min(RT_{CMO}, RT_C, RT_M, RT_O) \stackrel{P}{=} \min(RT_{CM}, RT_{CO}, RT_{MO})$ , just as the presence of CO neurons makes  $RT_{CO}$  likely shorter than predicted by the race equality  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$ . More specifically, our non-spurious equality was proven (in the [Methods](#) section) by using [Eq \(23\)](#) to write  $\min(RT_{CMO}, RT_C, RT_M, RT_O) \stackrel{P}{=} f[\max(Y)]$  and  $\min(RT_{CM}, RT_{CO}, RT_{MO}) \stackrel{P}{=} f[\max(Y)]$ , where  $Y$  is a list of responses from the single- and double-feature tuned neurons as specified in [Eq \(39\)](#). If CMO cells exist, then by [Eq \(23\)](#) four extra items  $r_{CMO}^{CMO}, r_C^{CMO}, r_M^{CMO}$ , and  $r_O^{CMO}$  should be added to the list  $Y$  for  $\min(RT_{CMO}, RT_C, RT_M, RT_O)$  and three extra items  $r_{CM}^{CMO}, r_{CO}^{CMO}$ , and  $r_{MO}^{CMO}$  to the same list  $Y$  for  $\min(RT_{CM}, RT_{CO}, RT_{MO})$ . This upsets the equality unless either the CMO responses satisfy

$$\max(r_{CMO}^{CMO}, r_C^{CMO}, r_M^{CMO}, r_O^{CMO}) \stackrel{P}{=} \max(r_{CM}^{CMO}, r_{CO}^{CMO}, r_{MO}^{CMO}), \tag{35}$$

or if all CMO responses are negligible relative to  $\max(Y)$ , the maximum response of the list of single- and double-feature tuned neurons. Iso-feature suppression would typically make  $r_{CMO}^{CMO}$  largest among  $r_z^{CMO}$  for all  $\alpha$ , making  $\langle \max(r_{CMO}^{CMO}, r_C^{CMO}, r_M^{CMO}, r_O^{CMO}) \rangle > \langle \max(r_{CM}^{CMO}, r_{CO}^{CMO}, r_{MO}^{CMO}) \rangle$  likely so that  $RT_{CMO}$  is likely smaller than predicted unless the CMO responses are immaterial.

Assuming that the extrastriate CMO responses are not negligible and do not satisfy [Eq \(35\)](#), then the experimental confirmation of our non-spurious race equality suggests that, at least for our singleton scenes, extrastriate cortices contribute little to the guidance of exogenous attention (excluding the contribution to maintaining the state of alertness of observers). This suggestion is consistent with our previous finding [21] that an eye-of-origin singleton is very salient despite a paucity of eye-of-origin signals in every cortical area beyond V1.

Meanwhile, we do not know enough to rule out the possibility that the responses of the extrastriate CMO cells satisfy [Eq \(35\)](#) or are negligible relative to the responses from cells tuned conjunctively to fewer feature dimensions. For example, [Eq \(35\)](#) could hold if CMO responses could be invariant to any changes in the contextual inputs outside the classical receptive fields of these cells, in particular, if the extrastriate CMO responses could be exempted from the ubiquitous iso-feature suppression. The current study can hopefully motivate experimental investigations of the response properties of these extrastriate cells.

## Visual search in complex scenes, top-down factors in visual search, saliency in lower animal species, and representation of saliency in various brain regions

There remains an empirical question to ask if extrastriate cortices participate in saliency computation in more complex scenes. When top-down guidance is not held constant, one can no longer assume that reaction time (across different trials and scenes) relates monotonously with the saliency at a target's location, making it difficult to test saliency hypotheses using reaction times. In a complex street scene for example, more than one saccade is typically required to search for, e.g., a person, whereas the singletons in our study can be typically located by the first saccade, leading to a manual reaction time to report the target less than a second in typical cases. Once the gist of a scene is comprehended within the first glimpse [59], the later saccades can be highly influenced by top-down knowledge [60, 61] (e.g., to direct gaze to the pavement but not the sky for finding a person). It is known that observers with and without object agnosia have very similar initial but not later saccades in viewing pictures [62], suggesting that initial saccades are relatively free from top-down factors via object-based knowledge [63]. Therefore, to answer our empirical question, we need more suitable measures than reaction times for a target not easily found by initial saccades. Meanwhile, having no neurons tuned to complex objects or features should not by itself exclude V1 from determining saliency in complex scenes. Most objects evoke V1 responses to their low level features, e.g., segments of a face contour. Such V1 responses could attract attention to objects before objects are recognized. A neural circuit model of V1 has showed that such responses could account for the fact that it is easier to find an ellipse among circles than a circle among ellipses [64, 7] and that angry faces tend to be more salient than happy ones [65].

Top-down factors can also affect short reaction times through expectation and goals. Krummenacher and Müller [66] showed that, CM singletons evoke a reaction time clearly shorter than predicted by the race model  $RT_{CM} = \frac{P}{1-P} \max(RT_C, RT_M)$  from the V1 saliency hypothesis when assuming no CM cells in V1. However, they data taking blocks had the C, M, and CM singleton trials exclusively, the target was red and/or moving while the non-targets were always stationary and green, enabling top-down feature-based attention to red and/or moving bars. Furthermore, their search array had only  $6 \times 6$  bars and the target was always within the central  $4 \times 4$  bars, i.e., within the attentional window around the fixation at the start of a search trial (previous work [67] suggests that the attentional window size during visual search has a radius of about two in the units of average distances between neighboring search items), making it easier to exert top-down, goal-driven, target selection. (In contrast, our observers could not guess beyond chance the type, features, or location (which was always far beyond the central fixation) of the singleton in the next trial [29]). Their finding can thus be viewed as evidence supporting their idea of signal integration processes for a top-down (feature) dimension-weighting account [68, 69]. Indeed, feature-based, goal-directed, selections evoke enhanced responses in neurons in the frontal eye fields and V4 to visual inputs sharing the target's features [70]. When they are useful for the task, repeated structures and details of visual inputs over trials can also guide attention [71] to contaminate behavioral measures for bottom-up saliency [72].

In lower animals like fish or frogs without a fully developed neocortex or V1, saliency computation is perhaps done in the retina or the optic tectum which is commonly called superior colliculus in mammals. Parallels of our saliency computation in singleton scenes are seen in archer fish preying on land-based insects by shooting them down with water [73]. The fish's reaction time to attack a motion singleton, unique in speed, motion direction, or both features, is roughly independent of the number of preys in the singleton scene (but not in non-singleton

scenes). Their tectum neurons exhibit iso-feature suppression in both feature dimensions of motion speed and motion direction, and some neurons are tuned conjunctively to both feature dimensions. Furthermore, the double-feature singletons attract attention more strongly while evoking stronger responses from the conjunctive cells. Hence, the V1 saliency map in primates may evolutionarily come from the tectum. It is of interest where the saliency map might be in animals such as rodents, whose V1 inputs to superior colliculus increase response magnitudes but not input selectivities of colliculus neurons [74].

As saliency affects behavior when read out for attentional shift (often combined with top-down factors for attentional guidance), it is unsurprising that neural correlates of saliency have been found in the superior colliculus [75] and in the parietal cortex [10, 76] and frontal eye field [77, 78], which also projects to the superior colliculus and are involved in top-down attentional control. In these downstream areas from V1 in the network for attentional control, saliency representation can be viewed as a copy or transformation of the saliency map in V1. For example, the map of graded saliency values can be transformed to a map of winner-take-all discrete values in which only the saccadic destination has a non-zero value. Indeed, in a color singleton search, the neural activities in superior colliculus [79], frontal eye field [80], and lateral intraparietal cortex [81] evolved from a map of activities at input locations of the search target and the non-targets to another map with activities merely or dominantly at the saccadic target destination. In the same vein, fMRI activities in the frontal eye field can be used to decode the most salient location in the visual field [78]. However, an explicit map of saliencies is computed and created in V1. Its content can be ignored, or combined with top-down factors, in the downstream areas such that the neural activities in all the three downstream areas are strongly affected by top-down, goal-directed, factors [77, 75, 76].

### Further discussions assuming no role in saliency by the extrastriate cortices

Although the current study cannot firmly establish the possibility that extrastriate cortices play no role in saliency, the implication of this possibility deserves pondering. The control of attentional selection, including exogenous selection, is traditionally thought to rest on a network of neural circuits comprising frontal and parietal areas [82, 10, 1]. The role of subcortical areas such as the superior colliculus has also been suggested [83]. An exclusion of extrastriate contributions from exogenous control should invite a fundamental revision of this network.

If exempted from guiding exogenous attention, extrastriate areas can focus on post-selective decoding and/or endogenous selection [84]. Furthermore, in light of exogenous selection by V1, and since attentional selection admits only a tiny fraction of sensory information to be processed in detail, visual information processed in the extrastriate areas is likely to have a much smaller amount than that fed to V1 from the retina. This consideration should shape our investigations and shed light on some past observations. Indeed, unlike those in extrastriate areas, V1 activities are more associated with sensory inputs than with perception (i.e., outcomes of visual inference) and is less influenced by top-down attention [85]. For example, V4 lesions impair visual selection of only non-salient objects [86] disfavored by exogenous selection, demonstrating V4's involvement in endogenous but not exogenous selection. Equally, neural responses in V4 but not V1 to binocularly rivalrous inputs are dominated by perceived input rather than the retinal images [87], contrasting V4 with V1 in perceptual decoding. Identifying V1's role in exogenous selection thus helps to crystallize the research questions and pave the way for investigating extrastriate cortical areas.

## Methods

### Behavioral data to test various race equalities

We test various race equalities using data from Koene and Zhaoping [29]. Each of their stimuli contained  $30 \text{ rows} \times 22 \text{ columns}$  of bars (each randomly jittered from the regular grid location), extending about  $39 \times 29$  degrees of visual angle. They collected about 300 samples of reaction times for each singleton category  $\alpha = C, M, O, CM, CO, MO$ , or  $CMO$  from each observer, whose task was to press a left or right button, respectively, to report as quickly as possible whether a singleton was in the left or right half of the display, regardless of the feature(s) distinguishing the singleton. Each stimulus bar was a rectangle about  $1 \times 0.2^\circ$  in visual angle, took one of the two possible colors (green and purple), tilted from vertical in either clockwise or anticlockwise direction by a constant amount, and moved left or right at a constant speed, see Fig 4. All background bars were identical to each other in color, orientation, and motion direction; the singleton is unique in color, tilt direction, or motion direction, or any combination of these features. The green and purple colors had equal luminance ( $14 \text{ cd/m}^2$  in a black background) and equal color saturation in opposite CIE 1976 direction (hue angle  $130^\circ$  and  $310^\circ$ , respectively) from neutral white at  $u' = 0.2$  and  $v' = 0.46$ . Given an observer, all bars had the same absolute angle from vertical, the same absolute motion speed, and the same color saturation; these absolute values were chosen for the observer and stayed fixed across all trials such that  $RT_\alpha$  for each single-feature singleton  $\alpha = C, M$ , or  $O$  was around 0.6 seconds on average (averaged across green and purple scenes). Different singleton scenes, in terms of the singleton type  $\alpha$  and the color (green or purple), motion direction, and tilt direction of the background bars, were randomly interleaved. In each trial (of the data for this study), the singleton was randomly near 1 of 18 (9 left, 9 right) grid locations (in the  $30 \times 22$  grid) at an eccentricity around  $12.8^\circ$  from the display center where observers fixated at the start of the trial.

Trials with incorrect button presses or with reaction times shorter than 0.2 seconds or longer than three standard deviations above the average reaction time (for the particular observer and singleton type) were excluded from data analysis. Two out of the eight observers (four of them male) lacked data on  $RT_{CMO}$  (since they completed only an earlier version of the experiment). More details about the experiment can be found in the original paper [29], which did not publish or use the  $RT_{CMO}$  data. For each observer, data are divided into two pools, one collected from the green scenes and the other from the purple scenes; and each pool has about 150  $RT_\alpha$  data samples (on average) for each  $\alpha$ . Results in Figs 3–9 are from analyzing data from purple scenes only. Fig 10 includes results from both types of scenes.

### Proof of the non-spurious race equality in Eq (30)

First, we use Eq (21) to write each  $RT_\alpha$  in this equality as

$$RT_\alpha = f[\max(\text{list of non-trivial neuron responses to the singleton } \alpha)]. \quad (36)$$

This generalizes Eq (20) to six types of V1 neurons  $X = C, M, O, CM, CO$ , and  $MO$ , of which none tuned to  $CMO$ , and to seven types singletons  $\alpha = C, M, O, CM, CO, MO$ , and  $CMO$ . The response of neuron type  $X$  to singleton type  $\alpha$  is  $r_X^\alpha$ . For example, by Eq (4) and analogous to Eq (17),

$$RT_C = f[\max(r_C^C, r_C^M, r_C^O, r_C^{CM}, r_C^{CO}, r_C^{MO})] \stackrel{P}{=} f[\max(r_C^C, r_B^M, r_B^O, r_C^{CM}, r_C^{CO}, r_B^{MO})]. \quad (37)$$

For the above, we used  $r_C^M \stackrel{P}{=} r_B^M, r_C^O \stackrel{P}{=} r_B^O$ , and  $r_C^{MO} \stackrel{P}{=} r_B^{MO}$  which, analogous to Eqs (15)–(16), arise because a neuron's responses to a singleton and a background bar are statistically the same unless the singleton is unique in at least one the feature dimensions to which this neuron

is tuned. Then, keeping only the non-trivial responses to the C singleton, we get

$$RT_C = f[\max(r_C^C, r_C^{CM}, r_C^{CO})]. \quad (38)$$

Analogously,

$$\begin{aligned} RT_M &= f[\max(r_M^C, r_M^M, r_M^O, r_M^{CM}, r_M^{CO}, r_M^{MO})] = f[\max(r_M^M, r_M^{CM}, r_M^{MO})], \\ RT_O &= f[\max(r_O^C, r_O^M, r_O^O, r_O^{CM}, r_O^{CO}, r_O^{MO})] = f[\max(r_O^O, r_O^{CO}, r_O^{MO})]. \end{aligned}$$

Meanwhile,

$$\begin{aligned} RT_{CM} &= f[\max(r_{CM}^C, r_{CM}^M, r_{CM}^O, r_{CM}^{CM}, r_{CM}^{CO}, r_{CM}^{MO})] \\ &\stackrel{P}{=} f[\max(r_C^C, r_M^M, r_B^O, r_{CM}^{CM}, r_C^{CO}, r_M^{MO})] = f[\max(r_C^C, r_M^M, r_{CM}^{CM}, r_C^{CO}, r_M^{MO})]. \end{aligned}$$

The second line above used  $r_{CM}^C \stackrel{P}{=} r_C^C$ ,  $r_{CM}^M \stackrel{P}{=} r_M^M$ ,  $r_{CM}^O \stackrel{P}{=} r_B^O$ ,  $r_{CM}^{CM} \stackrel{P}{=} r_C^{CO}$ , and  $r_{CM}^{MO} \stackrel{P}{=} r_M^{MO}$ , again because a neuron equates a unique feature with a background feature unless the neuron is tuned in this feature dimension. Analogously,

$$\begin{aligned} RT_{CO} &= f[\max(r_{CO}^C, r_{CO}^M, r_{CO}^O, r_{CO}^{CM}, r_{CO}^{CO}, r_{CO}^{MO})] \stackrel{P}{=} f[\max(r_C^C, r_O^O, r_C^{CM}, r_{CO}^{CO}, r_O^{MO})], \quad \text{and} \\ RT_{MO} &= f[\max(r_{MO}^C, r_{MO}^M, r_{MO}^O, r_{MO}^{CM}, r_{MO}^{CO}, r_{MO}^{MO})] \stackrel{P}{=} f[\max(r_M^M, r_O^O, r_M^{CM}, r_{MO}^{CO}, r_{MO}^{MO})]. \end{aligned}$$

Similarly,

$$RT_{CMO} = f[\max(r_{CMO}^C, r_{CMO}^M, r_{CMO}^O, r_{CMO}^{CM}, r_{CMO}^{CO}, r_{CMO}^{MO})] \stackrel{P}{=} f[\max(r_C^C, r_M^M, r_O^O, r_{CM}^{CM}, r_{CO}^{CO}, r_{MO}^{MO})].$$

Non-trivial responses to each singleton are listed under the scene schematics in Fig 4.

Using six types of V1 neurons (C, M, O, CM, CO, MO) instead of three types of V1 neurons (C, O, CO), one can generalize the derivations in Eqs (17)–(22) to verify that the race equality  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$  still does not hold in general.

Now, we apply Eq (23) to the left-hand side of our non-spurious equality,

$$\begin{aligned} \min(RT_{CMO}, RT_C, RT_M, RT_O) \\ \stackrel{P}{=} f[\max(r_C^C, r_C^C, r_O^O, r_O^O, r_M^M, r_M^M, r_C^{CM}, r_M^{CM}, r_{CM}^{CM}, r_C^{CO}, r_O^{CO}, r_{CO}^{CO}, r_M^{MO}, r_O^{MO}, r_{MO}^{MO})]. \end{aligned} \quad (39)$$

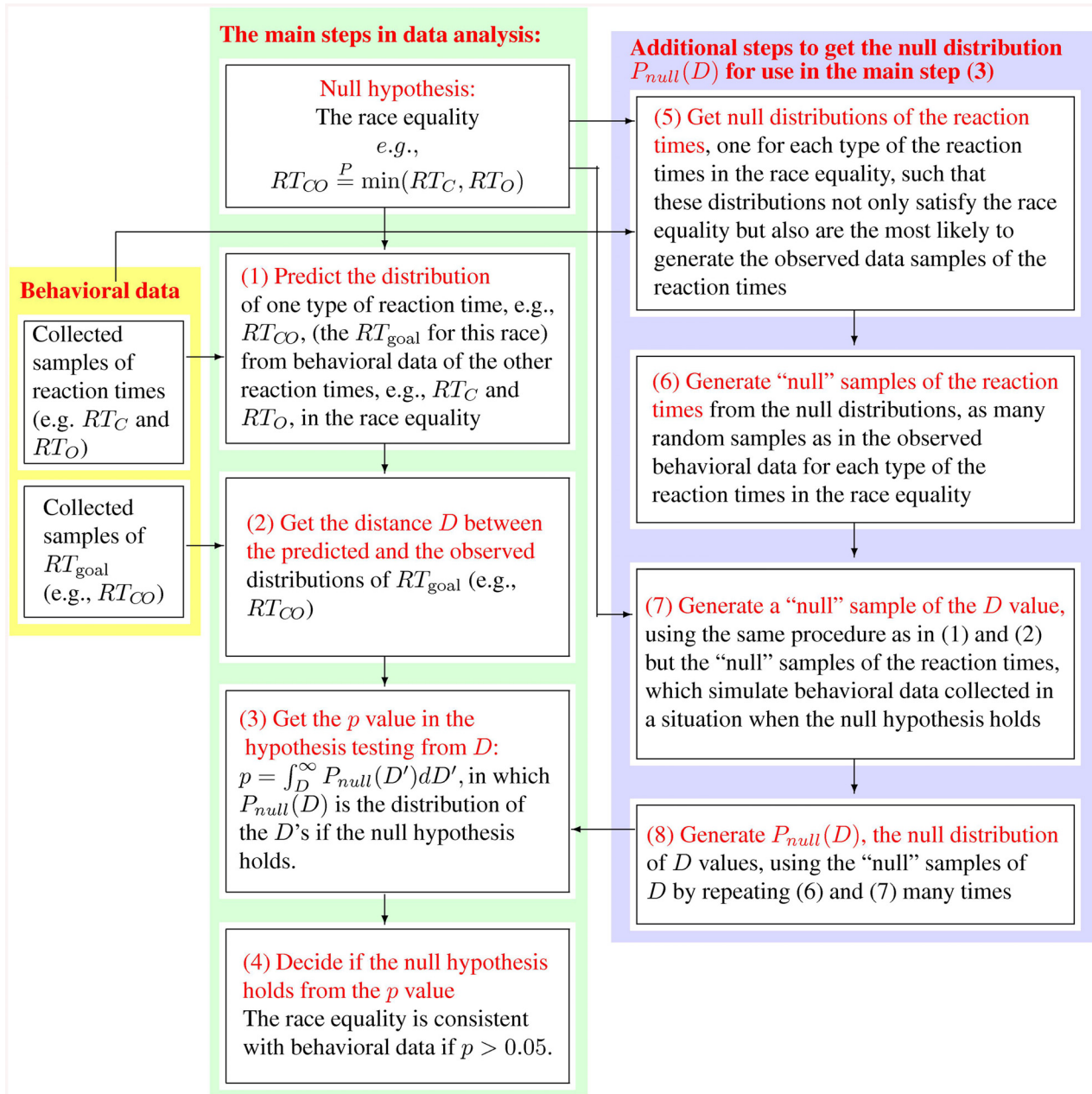
The list of the arguments in the  $f[\max(\dots)]$  above is the collection of all the non-trivial neural responses to the corresponding singletons. Similarly, writing  $\min(RT_{CM}, RT_{CO}, RT_{MO})$  from the right-hand side of our non-spurious equality as  $\min(RT_{CM}, RT_{CO}, RT_{MO}) \stackrel{P}{=} f[\max(\dots)]$  gives the same list of arguments in  $f[\max(\dots)]$  as in the equation above, thus proving the race equality.

In the list of arguments in  $f[\max(\dots)]$  of Eq (39), each of  $r_C^C$ ,  $r_O^O$  and  $r_M^M$  occurs twice as independent random samples. The list should not be simplified by deleting the repetitions, since the maximum of two random samples differs statistically from one random sample alone.

## Methods to test a race equality as a null hypothesis

Briefly, a race equality, e.g.,  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$ , is a null hypothesis. It is used to predict the distribution of  $RT_{goal}$ , the designated type of reaction times in the equality (e.g.,  $RT_{CO}$  is the  $RT_{goal}$  for  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$ ), from the behaviorally observed distributions of the other reaction times in the equality. A distance  $D$  is then calculated between the predicted distribution and the behaviorally observed one of  $RT_{goal}$ . Typically  $D$  is non-zero even when a race equality does hold, since finite numbers of data samples can only approximately represent the underlying distributions of various reaction times. A statistical test is devised to give a  $p$  value,





**Fig 11. Diagram outlining the methods to test each of our race equalities, *e.g.*,  $RT_{CO} \stackrel{P}{=} \min(RT_C, RT_O)$ .** The details of various components, in boxes (1)-(8), are described in the text.

doi:10.1371/journal.pcbi.1004375.g011

the probability that the  $D$  should be at least as big as observed if the null hypothesis holds. A  $p > 0.05$  is chosen to suggest that the race equality agrees with data. The details of the components of the hypothesis testing method, represented by the boxes in Fig 11, are described next.

**Methods to predict a distribution of reaction times from a race equality  $RE_i$ .** Here we describe the details for box (1) of Fig 11. First, given a race, *e.g.*,  $\min(RT_C, RT_O)$ ,  $\min(RT_C, RT_M, RT_O)$ , or  $\min(RT_{CMO}, RT_C, RT_M, RT_O)$ , the samples of the winner of the race are

winner samples of a race = the collection of the minimums, one from each of all possible combinations of the reaction time data samples from the racers, (40)



regardless of the number of racers. For example,  $m$  samples of  $RT_C$  and  $n$  samples of  $RT_O$  give us  $m \times n \min(RT_C, RT_O)$  samples from the  $m \times n$  possible combinations of  $RT_C$  and  $RT_O$  samples.

Each of our race equalities is in the format of  $RT1 \stackrel{P}{=} RT2$ , and a reaction time type, designated as  $RT_{goal}$  (listed in Table 1), in the equality is predicted from data samples of the other reaction times in the equality. In race equalities  $RE_i$  for  $i = 2-5$ ,  $RT_{goal}$  is  $RT1$  and is the reaction time of a double-feature or triple-feature singleton, its predicted distribution is that of the samples of the race winner  $RT2$  using data samples for the corresponding single-feature singletons (using Eq (40)).

In equalities  $RE_1$  and  $RE_i$  for  $i = 6-8$ , the  $RT_{goal}$  is always  $RT_{CMO}$ . We write  $RT1 \equiv \min(RT_{CMO}, RT_{part})$ , where, for  $RE_1, RE_6, RE_7$ , or  $RE_8$ , respectively,  $RT_{part}$  is  $\min(RT_C, RT_M, RT_O)$ ,  $\min(RT_M, RT_{CO})$ ,  $\min(RT_C, RT_{MO})$ , or  $\min(RT_O, RT_{CM})$ . Use Eq (40) to obtain samples for  $RT_{part}$  and  $RT2$  using behavioral data samples of  $RT_C, RT_M, RT_O, RT_{CM}, RT_{CO}$ , and  $RT_{MO}$ . Then, samples of  $RT_{part}, RT2$ , and data samples of  $RT_{CMO}$  are discretized into the same  $N$  time bins bounded by time values  $t_0 < t_1 < \dots < t_N$ . Different  $t_i$ 's are (in most data analysis) roughly evenly spaced except for very small and large  $t_i$ 's.  $N = 8-12$  is chosen to give sufficiently many behavioral data samples in each bin while maintaining a sufficiently large  $N$  for building a distribution.

Let distribution of any reaction times in  $N$  time bins be represented by an  $N$ -dimensional vector whose  $i^{th}$  component is  $n_i / (\sum_j n_j)$ , where  $n_i$  is the number of these reaction time samples in the  $i^{th}$  time bin. Let vectors  $\mathbf{P} \equiv (P_1, P_2, \dots, P_N)$  and  $\mathbf{Q} \equiv (Q_1, Q_2, \dots, Q_N)$  denote such distributions of  $RT1$  and  $RT2$ , respectively, and let  $\mathbf{p}$  and  $\mathbf{q}$  denote the distributions of  $RT_{CMO}$  and  $RT_{part}$ , respectively.  $RT1 \equiv \min(RT_{CMO}, RT_{part})$  means

$$P_i = p_i(1 - \sum_{j \leq i} q_j) + q_i(1 - \sum_{j \leq i} p_j) + p_i q_i, \quad \text{for all } i. \tag{41}$$

Then  $RT1 \stackrel{P}{=} RT2$  means  $P_i = Q_i$ , i.e.,

$$p_i(1 - \sum_{j \leq i} q_j) + q_i(1 - \sum_{j \leq i} p_j) + p_i q_i = Q_i, \quad \text{for all } i. \tag{42}$$

Given  $\mathbf{q}$  and  $\mathbf{Q}$  (obtained from samples of  $RT_{part}$  and  $RT2$ , respectively), solve for  $\mathbf{p}$  from the above linear equation. If this solution satisfies the probability constraints  $p_i \geq 0$  and  $\sum_i p_i = 1$ , it is our predicted distribution for  $RT_{CMO}$ . Otherwise (this can happen for example when  $q_i > Q_i$  for some  $i$  due to fluctuations in the limited data samples and/or due to a lack of the race equality in reality), the predicted  $\mathbf{p}$  is chosen as the one that minimizes a distance between  $\mathbf{P}$  and  $\mathbf{Q}$  under the constraints  $p_i \geq 0$  and  $\sum_i p_i = 1$  (through an optimization procedure, e.g., via the `fmincon` routine in MATLAB). The following four different distance measures (between  $\mathbf{P}$  and  $\mathbf{Q}$ ) were separately tried:

- (1) :  $|\mathbf{P} - \mathbf{Q}|^2$ , the squared Hemming distance,
- (2) :  $\sum_i (\sqrt{P_i} - \sqrt{Q_i})^2$ , the Hellinger distance,
- (3) :  $\sum_i |P_i - Q_i|$ , the 1-norm distance, and
- (4) :  $\sum_i \max(Q_i, \epsilon) \log \frac{\max(Q_i, \epsilon)}{\max(P_i, \epsilon)}$ , with a given  $\epsilon \ll 10^{-100}$ , the KL-like distance.

The last distance is the Kullback-Leibler divergence if all  $P_i$  and  $Q_i$  were larger than a very small  $\epsilon$ .

The boundaries  $t_i$  for the  $N$  time bins are determined as follows. Given a subject and a race equality, all the behavioral reaction time samples of all the singleton types in this race equality are put into a single pool. They are divided into  $L \gg N$  time bins ( $L = 100$  was used), whose boundaries

$$T_0 < T_1 < T_2 < \dots < T_L, \tag{44}$$

are such that all bins contain (as close as possible) an equal number of samples from this pool. For reasons that will be clear soon, each  $t_i$  is chosen from among these  $T_i$ 's as follows. Let  $RT(\max)$  and  $RT(\min)$ , respectively, denote the largest and smallest data samples of the collective pool of  $RT_{\text{goal}}$ ,  $RT_2$ , and (for  $RE_1$  and  $RE_i$  for  $i = 6-8$ )  $RT_{\text{part}}$  data samples. Given  $(T_0, T_1, \dots, T_L)$ ,  $t_0$  is the largest  $T_j$  smaller than  $RT(\min)$  and  $t_N$  is the smallest  $T_j$  larger than  $RT(\max)$ . Then, let  $RT'(\max)$  and  $RT'(\min)$  denote the largest and smallest  $RT_{\text{goal}}$  data samples, respectively. If  $RT'(\min) > RT(\min)$  and the largest  $T_j$  smaller than  $RT'(\min)$  is larger than  $t_0$ , then this  $T_j$  is assigned to  $t_1$ . If  $RT'(\max) < RT(\max)$  and the smallest  $T_j$  larger than  $RT'(\max)$  is smaller than  $t_N$ , then this  $T_j$  is assigned to  $t_{N-1}$ . Depending on whether  $t_1$  and  $t_{N-1}$  have just been assigned, there are now  $N' = N - 1, N - 2$ , or  $N - 3$  of the unassigned  $t_i$ , which will be assigned in ascending order to  $\tau_1 < \tau_2 < \dots < \tau_{N'}$ . Each  $\tau_i$  is the  $T_j$  value not yet assigned to any  $t_j$  for any  $j$  and is closest to the value  $\tau'_i$  which is larger than a fraction  $F_i$  (with  $F_1 < F_2 < \dots < F_{N'}$ ) of the  $RT_{\text{goal}}$  data samples. We tried each of the following four ways to choose  $F_i$ 's. One is  $F_i = i/(N' + 1)$ . The others are

$$F_i = \frac{1}{2} \left( \operatorname{erf} \left( -x + 2x \frac{i-1}{N'-1} \right) + 1 \right), \tag{45}$$

in which  $\operatorname{erf}(\cdot)$  is the error function and  $x$  is a parameter with value  $x = 1.25, 1.35$ , or  $1.45$ .

**The statistical test for the null hypothesis.** The Kolmogorov-Smirnov test cannot be used to test whether  $RT_1$  samples and  $RT_2$  samples are generated from the same underlying distribution, because the samples of at least  $RT_2$  are not independently generated. The following describes the methods in boxes (2)-(8) of Fig 11 for testing whether the predicted and observed distributions of  $RT_{\text{goal}}$  arise from the same underlying entity.

Given an observer and a race equality, let  $\mathbf{p}$  and  $\tilde{\mathbf{p}}$  be the  $N$ -dimensional vectors for predicted and observed distributions of  $RT_{\text{goal}}$ , respectively, in our  $N$  time bins. The distance  $D$  between  $\mathbf{p}$  and  $\tilde{\mathbf{p}}$  (for box (2) of Fig 11) is measured by one of the four distance metrics in Eq (43), substituting  $\tilde{\mathbf{p}}$  and  $\mathbf{p}$  for  $\mathbf{P}$  and  $\mathbf{Q}$ , respectively.

To test whether  $\tilde{\mathbf{p}}$  and  $\mathbf{p}$  are statistically the same, we generated  $m = 500$  other, simulated, distances  $D$  (box (8) of Fig 11). Each simulated  $D$  is a “null” sample for box (7) of Fig 11. It is obtained from a set of simulated samples of reaction times collected from a simulated behavioral experiment in a hypothetical situation when the race equality holds while the simulated data samples resemble the real behavioral data samples in terms of their distributions. Given the fixed time boundaries  $T_0 < T_1 < T_2 < \dots < T_L$  (Eq (44)) obtained from the real behavioral data, the procedure to obtain a (simulated)  $D$  value using the simulated data samples is the same as that when the real data samples are used. The  $p$  value of the statistical test (box (3) of Fig 11) is the fraction of the simulated  $D$  values that are larger than the real  $D$  value (obtained using the real behavioral data), a  $p < 1/m = 0.002$  is given when this fraction is zero. Our predicted and observed distributions of  $RT_{\text{goal}}$  are said to be significantly different from each other, i.e., not arising from the same underlying entity, and we declare that the race equality is broken, when  $p < 0.05$  (box (4) of Fig 11)

To obtain simulated samples of reaction times for a race equality (box (6) of Fig 11), we should have already constructed (detailed in the next paragraph) a set of probability

distributions, called the null distributions (of box (5) in Fig 11), for the reaction times involved in this race equality. The null distributions satisfy the race equality while being most likely to be the underlying distributions from which the behaviorally observed samples of reaction times could be generated. For example, for equality  $RT_{CO} \stackrel{p}{=} \min(RT_C, RT_O)$ , the null distributions include three distributions, one each for  $RT_{CO}$ ,  $RT_C$ , and  $RT_O$ , respectively. From each of these null distributions, as many simulated samples of reaction times as the corresponding real behavioral samples of reaction times (for the corresponding singleton type) are randomly generated.

The null distributions in box (5) of Fig 11 are constructed as follows. Given a subject and a race equality, the real  $RT_\alpha$  samples for all the singleton types  $\alpha$  in the equality are discretized into  $L$  time bins using time boundaries  $T_0 < T_1 < \dots < T_L$  in Eq (44). Let  $\mathbf{n}_\alpha \equiv [(n_\alpha)_1, (n_\alpha)_2, \dots, (n_\alpha)_L]$ , and  $(n_\alpha)_i$  is the number of  $RT_\alpha$  samples in the  $i^{th}$  time bin. The likelihood, or probability, that an underlying distribution  $\hat{\mathbf{p}}_\alpha \equiv (\hat{p}_{\alpha_1}, \hat{p}_{\alpha_2}, \dots, \hat{p}_{\alpha_L})$  over these bins is the generator of  $\mathbf{n}_\alpha$  is proportional to  $\prod_{i=1}^L (\hat{p}_{\alpha_i})^{n_{\alpha_i}}$ , whose logarithm is  $\sum_{i=1}^L n_{\alpha_i} \ln \hat{p}_{\alpha_i} + \text{constant}$ . We construct null distributions  $\hat{\mathbf{p}}_\alpha$ , one for each singleton type  $\alpha$  in the race equality, such that the total log-likelihood

$$\sum_{\alpha} \sum_{i=1}^L n_{\alpha_i} \ln \hat{p}_{\alpha_i} + \text{constant} \tag{46}$$

is maximized, subject to the constraints that the race equality  $RT1 \stackrel{p}{=} RT2$  (which takes the form like Eqs (41)–(42)) is satisfied by these  $\hat{\mathbf{p}}_\alpha$ s and, for each  $\alpha$ ,  $\sum_{i=1}^L \hat{p}_{\alpha_i} = 1$  and  $\hat{p}_{\alpha_i} \geq 0$ . The resulting  $\hat{\mathbf{p}}_\alpha$ 's obtained through an optimization procedure (e.g., using `fmincon` in MATLAB) were verified to satisfy the race equality and sufficiently resemble the respective histograms of behavioral data  $RT_\alpha$ .

When each  $\hat{\mathbf{p}}_\alpha$  is viewed through coarser time bins for predicting the  $RT_{\text{goal}}$  distribution, the race equality remains satisfied since the boundaries  $t_i$  for these coarser time bins were chosen from those  $T_j$ 's for the finer time bins. Although irrelevant to our outcome, the null distributions over continuous time can be approximated by (for each  $RT_\alpha$ ) a uniform probability density  $\hat{p}_{\alpha_i} / (T_i - T_{i-1})$  within the time window  $[T_{i-1}, T_i]$  and zero outside  $(T_0, T_L)$ .

## Acknowledgments

We like to thank Peter Dayan, Peter Latham, and Keith May for careful reading of the manuscript with very helpful comments for presentations.

## Author Contributions

Analyzed the data: LZhe LZha. Wrote the paper: LZha. Conceived and designed the project: LZha.

## References

1. Itti L, Koch C. Computational modelling of visual attention. *Nature Reviews Neuroscience*. 2001; 2(3):194–203. doi: [10.1038/35058500](https://doi.org/10.1038/35058500) PMID: [11256080](https://pubmed.ncbi.nlm.nih.gov/11256080/)
2. Treisman AM, Gelade G. A feature-integration theory of attention. *Cognitive Psychology*. 1980; 12(1):97–136. doi: [10.1016/0010-0285\(80\)90005-5](https://doi.org/10.1016/0010-0285(80)90005-5) PMID: [7351125](https://pubmed.ncbi.nlm.nih.gov/7351125/)
3. Müller HJ, Rabbitt PM. Reflexive and voluntary orienting of visual attention: time course of activation and resistance to interruption. *Journal of Experimental Psychology: Human Perception and Performance*. 1989; 15(2):315–330. PMID: [2525601](https://pubmed.ncbi.nlm.nih.gov/2525601/)

4. Nakayama K, Mackeben M. Sustained and transient components of focal visual attention. *Vision Research*. 1989; 29(11):631–47. doi: [10.1016/0042-6989\(89\)90144-2](https://doi.org/10.1016/0042-6989(89)90144-2)
5. Duncan J, Humphreys GW. Visual search and stimulus similarity. *Psychological Review*. 1989; 96(3):433–58. doi: [10.1037/0033-295X.96.3.433](https://doi.org/10.1037/0033-295X.96.3.433) PMID: [2756067](https://pubmed.ncbi.nlm.nih.gov/2756067/)
6. Wolfe JM, Cave KR, Franzel SL. Guided search: an alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*. 1989; 15:419–433. PMID: [2527952](https://pubmed.ncbi.nlm.nih.gov/2527952/)
7. Li Z. Contextual influences in V1 as a basis for pop out and asymmetry in visual search. *Proceedings of the National Academy of Sciences of the USA*. 1999; 96(18):10530–10535. doi: [10.1073/pnas.96.18.10530](https://doi.org/10.1073/pnas.96.18.10530) PMID: [10468643](https://pubmed.ncbi.nlm.nih.gov/10468643/)
8. Li Z. A saliency map in primary visual cortex. *Trends in Cognitive Sciences*. 2002; 6(1):9–16. doi: [10.1016/S1364-6613\(00\)01817-9](https://doi.org/10.1016/S1364-6613(00)01817-9) PMID: [11849610](https://pubmed.ncbi.nlm.nih.gov/11849610/)
9. Zhaoping L. *Understanding Vision: theory, models, and data*. Oxford University Press; 2014.
10. Gottlieb JP, Kusunoki M, Goldberg ME. The representation of visual salience in monkey parietal cortex. *Nature*. 1998; 391(6666):481–4. doi: [10.1038/35135](https://doi.org/10.1038/35135) PMID: [9461214](https://pubmed.ncbi.nlm.nih.gov/9461214/)
11. Schiller PH. The neural control of visually guided eye movements. In: Richards JE, editor. *Cognitive Neuroscience of Attention, a Developmental Perspective*. Mahwah, New Jersey, USA: Lawrence Erlbaum Associates, Inc.; 1998. p. 3–50.
12. Allman J, Miezin F, McGuinness E. Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. *Annual Review of Neuroscience*. 1985; 8:407–30. doi: [10.1146/annurev.ne.08.030185.002203](https://doi.org/10.1146/annurev.ne.08.030185.002203) PMID: [3885829](https://pubmed.ncbi.nlm.nih.gov/3885829/)
13. Knierim JJ, Van Essen DC. Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology*. 1992; 67(4):961–80. PMID: [1588394](https://pubmed.ncbi.nlm.nih.gov/1588394/)
14. Li CY, Li W. Extensive integration field beyond the classical receptive field of cat's striate cortical neurons—classification and tuning properties. *Vision Research*. 1994; 34(18):2337–55. doi: [10.1016/0042-6989\(94\)90280-1](https://doi.org/10.1016/0042-6989(94)90280-1) PMID: [7975275](https://pubmed.ncbi.nlm.nih.gov/7975275/)
15. Hubel DH, Wiesel TN. Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*. 1968; 195(1):215–43. doi: [10.1113/jphysiol.1968.sp008455](https://doi.org/10.1113/jphysiol.1968.sp008455) PMID: [4966457](https://pubmed.ncbi.nlm.nih.gov/4966457/)
16. Livingstone MS, Hubel DH. Anatomy and physiology of a color system in the primate visual cortex. *The Journal of Neuroscience*. 1984; 4(1):309–56. PMID: [6198495](https://pubmed.ncbi.nlm.nih.gov/6198495/)
17. Wachtler T, Sejnowski TJ, Albright TD. Representation of color stimuli in awake macaque primary visual cortex. *Neuron*. 2003; 37(4):681–91. doi: [10.1016/S0896-6273\(03\)00035-7](https://doi.org/10.1016/S0896-6273(03)00035-7) PMID: [12597864](https://pubmed.ncbi.nlm.nih.gov/12597864/)
18. DeAngelis GC, Freeman RD, Ohzawa I. Length and width tuning of neurons in the cat's primary visual cortex. *Journal of Neurophysiology*. 1994; 71:347–374. PMID: [8158236](https://pubmed.ncbi.nlm.nih.gov/8158236/)
19. Jones HE, Grieve KL, Wang W, Sillito AM. Surround suppression in primate V1. *Journal of Neurophysiology*. 2001; 86(4):2011–28. PMID: [11600658](https://pubmed.ncbi.nlm.nih.gov/11600658/)
20. Cavanaugh JR, Bair W, Movshon JA. Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. *Journal of Neurophysiology*. 2002; 88(5):2547–2556. doi: [10.1152/jn.00693.2001](https://doi.org/10.1152/jn.00693.2001) PMID: [12424293](https://pubmed.ncbi.nlm.nih.gov/12424293/)
21. Zhaoping L. Attention capture by eye of origin singletons even without awareness—a hallmark of a bottom-up saliency map in the primary visual cortex. *Journal of Vision*. 2008; 8(5):article 1. Available from: <http://journalofvision.org/8/5/1/> doi: [10.1167/8.5.1](https://doi.org/10.1167/8.5.1)
22. Rockland KS, Lund JS. Intrinsic laminar lattice connections in primate visual cortex. *The Journal of Comparative Neurology*. 1983; 216(3):303–18. doi: [10.1002/cne.902160307](https://doi.org/10.1002/cne.902160307) PMID: [6306066](https://pubmed.ncbi.nlm.nih.gov/6306066/)
23. Gilbert CD, Wiesel TN. Clustered intrinsic connections in cat visual cortex. *The Journal of Neuroscience*. 1983; 3(5):1116–33. PMID: [6188819](https://pubmed.ncbi.nlm.nih.gov/6188819/)
24. Li Z. A neural model of contour integration in the primary visual cortex. *Neural Computation*. 1998; 10(4):903–40. doi: [10.1162/089976698300017557](https://doi.org/10.1162/089976698300017557) PMID: [9573412](https://pubmed.ncbi.nlm.nih.gov/9573412/)
25. Li Z. Visual segmentation by contextual influences via intra-cortical interactions in primary visual cortex. *Network: Computation in Neural Systems*. 1999; 10(2):187–212. doi: [10.1088/0954-898X/10/2/305](https://doi.org/10.1088/0954-898X/10/2/305)
26. Li Z. Pre-attentive segmentation in the primary visual cortex. *Spatial Vision*. 2000; 13(1):25–50. doi: [10.1163/156856800741009](https://doi.org/10.1163/156856800741009) PMID: [10688228](https://pubmed.ncbi.nlm.nih.gov/10688228/)
27. Zhaoping L. V1 mechanisms and some figure-ground and border effects. *Journal of Physiology, Paris*. 2003; 97(4-6):503–515. doi: [10.1016/j.jphysparis.2004.01.008](https://doi.org/10.1016/j.jphysparis.2004.01.008)
28. Zhaoping L, Snowden RJ. A theory of a saliency map in primary visual cortex (V1) tested by psychophysics of color-orientation interference in texture segmentation. *Visual Cognition*. 2006; 14(4-8):911–933. doi: [10.1080/13506280500196035](https://doi.org/10.1080/13506280500196035)

29. Koene AR, Zhaoping L. Feature-specific interactions in saliency from combined feature contrasts: Evidence for a bottom-up saliency map in V1. *Journal of Vision*. 2007; 7(7):article 6. Available from: <http://journalofvision.org/7/7/6/> doi: [10.1167/7.7.6](https://doi.org/10.1167/7.7.6)
30. Zhaoping L, May KA. Psychophysical tests of the hypothesis of a bottom-up saliency map in primary visual cortex. *PLoS Computational Biology*. 2007; 3(4):e62. doi: [10.1371/journal.pcbi.0030062](https://doi.org/10.1371/journal.pcbi.0030062) PMID: [17411335](https://pubmed.ncbi.nlm.nih.gov/17411335/)
31. Jingling L, Zhaoping L. Change detection is easier at texture border bars when they are parallel to the border: Evidence for V1 mechanisms of bottom-up saliency. *Perception*. 2008; 37(2):197–206. doi: [10.1068/p5829](https://doi.org/10.1068/p5829) PMID: [18456924](https://pubmed.ncbi.nlm.nih.gov/18456924/)
32. Zhang X, Zhaoping L, Zhou T, Fang F. Neural activities in V1 create a bottom-up saliency map. *Neuron*. 2011; 73:183–192. doi: [10.1016/j.neuron.2011.10.035](https://doi.org/10.1016/j.neuron.2011.10.035)
33. Zhaoping L. Gaze capture by eye-of-origin singletons: Interdependence with awareness. *Journal of Vision*. 2012; 12(2):article 17. Available from: <http://journalofvision.org/12/2/17/> doi: [10.1167/12.2.17](https://doi.org/10.1167/12.2.17)
34. Hubel DH, Wiesel TN. Ferrier lecture: Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London Series B, Biological Sciences*. 1977; 198(1130):1–59. doi: [10.1098/rspb.1977.0085](https://doi.org/10.1098/rspb.1977.0085) PMID: [20635](https://pubmed.ncbi.nlm.nih.gov/20635/)
35. Burkhalter A, Van Essen DC. Processing of color, form and disparity information in visual areas VP and V2 of ventral extrastriate cortex in the macaque monkey. *The Journal of Neuroscience*. 1986; 6:2327–2351. PMID: [3746412](https://pubmed.ncbi.nlm.nih.gov/3746412/)
36. Nelson JL, Frost BJ. Intracortical facilitation among co-oriented, co-axially aligned simple cells in cat striate cortex. *Experimental Brain Research*. 1985; 61(1):54–61. doi: [10.1007/BF00235620](https://doi.org/10.1007/BF00235620) PMID: [4085603](https://pubmed.ncbi.nlm.nih.gov/4085603/)
37. Kapadia MK, Ito M, Gilbert CD, Westheimer G. Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron*. 1995; 15(4):843–56. doi: [10.1016/0896-6273\(95\)90175-2](https://doi.org/10.1016/0896-6273(95)90175-2) PMID: [7576633](https://pubmed.ncbi.nlm.nih.gov/7576633/)
38. Levitt JB, Lund JS. Contrast dependence of contextual effects in primate visual cortex. *Nature*. 1997; 387(6628):73–76. doi: [10.1038/387073a0](https://doi.org/10.1038/387073a0) PMID: [9139823](https://pubmed.ncbi.nlm.nih.gov/9139823/)
39. Polat U, Mizobe K, Pettet MW, Kasamatsu T, Norcia AM. Collinear stimuli regulate visual responses depending on cell's contrast threshold. *Nature*. 1998; 391(6667):580–584. doi: [10.1038/35372](https://doi.org/10.1038/35372) PMID: [9468134](https://pubmed.ncbi.nlm.nih.gov/9468134/)
40. Zhaoping L, Zhe L. V1 saliency theory makes quantitative, zero parameter, prediction of reaction times in visual search of feature singletons. *Journal of Vision*. 2012; 12(9):1160–1160. doi: [10.1167/12.9.1160](https://doi.org/10.1167/12.9.1160)
41. Zhaoping L. A theory of the primary visual cortex (V1): Predictions, experimental tests, and implications for future research. *Perception*. 2013; 42:ECVP Abstract Supplement, page 84. Available from: <http://www.perceptionweb.com/abstract.cgi?id=v132001>
42. Zhaoping L. Theoretical understanding of the early visual processes by data compression and data selection. *Network: Computation in Neural Systems*. 2006; 17(4):301–334. doi: [10.1080/09548980600931995](https://doi.org/10.1080/09548980600931995)
43. Rabb DH. Statistical facilitation of simple reaction times. *Trans N Y Acad Sci*. 1962; 24:574–90. doi: [10.1111/j.2164-0947.1962.tb01433.x](https://doi.org/10.1111/j.2164-0947.1962.tb01433.x)
44. Webster MA, Miyahara E, Malkoc G, Raker VE. Variations in normal color vision. I. Cone-opponent axes. *Journal of Optical Society of America, A*. 2000; 17(9):1535–1544. doi: [10.1364/JOSAA.17.001535](https://doi.org/10.1364/JOSAA.17.001535) PMID: [10975363](https://pubmed.ncbi.nlm.nih.gov/10975363/)
45. Kanai R, Rees G. The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews Neuroscience*. 2011; 12:231–242. doi: [10.1038/nrn3000](https://doi.org/10.1038/nrn3000) PMID: [21407245](https://pubmed.ncbi.nlm.nih.gov/21407245/)
46. Horwitz GD, Albright TD. Paucity of chromatic linear motion detectors in macaque V1. *Journal of Vision*. 2005; 5(6):article 4. doi: [10.1167/5.6.4](https://doi.org/10.1167/5.6.4)
47. Michael CR. Color vision mechanisms in monkey striate cortex: simple cells with dual opponent-color receptive fields. *Journal of Neurophysiology*. 1978; 41(5):1233–1249. PMID: [100586](https://pubmed.ncbi.nlm.nih.gov/100586/)
48. Tamura H, Sato H, Katsuyama N, Hata Y, Tsumoto T. Less segregated processing of visual information in V2 than in V1 of the monkey visual cortex. *The European Journal of Neuroscience*. 1996; 8(2):300–9. doi: [10.1111/j.1460-9568.1996.tb01214.x](https://doi.org/10.1111/j.1460-9568.1996.tb01214.x) PMID: [8714701](https://pubmed.ncbi.nlm.nih.gov/8714701/)
49. Krummenacher J, Müller HJ, Heller D. Visual search for dimensionally redundant pop-out targets: Evidence for parallel-coactive processing of dimensions. *Perception & Psychophysics*. 2001; 63(5):901–917. doi: [10.3758/BF03194446](https://doi.org/10.3758/BF03194446)
50. Zhaoping L, Zhe L. Properties of V1 neurons tuned to conjunctions of visual features: application of the V1 saliency hypothesis to visual search behavior. *PLoS One*. 2012; 7(6):e36223. doi: [10.1371/journal.pone.0036223](https://doi.org/10.1371/journal.pone.0036223) PMID: [22719829](https://pubmed.ncbi.nlm.nih.gov/22719829/)



51. Heeger DJ. Normalization of cell responses in cat striate cortex. *Visual Neuroscience*. 1992; 9(02):181–197. doi: [10.1017/S0952523800009640](https://doi.org/10.1017/S0952523800009640) PMID: [1504027](https://pubmed.ncbi.nlm.nih.gov/1504027/)
52. Albrecht DG, Hamilton DB. Striate cortex of monkey and cat: contrast response function. *Journal of Neurophysiology*. 1982; 48(1):217–237. PMID: [7119846](https://pubmed.ncbi.nlm.nih.gov/7119846/)
53. Hegdé J, Van Essen DC. A comparative study of shape representation in macaque visual areas V2 and V4. *Cerebral Cortex*. 2007; 17(5):1100–1116. PMID: [16785255](https://pubmed.ncbi.nlm.nih.gov/16785255/)
54. Desimone R, Schein SJ. Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *Journal of Neurophysiology*. 1987; 57(3):835–868. PMID: [3559704](https://pubmed.ncbi.nlm.nih.gov/3559704/)
55. Schein SJ, Desimone R. Spectral properties of V4 neurons in the macaque. *The Journal of Neuroscience*. 1990; 10(10):3369–3389. PMID: [2213146](https://pubmed.ncbi.nlm.nih.gov/2213146/)
56. Van Essen DC, DeYoe EA, Olavarria JF, Knierim JJ, Fox JM. Neural responses to static and moving texture patterns in visual cortex of the macaque monkey. In: Lam DK, Gilbert C, editors. *Models of the Visual Cortex*. Portfolio Publishing: Woodlands, TX.; 1989. p. 137–154.
57. Gegenfurtner KR, Kiper DC, Fenstemaker SB. Processing of color, form, and motion in macaque area V2. *Visual Neuroscience*. 1996; 13:161–172. doi: [10.1017/S0952523800007203](https://doi.org/10.1017/S0952523800007203) PMID: [8730997](https://pubmed.ncbi.nlm.nih.gov/8730997/)
58. Shipp S, Adams D, Moutoussis K, Zeki S. Feature binding in the feedback layers of area V2. *Cerebral Cortex*. 2009; 19(10):2230–9. doi: [10.1093/cercor/bhn243](https://doi.org/10.1093/cercor/bhn243) PMID: [19153106](https://pubmed.ncbi.nlm.nih.gov/19153106/)
59. Thorpe S, Fize D, Marlot C. Speed of processing in the human visual system. *Nature*. 1996; 381(6582):520–522. doi: [10.1038/381520a0](https://doi.org/10.1038/381520a0) PMID: [8632824](https://pubmed.ncbi.nlm.nih.gov/8632824/)
60. Henderson JM, Weeks JPA, Hollingworth A. The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*. 1999; 25(1):210–228.
61. Torralba A, Oliva A, Castelano MS, Henderson JM. Contextual guidance of eye movements and attention in real-world scenes: The role of global features on object search. *Psychological Review*. 2006; 113(4):766–786. doi: [10.1037/0033-295X.113.4.766](https://doi.org/10.1037/0033-295X.113.4.766) PMID: [17014302](https://pubmed.ncbi.nlm.nih.gov/17014302/)
62. Mannan SK, Kennard C, Husain M. The role of visual salience in directing eye movements in visual object agnosia. *Current Biology*. 2009; 19(6):R247–8. doi: [10.1016/j.cub.2009.02.020](https://doi.org/10.1016/j.cub.2009.02.020) PMID: [19321139](https://pubmed.ncbi.nlm.nih.gov/19321139/)
63. Einhäuser W, Spain M, Perona P. Objects predict fixations better than early saliency. *Journal of Vision*. 2008 11; 8(14):article 18. Available from: <http://jov.journalofvision.org/8/14/18/> doi: [10.1167/8.14.18](https://doi.org/10.1167/8.14.18)
64. Treisman A, Gormican S. Feature analysis in early vision: evidence from search asymmetries. *Psychological Review*. 1988; 95(1):15–48. doi: [10.1037/0033-295X.95.1.15](https://doi.org/10.1037/0033-295X.95.1.15) PMID: [3353475](https://pubmed.ncbi.nlm.nih.gov/3353475/)
65. Kennett MJ, Wallis G, Zhaoping L. The schematic angry face effect: threat detection or V1 processes? *Perception 43 ECVF Abstract Supplement*. 2014;p. 65.
66. Krummenacher J, Müller HJ. Visual search for singleton targets redundantly defined in two feature dimensions: Coactive processing of color-motion targets? *Journal of Experimental Psychology: Human Perception and Performance*. 2014; 40(5):1926–1939. PMID: [25089576](https://pubmed.ncbi.nlm.nih.gov/25089576/)
67. Motter BC, Belky EJ. The zone of focal attention during active visual search. *Vision Research*. 1998; 38(7):1007–22. doi: [10.1016/S0042-6989\(97\)00252-6](https://doi.org/10.1016/S0042-6989(97)00252-6) PMID: [9666983](https://pubmed.ncbi.nlm.nih.gov/9666983/)
68. Müller HJ, Heller D, Ziegler J. Visual search for singleton feature targets within and across feature dimensions. *Perception & Psychophysics*. 1995; 57(1):1–17. doi: [10.3758/BF03211845](https://doi.org/10.3758/BF03211845)
69. Müller HJ, Reimann B, Krummenacher J. Visual search for singleton feature targets across dimensions: Stimulus- and expectancy-driven effects in dimensional weighting. *Journal of Experimental Psychology: Human Perception and Performance*. 2003; 29(5):1021–1035. PMID: [14585020](https://pubmed.ncbi.nlm.nih.gov/14585020/)
70. Zhou H, Desimone R. Feature-based attention in the frontal eye field and area V4 during visual search. *Neuron*. 2011; 70(6):1205–1217. doi: [10.1016/j.neuron.2011.04.032](https://doi.org/10.1016/j.neuron.2011.04.032) PMID: [21689605](https://pubmed.ncbi.nlm.nih.gov/21689605/)
71. Chun MM, Jiang Y. Contextual cueing: implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*. 1998; 36:28–71. doi: [10.1006/cogp.1998.0681](https://doi.org/10.1006/cogp.1998.0681) PMID: [9679076](https://pubmed.ncbi.nlm.nih.gov/9679076/)
72. Betz T, Wilming N, Bogler C, Haynes JD, König P. Dissociation between saliency signals and activity in early visual cortex. *Journal of Vision*. 2013; 13(6):article 6. doi: [10.1167/13.14.6](https://doi.org/10.1167/13.14.6)
73. Ben-Tov M, Donchin O, Ben-Shahar O, Segev R. Pop-out in Visual search of moving targets in the archer fish. *Nature Communications*. 2015; 6:article number 6476. doi: [10.1038/ncomms7476](https://doi.org/10.1038/ncomms7476) PMID: [25753807](https://pubmed.ncbi.nlm.nih.gov/25753807/)
74. Zhao X, Liu M, Cang J. Visual cortex modulates the magnitude but not the selectivity of looming-evoked responses in the superior colliculus of awake mice. *Neuron*. 2014; 84(1):202–213. doi: [10.1016/j.neuron.2014.08.037](https://doi.org/10.1016/j.neuron.2014.08.037) PMID: [25220812](https://pubmed.ncbi.nlm.nih.gov/25220812/)
75. Fecteau JH, Munoz DP. Saliency, relevance, and firing: a priority map for target selection. *Trends in Cognitive Sciences*. 2006; 10(8):382–390. doi: [10.1016/j.tics.2006.06.011](https://doi.org/10.1016/j.tics.2006.06.011) PMID: [16843702](https://pubmed.ncbi.nlm.nih.gov/16843702/)



76. Bisley JW, Goldberg ME. Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience*. 2011; 33:1–21. doi: [10.1146/annurev-neuro-060909-152823](https://doi.org/10.1146/annurev-neuro-060909-152823)
77. Thompson KG, Bichot NP. A visual salience map in the primate frontal eye field. *Progress in Brain Research*. 2005; 147:249–262. doi: [10.1016/S0079-6123\(04\)47019-8](https://doi.org/10.1016/S0079-6123(04)47019-8)
78. Bogler C, Bode S, Haynes JD. Decoding successive computational stages of saliency processing. *Current Biology*. 2011; 21(19):1667–1671. doi: [10.1016/j.cub.2011.08.039](https://doi.org/10.1016/j.cub.2011.08.039) PMID: [21962709](https://pubmed.ncbi.nlm.nih.gov/21962709/)
79. McPeck RM, Keller EL. Saccade target selection in the superior colliculus during a visual search task. *Journal of Neurophysiology*. 2002; 88(4):2019–2034. PMID: [12364525](https://pubmed.ncbi.nlm.nih.gov/12364525/)
80. McPeck RM. Incomplete suppression of distractor-related activity in the frontal eye field results in curved saccades. *Journal of Neurophysiology*. 2006; 96(5):2699–2711. doi: [10.1152/jn.00564.2006](https://doi.org/10.1152/jn.00564.2006) PMID: [16885521](https://pubmed.ncbi.nlm.nih.gov/16885521/)
81. Thomas NW, Paré M. Temporal processing of saccade targets in parietal cortex area LIP during visual search. *Journal of Neurophysiology*. 2007; 97(1):942–947. doi: [10.1152/jn.00413.2006](https://doi.org/10.1152/jn.00413.2006) PMID: [17079346](https://pubmed.ncbi.nlm.nih.gov/17079346/)
82. Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*. 2002; 3:201–15. doi: [10.1038/nrn755](https://doi.org/10.1038/nrn755) PMID: [11994752](https://pubmed.ncbi.nlm.nih.gov/11994752/)
83. Kustov AA, Robinson DL. Shared neural control of attentional shifts and eye movements. *Nature*. 1996; 384:74–77. doi: [10.1038/384074a0](https://doi.org/10.1038/384074a0) PMID: [8900281](https://pubmed.ncbi.nlm.nih.gov/8900281/)
84. Desimone R, Duncan J. Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*. 1995; 18:193–222. with appropriate other details see <http://www.annualreviews.org/doi/abs/10.1146/annurev.ne.18.030195.001205?journalCode=neuro>
85. Crick F, Koch C. Are we aware of neural activity in primary visual cortex? *Nature*. 1995; 375(6527):121–3. doi: [10.1038/375121a0](https://doi.org/10.1038/375121a0) PMID: [7753166](https://pubmed.ncbi.nlm.nih.gov/7753166/)
86. Schiller PH, Lee K. The role of the primate extrastriate area V4 in vision. *Science*. 1991; 251(4998):1251–1253. doi: [10.1126/science.2006413](https://doi.org/10.1126/science.2006413) PMID: [2006413](https://pubmed.ncbi.nlm.nih.gov/2006413/)
87. Leopold DA, Logothetis NK. Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature*. 1996; 379(6565):549–553. doi: [10.1038/379549a0](https://doi.org/10.1038/379549a0) PMID: [8596635](https://pubmed.ncbi.nlm.nih.gov/8596635/)