



Distinguishing target biases and strategic guesses in visual working memory

Liqliang Huang¹

Published online: 22 November 2019
© The Psychonomic Society, Inc. 2019

Abstract

Previous studies have shown that when there are statistical regularities in the items stored in visual working memory, the responses are biased toward the ensemble average. This statistical-regularity-induced bias could happen in two ways: (1) a *target bias*, where the individual memory representations are pulled toward the ensemble average; or (2) a *strategic guess*, for items that are not memorized, other information in the ensemble (e.g., another item) is reported as a substitute. Here, these two mechanisms are distinguished on the basis of a three-part model (target responses + swap responses + random guesses; e.g., Bays, Catalao, & Husain, 2009, *Journal of Vision*, 9, 7). The strategic guess is operationalized as swap responses, whereas the target bias is reflected by a bias parameter in the target responses. This model was applied on 8 data sets (22 observers each). In this model, contributions of target biases and strategic guesses can be clearly distinguished from each other because they lead to distinctive patterns in the distribution of responses. In the present results, strategic guesses always contributed substantially to the statistical-regularity-induced biases, whereas target biases were limited to specific conditions. All in all, the Bayesian inference in visual working memory is much more limited than what is previously advocated.

Keywords Working memory · Ensemble · Memory bias

Statistical-regularity-induced bias

The human visual system constantly faces many objects and processes a large amount of information. However, only a small subset of this information can be temporarily retained in the mind. This mechanism—namely, visual working memory—has been studied extensively by cognitive psychologists (e.g., Alvarez & Cavanagh, 2004; Huang, 2010b; Olson

Public significance Previous studies have indicated that statistical regularity can lead to systematic biases in observers' responses. In this study, I tried to conduct a systematic exploration of the target biases (i.e., the individual memory representations are pulled toward the ensemble average) and the strategic guesses (e.g., for items that are not memorized, another item is reported as a substitute) in a broader range of stimuli. In the present study, by using a three-part model, contributions of target biases and strategic guesses can be clearly distinguished from each other because they lead to distinctive patterns in the distribution of responses. In the present results, strategic guesses always contributed substantially to the statistical-regularity-induced biases, whereas the target biases were limited to specific conditions.

✉ Liqliang Huang
lqhuang@psy.cuhk.edu.hk

¹ Department of Psychology, The Chinese University of Hong Kong, Hong Kong, China

& Jiang, 2002; Pashler, 1988; Phillips, 1974; Wheeler & Treisman, 2002; Xu, 2002; Zhang & Luck, 2008).

The amount of information in the visual world is often very large and goes well beyond the capacity limit of visual working memory. However, it is also often highly structured and therefore contains a lot of redundant information that can potentially be exploited. For example, the sizes of the apples hanging on a tree may be fairly similar to each other, so the “average size” of all apples gives a sometimes imprecise, but nevertheless informative and efficient, description of the sizes of individual apples. The studies on visual statistical properties (e.g., Alvarez, 2011; Chong & Treisman, 2003; Huang, 2015b) have shown that human observers can indeed efficiently extract the statistical properties of a set of objects, and more specifically have shown that these statistical regularities can be exploited to memorize more stimulus items (e.g., Brady, Konkle, & Alvarez, 2009; see also Sims, 2016; Sims, Jacobs, & Knill, 2012; but see Huang & Awh, 2018; Ngiam, Brissenden, & Awh, 2019).

Many previous studies have reported statistical-regularity-induced biases and attributed these biases to Bayesian inference (e.g., see Brady & Alvarez, 2011; Huang & Sekuler, 2010; Orhan & Jacobs, 2013; Son, Oh, Kang, & Chong, 2019). These studies found that the visual working memory

for an individual item is sometimes biased toward the “ensemble average” of the group the item belongs to. For example, Brady and Alvarez (2011) found that if an item belongs to a group whose members are generally larger than the item itself, then the report of the memorized size of the item is systematically biased toward being larger. These studies generally suggested that there are Bayesian inferences across different levels of representations.

Target bias versus strategic guesses

The abovementioned studies (e.g., Brady & Alvarez, 2011; Orhan & Jacobs, 2013) used the notion of Bayesian inference to account for the *bias* in the mnemonic representation of individual items. Briefly speaking, in a Bayesian inference mechanism, the visual system has a prior assumption about the distribution of the stimulus value and combines that with the observed data to produce a posterior distribution, and this combination introduces systematic bias in mental representations of the stimulus values. This notion of Bayesian inference has provided excellent accounts in various cases. For example, in the prototype effect (e.g., Bae, Olkkonen, Allred, & Flombaum, 2015; Huttenlocher, Hedges, & Duncan, 1991), biases are introduced by combining the prototype (i.e., prior distribution) and the actual feature value (i.e., data). If the abovementioned statistical-regularity-induced bias occurs as a result of Bayesian inference, then that means the individual features are genuinely memorized in a way that is systematically biased from the actual values. For example, if someone attempts to remember the sizes of 10 apples and is tested on a target apple that is smaller than the average of the set, the actual memory of the size of this tested apple is larger than it should be because it is biased toward the average. Hereinafter, this *individual-item-level bias in target-based responses* is addressed as the *target bias*.

Another possible mechanism is a *strategic guess*. Perhaps, an observer fails to remember the feature of a target item, but can use some strategies to make educated guesses. Generally speaking, various types of such guessing strategies (e.g., using a nontarget item as a substitute for the target item, reporting of the ensemble mean as a substitute of the target) will lead to “response biases” toward the ensemble mean. Hereinafter, these will be addressed as the *strategic guesses*. For a specific example, if someone attempts to remember the sizes of 10 apples and is tested on a target apple that is smaller than the average of the set, the observer may very well remember nothing about this particular target apple at all, but can nevertheless make an educated guess on the basis of what he or she knows about another apple. Then, on average, this guess would likely be larger than the actual size of the target apple, even if there is no individual-item bias in any of the mnemonic

representations. Hereinafter, this swap-based strategic guess is briefly addressed as the *swap*.

The target bias and swap are illustrated in Fig. 1. Assuming an ensemble of six values, target bias refers to the degree of how individual memory representations are pulled toward the center of the ensemble. On the other hand, swap responses could be made as substitutes for target responses, and that would also pull the responses toward the center of the ensemble. These two mechanisms have one thing in common: In terms of the mean of the distributions, the “output distributions” will be biased toward the ensemble mean. But, conceptually, they are two very different mechanisms.

Distinguishing between target biases and strategic guesses

To distinguish between the target bias and the strategic guess, the present study employs the three-part mixture model (e.g., Bays, Catalao, & Husain, 2009; Oberauer, Stoneking, Wabersich, & Lin, 2017), which was developed on the basis of Zhang and Luck’s (2008) two-part mixture model. In a typical trial of these studies, the observers were asked to memorize a set of colors and were then tested on one of them by choosing a color on a color wheel. In Zhang and Luck (2008),

Target bias vs. swap responses

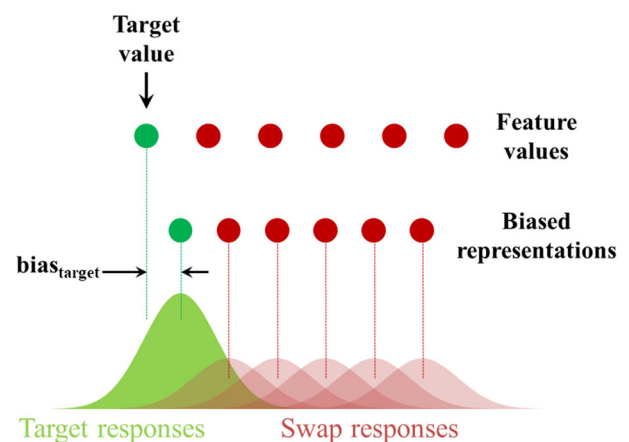


Fig. 1 Target-based responses and strategic guesses. Observers were asked to memorize six orientations (or six colors) that were evenly distributed over a certain range in the orientation space (or color wheel) and were tested on one of them. Reasonably, the target-based responses (i.e., responses based on knowledge about individual items) were expected to be normally distributed, centering on the target feature value plus a possible bias (green distribution). When observers know nothing about the target, they might make strategic guesses. For example, they might report another item from the ensemble that they know is likely to be similar to the target. Such swap responses were also expected to be normally distributed, centering on the other items (red distributions). (Color figure online)

knowledge-based responses were assumed to be normally distributed and centered on the target value, whereas random guesses were assumed to be evenly distributed on the entire color wheel. A mixture of the two distributions makes a family of distinctive distribution shapes that allows us to independently determine (1) the probability of memorizing an item and (2) the precision of the mnemonic representation. This mixture model has been widely applied to study various aspects of visual working memory (e.g., Hardman, Vergauwe, & Ricker, 2017; Murray, Nobre, Clark, Cravo, & Stokes, 2013; Ricker & Hardman, 2017; Zhang & Luck, 2008). Subsequent studies (e.g., Bays et al., 2009; Oberauer et al., 2017) added swap responses to Zhang and Luck's (2008) model and considered a mixture of three parts.

The present study adopted this three-part model and tried to independently determine the magnitudes of the target bias and strategic guess. Specifically, the magnitude of the target bias is reflected by the bias parameter in the target responses, whereas the magnitude of the strategic guess is reflected by the portion of the swaps.

As mentioned above, the swap is a plausible strategy of making educated guesses. Therefore, it is used as a convenient operational definition of the strategic guess in the present three-part model. It should be stated that these two can be different from each other. On the one hand, as mentioned above, one can easily imagine other strategies of making educated guesses, such as reporting the mean of the perceived ensemble. On the other hand, there are other potential reasons for swap responses to occur, such as perceptual binding errors. The potential implications of this divergence will be elaborated in the general discussion.

The present study

To summarize, the target biases and the strategic guess (i.e., operationalized as swaps) are two ways of using the ensemble information,¹ and they can be distinguished in a three-part model. Recently, many studies have reported target biases. Some of these studies (e.g., Brady & Alvarez, 2011; Orhan & Jacobs, 2013) have considered strategic guess as a contributing factor and have always concluded that it does not play a significant role.

Is it generally true that the target bias plays a dominant role, but the strategic guess plays no role in ensemble-induced bias? The present study attempted to readdress this question in broader conditions and in a more effective way. First, the present study adopted a greater memory load, manipulated

the range of ensemble (i.e., the similarity between members in an ensemble), and tried both color and orientation stimuli.

Second, the present study used a regular stimuli set which allows a direct visual analysis of the distribution of responses. In the present study, the orientations or colors were designed to be evenly distributed over a certain range in the featural space. These very regular stimuli were used for two reasons. First, this very clear-cut regularity will hopefully facilitate the use of ensemble information in the designed way so that the statistical-regularity-induced biases will be clearer and more pronounced. For example, if the ensemble consists of items that are unevenly distributed in the feature space, then it seems plausible that sometimes the actual “perceived ensemble” will deviate from the designed ensemble in some ways (e.g., skewed distribution, or forming two clusters) and these will add ambiguities on the predictions. Second, using regular stimuli, the predicted distribution of strategic guesses is known and remains constant from trial to trial (depending on where the target is in the range of features; see below), thus we can visually analyze the distribution of responses (see Fig. 4). This gives a more transparent analysis of the research question in addition to the statistical indexes.

Third, the present study used circular feature spaces (colors and orientations), which, when compared with linear feature spaces (e.g., sizes; see Brady & Alvarez, 2011), allow clearer interpretation of random guesses. On the one hand, there is a fairly straightforward way of modeling random guesses in circular spaces (i.e., uniform distribution in the full range). On the other hand, in linear spaces, the observer probably will inevitably consider the “usual range” of the stimuli when making a response. Therefore, when sizes are used, the response will appear to contain some information even if the observer knows nothing about the stimuli in a trial. Accordingly, all responses will be more or less affected, adding ambiguities on the interpretations of the responses.

In the present study, the observers were asked to memorize six orientations that were evenly distributed over a certain range in the orientation space (or six colors that were evenly distributed over a certain range in the color wheel) and were tested on one of them (see Fig. 3); their responses were then analyzed. Reasonably, the target-based responses (i.e., responses based on knowledge about individual items) were expected to be normally distributed, centering on the target feature value plus a possible bias (green distribution in Fig. 1); the swap responses were also expected to be normally distributed, centering on one of the other items (red distributions in Fig. 1).

Figure 2 illustrates how the target biases and the swap differently affect the distribution of responses. The overall responses (yellow curve) consist of target-based responses (green curve) and swap responses (red curve; i.e., the mixture of the red distributions in Fig. 1). In the two top examples, the target biases manifest as a shift in the average

¹ The target bias is clearly a Bayesian inference process. It should be mentioned that the strategic guess can also be viewed as a specific type of Bayesian inference process: There is no information at all in the data, so the posterior distribution is equal to the prior distribution. This point will be elaborated in the Discussion.

Mixture of Responses

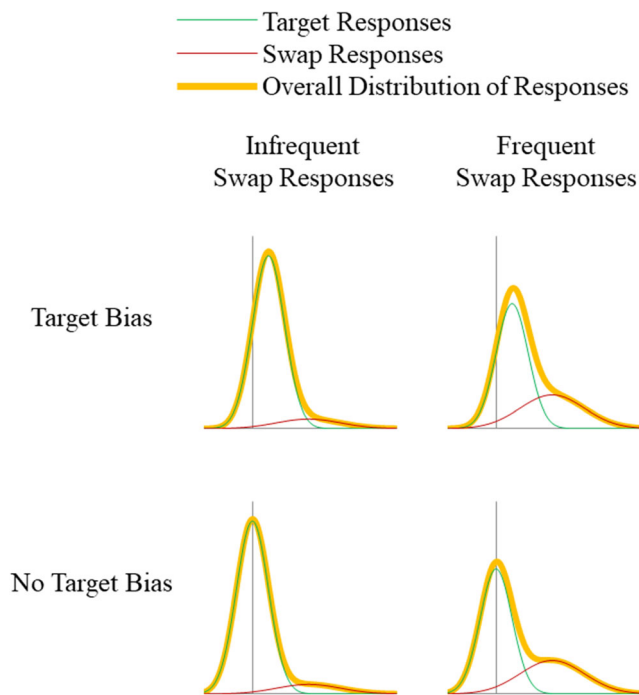


Fig. 2 How the target biases and strategic guesses differently affect the distribution of responses. The overall responses (yellow curve) consist of target-based responses (green curve) and strategic guesses (red curve, i.e., the mixture of the red distributions in Fig. 1). In the two top examples, the target biases manifest as a shift in the average of the distribution of target-based responses. In the two right-side examples, the strategic guesses manifest as the emergence of an additional distribution (i.e., the red curve, which is the mixture of the red distributions in Fig. 1). These two types of biases could certainly occur simultaneously (top-right corner). Then, the overall distribution of responses will vary in two distinctive ways when (a) there is a greater proportion of strategic guesses and (b) when there is a larger target bias. Specifically, although both will lead to a larger bias toward the ensemble average in the responses, the former manifests as a “fat tail” toward the ensemble average, whereas the latter manifests as a “shift of peak” toward the ensemble average. Then, both the proportion of swap responses and the magnitude of target bias can be simultaneously determined by fitting the experimental results to these predicted distributions. (Color figure online)

of the distribution of target-based responses. In the two right-side examples, the swap responses manifest as the emergence of an additional distribution centering on the ensemble average (i.e., the red curve). These two types of biases could occur with or without each other.

In Fig. 2, the overall distribution of responses will vary in two distinctive ways: (a) when there is a greater proportion of swap responses and (b) when there is a larger target bias. Specifically, although both will lead to a larger bias toward the ensemble average in the responses, the former manifests as a “fat tail” toward the ensemble average, whereas the latter manifests as a “shift of peak” toward the ensemble average. Then, both the proportion of swap responses and the magnitude

of target bias can be simultaneously determined by fitting the experimental results to these predicted distributions.

Method

Participants

A total of 176 participants (university students from the Chinese University of Hong Kong, average age = 20 years, 111 females), all of whom had a normal or corrected-to-normal vision, participated in this study. Each was paid HK\$50 for participating.

Participants were divided into eight separate groups (22 each), four groups for four different values of orientation ranges (75°, 100°, 125°, 150°; see below), and the other four groups for four different values of color ranges (75°, 100°, 125°, 150°; see below).

All experiments of the present study were carried out in accordance with approved guidelines. The consent form and experimental procedures received prior ethical approval from the research ethics committee of the Chinese University of Hong Kong. Informed consent was obtained from each participant.

Apparatus

In this study’s experiment, the participants viewed the display from a distance of about 60 cm. The participants were asked to make responses by clicking on a wheel to indicate the memorized orientation (or color) of the target. They were asked to respond as accurately as possible, but were under no time pressure (i.e., unspeeded responses).

Stimuli

Sample stimulus displays are shown in Fig. 3. For orientations, a stimulus display consisted of six isosceles triangles that were evenly placed around the center (on corners of a virtual regular hexagon) and 3.3 cm away from the center. For colors, a stimulus display consisted of six colored disks that were evenly placed around the center (on corners of a virtual regular hexagon) and 2.3 cm away from the center. The orientations of triangles (or colors of disks) were chosen so that they took six evenly distributed values over a certain range in the orientation space (or color wheel²). For example, in a trial, they could be 3°, 23°, 43°, 63°, 83°, and 103°. These orientations or colors were randomly assigned to the six items (i.e., locations).

² For colors, these degrees correspond to the degree on a color wheel, such as the one shown in Fig. 3.

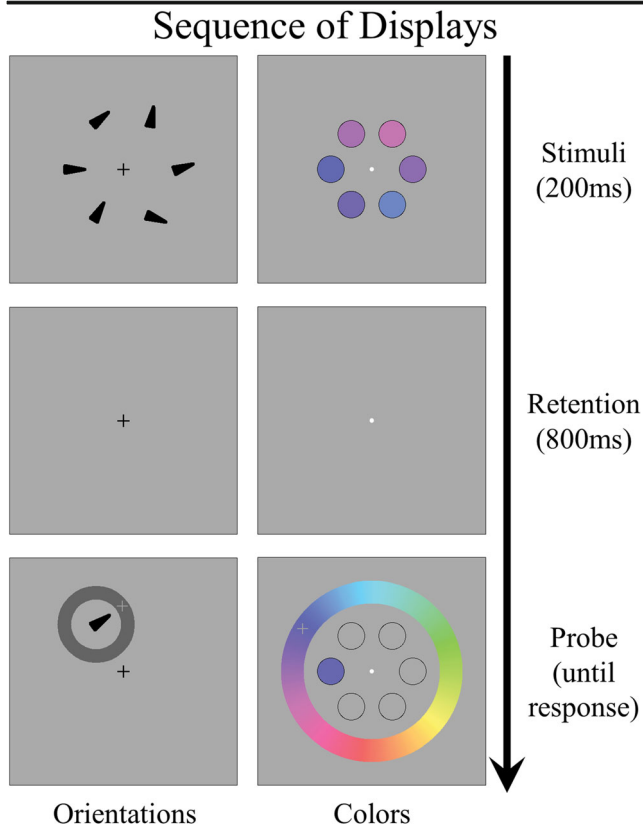


Fig. 3 Stimuli and procedure. A stimulus display consisted of six isosceles triangles (or six colored disks) that were evenly placed around the center. The orientations of these triangles (colors of disks) were evenly distributed values over a certain range in the orientation space (color wheel) and were randomly assigned to the six items. The range of the orientations (colors) could be 75°, 100°, 125°, or 150°, each for a separate group of 22 participants. The stimuli were presented for 200 ms and then disappeared. After a retention interval of 800 ms, one of the six items was randomly chosen for the memory test and was indicated by a probe. The participants then attempted to report, according to their memory, the orientation (color) of this item by choosing a position on the probe ring (or color wheel). (Color figure online)

The range of orientations (or colors) could be 75°, 100°, 125° or 150°. Each of the orientation range value was used in a separate group of 22 participants. Perceptually, an ensemble appeared fairly coherent when the orientation (or color) range was 75° (e.g., the items could be 3°, 18°, 33°, 48°, 63°, 78°), but appeared rather heterogeneous when the orientation (or color) range was 150° (e.g., the items could be 6°, 36°, 66°, 96°, 126°, 156°).

This manipulation on orientation (or color) range was introduced to see how target bias and swap responses are affected. The range manipulation is likely to be important because it directly determines the similarity between within-ensemble items, which plausibly affects the usefulness of the ensemble information. The orientations (or colors) were randomly chosen in the full 360° range, but can only vary as integers. For example, an orientation could be 341° or 342°, but can never be 341.5°.

Procedure

The procedure is illustrated in Fig. 3. A trial started with a fixation in the center of the display, which was presented for 800 ms and then followed by the stimulus display. The stimuli were presented for 200 ms and then disappeared. This stimulus duration of 200 ms was designed with reference to previous studies (e.g., Brady, Stormer, & Alvarez, 2016), which showed that simple features can be sufficiently encoded in 200 ms.

After a retention interval of 800 ms, one of the six items was randomly chosen for the memory test and was indicated by a probe (a dark-gray ring for reporting an orientation, or a global color wheel for reporting a color). The participants then attempted to report, according to their memory, the orientation or color of this item by choosing a position on the probe ring (or color wheel) using a mouse. The mouse cursor was set as a cross (rather than an arrow) for unambiguous localization. A “preview” of the chosen orientation (or chosen color) was shown and was continuously updated with mouse movement. After the participants were satisfied with a chosen response, they clicked the mouse to make a report. Each participant completed 10 blocks (56 trials per block). The first block was regarded as practice and excluded from the analysis.

Results

The results of the experiment are presented in Fig. 4. The data are presented as gray histograms, whereas the predictions (see below) are presented as black curves.

Following Zhang and Luck (2008), a *report-target difference* was calculated: The relative difference between the reported value and the target value, which falls within the range between -180° and 179° . The less the absolute value of this report-target difference is, the more precise a report is.

Figure 4 only plotted trials in which the target is an extreme value in the range of colors or orientations. For example, if, in a trial, the orientations were 147°, 167°, 187°, 207°, 227°, and 247°, then this trial is included only if the target is one of the “extremes values” of the range (147° and 247°)³. This is because the effects of both target biases and strategic guesses are expected to be more pronounced in these extreme values.⁴

³ In Fig. 4, the results are presented in the way that the target is less than the ensemble average. The data (i.e., report–target differences) on the other side (i.e., target greater than ensemble average) has been reversed ($\times -1$) for simpler presentation. In this example, the report–target differences for the targets 247° has been reversed and shown together with the target 147°.

⁴ For swap responses, the average swapped-induced difference (i.e., the difference between an item and the average of the other five items) was 60° for the extremes values of the range (147° and 247°), 36° for the intermediate values (167° and 227°), and 12° for the close-to-center values (187° and 207°). For target bias, the parameter $bias_{target}$ is used to represent the “magnitude of target bias” for the extreme values, and the magnitude of target bias should, respectively, be $0.6 \times bias_{target}$ for intermediate values, and $0.2 \times bias_{target}$ for close-to-center values. See main text for the details of modeling.

This restriction to the extreme values is only adopted for clearer presentation in Fig. 4, and all data were included for modeling below.

Implementation of the model

Platform, software, and type of the model

The data of the present experiment was modeled using matjags1.3 as a MATLAB interface for running JAGS (http://psiexp.ss.uci.edu/research/programs_data/jags/). The MATLAB script was run in MATLAB R2018b. The MATLAB script and the model specifications are available on the Open Science Framework project page (<https://osf.io/wscpv/>).

The four orientation ranges and four color ranges were modeled separately. For each there were 22 participants and 11,088 trials in total.

The model

Two parameters, $portion_{guess}$ and $portion_{swap}$ are used to describe respectively the “proportion of random guesses among all responses” and the “proportion of swap responses after excluding random guesses.”

Next, both the target-based responses and swap responses are assumed to be normally distributed, so one parameter (sd_{target}) is needed to describe the precision of an individual mnemonic representation (i.e., the standard deviation of the green distribution, or that of a red distribution in Fig. 1). In addition, one parameter ($bias_{target}$) is needed to describe the magnitude of target bias (see Fig. 1).

The specific model is as follows. The first part uses von Mises distributions⁵ (i.e., normal distribution in a circular space) to describe both the target-based responses and swap responses. The second part is the random guesses.

$$error_{response} \sim \sum_{item=0}^5 VM(mean_{item}, sd_{target}) \cdot portion_{item} + \frac{portion_{guess}}{2\pi}$$

The $portion_{item}$ (i.e., portion of contribution of an item) depends on whether it is the target or not. It is determined as

the following equation⁶:

$$If (item == item_{target}) \{ portion_{item} = (1 - portion_{guess}) \cdot (1 - portion_{swap}) \} \\ else \{ portion_{item} = (1 - portion_{guess}) \cdot (portion_{swap}) / 5 \}$$

The $mean_{item}$ (i.e., the difference between the “reported feature value of an item” and the “actual target value”) is determined as the following equation⁷:

$$mean_{item} = 0.2 \cdot range \cdot (item - item_{target}) + bias_{target} \cdot (1 - 0.4 \cdot item)$$

In this study, I adopted a Bayesian hierarchical model so that the between-subjects variations of the parameters can be appropriately incorporated. Then, for this purpose, each observer is assigned a different value on the four parameters. The distribution of individual subjects’ values of $bias_{target}$ was modeled by a normal distribution, whereas the distributions of individual subjects’ values of the other three parameters (sd_{target} , $portion_{guess}$, $portion_{swap}$) were modeled by lognormal distributions because they are always positive by definition.

The priors of the overall mean of these four parameters were chosen in the following ways. The priors of $portion_{guess}$ and $portion_{swap}$ were both uniform distribution [0, 1]. The prior of $bias_{target}$ was $N(\mu = 0, \lambda = 4)$: a moderate spread around zero. The prior of sd_{target} was LogNormal ($\mu = -1, \lambda = 0.001$): the mean is $\ln(22 \times \pi / 180) \approx -1$ because 22° is a typical sd in previous studies (e.g., Zhang & Luck, 2008), whereas the very small λ indicates uninformed prior distributions. The priors of the between-subjects variations (sds) of all the four parameters

⁶ Three points need to be clarified: (1) Both $item$ and $item_{target}$ mean not the physical location, but the relative position of an item in the range of features. For example, in a set of $[3^\circ, 23^\circ, 43^\circ, 63^\circ, 83^\circ, 103^\circ]$, $item = 0$ means 3° , $item = 1$ means 23° , and so forth. (2) A nontarget (i.e., swap) response is divided by 5 because it is assumed that one of the 5 nontarget items is randomly chosen. (3) $portion_{swap}$ is defined as the portion of swap responses after excluding random guesses. In other words, it is the portion of swap responses among “target-based responses and swap responses.” This parameter is defined this way so that $portion_{guess}$ and $portion_{swap}$ can vary as two separable parameters each in the full range of [0, 1].

⁷ The first part of this equation describes the difference caused by swapping in actual values, whereas the second part of this equation describes the target bias on the reported item. In the Bayesian integration of normal distributions, the mean of the posterior distribution is a weighted average of the mean of data and the mean of prior distribution: $mean_{posterior} = \frac{\lambda_{prior}}{\lambda_{prior} + \lambda_{data}} \cdot mean_{prior} + \frac{\lambda_{data}}{\lambda_{prior} + \lambda_{data}} \cdot mean_{data}$. Following this linear principle of Bayesian inference, the magnitude of the target bias should be proportional to the “target-ensemble distance.” Here, I do not make specific assumptions on the precisions of the prior information and data, but do assume they are the same regardless of where the target is in the range. In Fig. 1, these mean that all mnemonic representations are pulled toward the center of the ensemble as a proportional shrinking. Therefore, the parameter $bias_{target}$ is used to represent the magnitude of bias for $item = 0$, and the actual magnitude of bias in the other items are calculated as $bias_{target} \cdot (1 - 0.4 \cdot item)$, so that the bias is $-bias_{target}$ for $item = 5$.

⁵ The von Mises probability density function for a given angle x is $f(x | \mu, k) = e^{k \cos(x - \mu)} / (2\pi I_0(k))$, in which $I_0(k)$ is the modified Bessel function of order 0. The parameters μ and $1/k$ are analogous to μ and σ^2 (the mean and variance) in the normal distribution. The von Mises distribution is not provided in standard jags, so I used a module provided by Oberauer et al. (2017).

Best Fitting Models

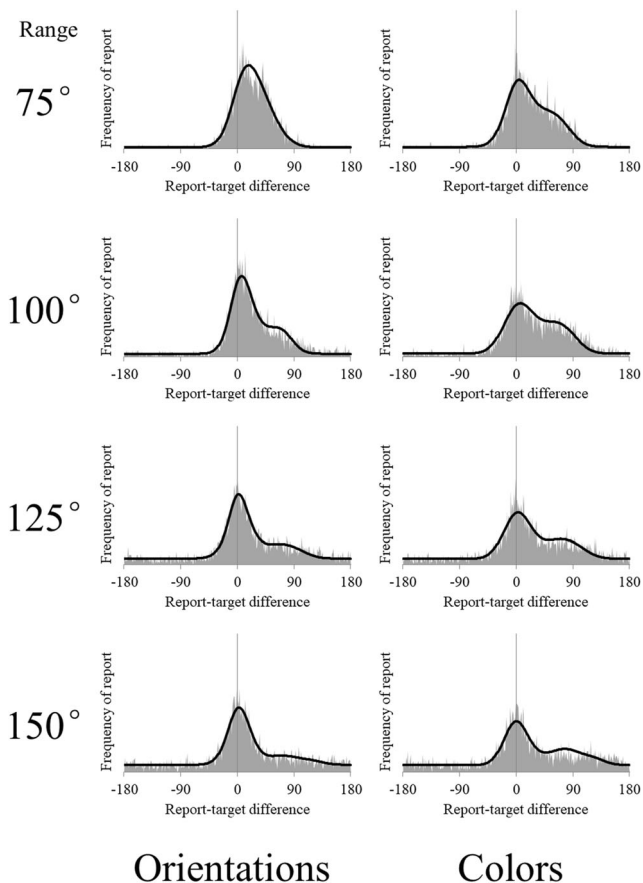


Fig. 4 The distributions of response-target differences and the best fitting models. The data are presented as gray histograms, whereas the predictions are presented as black curves. Only the extremes values are presented here because the effects are more pronounced. Clearly, the predictions provide an excellent fit to the overall trend of the data. Specifically, orientation ranges 75° and 100°, the black curves capture the shift of peaks in the data. More obviously, in all cases, the black curves capture the fat tails in the data

were LogNormal ($\mu = 0, \lambda = 0.001$) for the purpose of giving uninformed prior distributions.

Results of modeling

The JAGS gave posterior distributions of the mean of the population-level value of the four parameters ($portion_{guess}$, $portion_{swap}$, sd_{target} , $bias_{target}$). The mean and 95% confidence intervals (2.5% and 97.5% percentiles) are calculated from these distributions and plotted in Fig. 5. It is clear that all parameters can be determined fairly precisely.

As shown in Fig. 5a, $bias_{target}$ clearly differs from zero in two orientation ranges (75° and 100°), but not in the other two orientation ranges (125° and 150°) or any of the four color ranges. Therefore, target bias is statistically reliable in only two of eight situations.

As shown in Fig. 5b, sd_{target} is basically constant for all orientation ranges and constant for all color ranges. These

values fit very well with the typical sd values reported in previous studies (Bays et al., 2009; Zhang & Luck, 2008).

As shown in Fig. 5c, $portion_{guess}$ increases with the range of orientation, whereas $portion_{swap}$ decreases with the range of features. More critically for the present purpose, $portion_{swap}$ is always very far from zero, implying that swap is always an essential part of the responses.

To see whether the three-part model provides a sufficient fitting to the data, predicted distributions are plotted in Fig. 4

Parameter Estimation

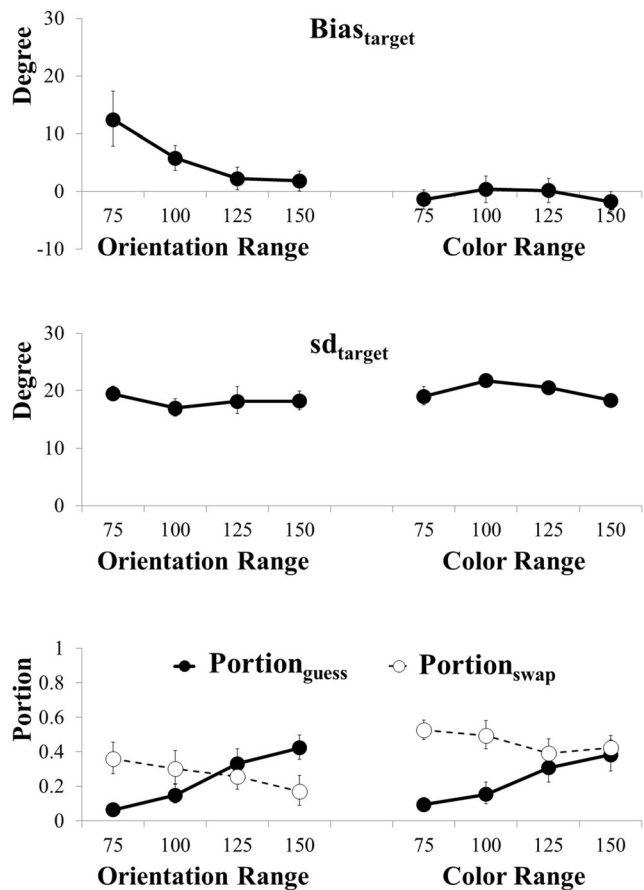


Fig. 5 Posterior distribution of parameters. The Bayesian modeling gave posterior distributions of the mean of the population-level value of the four parameters ($bias_{target}$, sd_{target} , $portion_{guess}$, $portion_{swap}$). The mean and 95% confidence intervals (2.5% & 97.5% percentiles) are calculated from these distributions and plotted. The $bias_{target}$ clearly differs from zero for two orientation ranges (75° & 100°), but not the other two orientation ranges (125° & 150°) or any of the four color ranges. Therefore, target bias is statistically reliable in some situations, but not in all situations. The sd_{target} is basically constant in all situations and this fits very well with the typical standard deviations values reported in previous studies (Bays et al., 2009; Zhang & Luck, 2008). The $portion_{guess}$ increases with the range of features, whereas $portion_{swap}$ decreases with the range of features. It seems that when a set of stimuli are less coherent, observers will be less likely to make strategic guesses and simply make a random response (see also Fig. 7)

as black curves.⁸ Clearly, they provide an excellent fit to the overall trend of the data. Specifically, in two cases (orientations 75° and 100°), the black curves capture the shift of peaks in the data. More obviously, in all cases, the black curves capture the fat tails in the data.

Parameter estimations and model comparisons are two different Bayesian approaches to the assessment of null values (Kruschke, 2011). To assess the necessity of the target biases and swaps, I considered three other models by removing target bias, swap responses, or both from the three-part model. DICs⁹ of these models are compared in Fig. 6. Basically, the no-swap model is always worse than the three-part model, and the no-target-bias-no-swap model is always worse than the no-target-bias model. Clearly, the swap is always an essential part of accounting for the pattern of responses.

On the other hand, the no-target-bias model is similar to, and sometimes even better than, the three-part model, whereas the no-target-bias-no-swap model is similar to, and sometimes even better than, the no-swap model. So, it seems fair to say that the target bias is not as important as the swap for accounting for the pattern of responses.

It should be pointed out that there is a disagreement between the results of parameter estimation and those of model comparisons in the orientation range 75° and 100°: The results of parameter estimation suggests a statistically reliable target bias, whereas results of model comparison suggest a preference for the no-bias model. It worth mentioning that the target biases are visually compelling in the orientation range 75° and 100° (see Fig. 4). Therefore, combining model comparison, parameter estimation, and visual assessment of the data, it seems reasonable to say that the results offer support for target biases in these cases.

Discussion

To explore the nature of statistical-regularity-induced bias in visual working memory, the present study attempted to predict the shapes of the distributions of responses when target-based responses are mixed with random guesses and swap responses. The predictions of this three-part model were then compared with the results of experiments in a Bayesian hierarchical model to determine the parameters of these distributions. This three-part model provides an excellent fit to the

⁸ These predicted distributions are generated by the three following steps. First, I took the mean of posterior distributions for each of the four parameters of each individual participant. Second, for each individual participant, the predicted distributions were generated by using these means as parameters. Third, these individual-participant distributions were averaged to generate the predicted aggregated distributions.

⁹ The DIC (deviance information criterion) is an estimator commonly used for model comparison in Bayesian hierarchical modeling. Conceptually, it is similar to AIC (Akaike information criterion) in nonhierarchical modeling.

Model Comparison

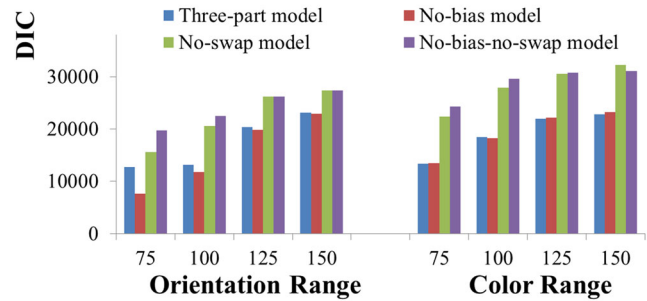


Fig. 6 Model comparison. To assess the necessity of the target biases and swaps, I considered three alternative models by removing target bias, swap responses, or both from the three-part model. DICs of the models are compared. Basically, the no-swap model is always worse than the three-part model, and the no-target-bias-no-swap model is always worse than the no-target-bias model. Clearly, strategic guess (swap) is always an essential part of accounting for the pattern of responses. (Color figure online)

data. In the present results, strategic guesses always contributed substantially to the statistical-regularity-induced biases, whereas target biases were limited to specific conditions. All in all, the Bayesian inference in visual working memory is much more limited than what is previously advocated.

Swap as strategic guesses

In the present study, the strategic guess is operationalized as swap responses. These two can be different from each other. Therefore, the potential implications of this divergence need to be elaborated.

Other strategies of making educated guesses

Aside from the swap, there are certainly other strategies of making educated guesses. For example, the observers could use the strategy of reporting the “perceived ensemble average” as a substitute for the target. Across many trials, the swap strategy would lead to a distribution that is essentially the same as the ensemble of all items. Therefore, in the present data, it is difficult to empirically distinguish swap responses and ensemble-average-based responses. Future studies will be needed to specifically target the distinction between these two (and perhaps other) types of strategic guesses. It should be clearly stated that for the present purpose of distinguishing target biases and strategic guesses, both the swap responses and ensemble-based responses are reasonable operational definitions of strategic guesses, and they serve the same conceptual purpose.

Alternative reasons for swap responses

One potential reason for swap responses is perceptual or mnemonic binding errors. As a binding error, the observers

do see (or remember) the item in a wrong location and report that item in the same confident way as they report correctly located items. As strategic guesses, the observers make speculations purposely. These two cannot be empirically distinguished from each other in the present data. However, it seems unlikely that binding errors play a large role. Previous studies have shown that from perceptual encoding (Johnston & Pashler, 1990) to working memory (Chen & Wyble, 2018; Jiang, Olson, & Chun, 2000; see also Pertzov & Husain, 2014; Rajsic & Wilson, 2014; Schneegans & Bays, 2017), the dimension of location is always the primary dimension that other features are based upon, so it is unlikely that one will often remember a color or an orientation in the wrong location. Recently, Pratte (2019) has explicitly tested this issue by presenting false-probe trials in which the probed color has not been actually presented. The responses, which are necessarily strategic guesses, are still only given to the locations of other items. In addition, the confidence ratings for the swap responses are low and comparable to random guesses. Altogether, Pratte (2019) suggested that swap responses are strategic guesses rather than real binding errors.

Recent studies (Oberauer & Lin, 2017; Schneegans & Bays, 2017) reported another potential reason for swap responses: the limited memory precision for the cue feature. However, these studies that have shown a serious problem on the cue precision usually involve more challenging situations (e.g., many locations or nonlocation feature cue). In the present study, the cue is manipulated as one out of six prototypical locations (i.e., corners of a virtual regular hexagon), so it is trivially easy to distinguish between these cued locations.

Even if these abovementioned factors do occasionally contribute to swap responses, there is no way they can account for the $portion_{swap}$, which can go up to as much as 50% in the present study. So they do not throw any serious doubt on the interpretations of the present results.

Strategic guess as a special case of Bayesian inference

It should be mentioned that the strategic guess concept can be viewed as a special case of Bayesian inference. If there is no observed data at all, then the posterior distribution is equal to the prior distribution. In other words, some types of prior knowledge (e.g., another known item) can be used as a report. Therefore, the strategic guess is not the opposite of the Bayesian inference, but can be viewed as a special case of the latter.

Nevertheless, the strategic guess is clearly only a special type of Bayesian inference in limited situations. On the one hand, if the Bayesian inference is a generally applicable mechanism, then the target bias should be observed in all situations (e.g., all the eight data sets of the present study). On the other hand, if one views the strategic guess (e.g., swap) as a special

type of Bayesian inference, then it is a much narrower solution that is only limited to the situation of “no observed data.” Clearly, if the Bayesian inference is mainly limited to the “no observed data” situation, then that is already consistent with the present conclusion, which is that the Bayesian inference in visual working memory is much more limited than what is previously advocated.

Contributions of target biases and strategic guesses

As mentioned above, the presence of strategic guesses (i.e., swap responses) is statistically very reliable in all eight situations, whereas the presence of target bias is statistically reliable in two of the eight situations. More generally, we want to distinguish the target biases and strategic guesses rather than merely testing against a specific null hypothesis. Therefore, it is useful to compare the contributions of these two types of biases to the overall statistical-regularity-induced bias.

In the sense of contributions to the weighted average, the target biases account for 19.8% and 10.5%, respectively, of overall biases for the orientation range of 75° and 100°, and approximately zero for the orientation range of 125° and 150° and all color ranges. Clearly, in the cases that have been tried in the present study, strategic guesses play a dominant role, whereas the target biases play a relatively small and inconsistent role.

Of course, this description of the dominance of strategic guesses is not meant to deny the importance of target biases in some cases. In Bayesian inference, it is useful to ask whether the integration mechanism has “fully used” the prior information. As mentioned above, in the design of the present study, we do not make specific assumptions on the sd of prior distributions. Nevertheless, an approximate estimation can be made. In the Bayesian integration of normal distributions, the mean of the posterior distribution is a weighted average of the mean of data and the mean of prior distribution:

$$\begin{aligned} mean_{posterior} = & \frac{\lambda_{prior}}{\lambda_{prior} + \lambda_{data}} \cdot mean_{prior} \\ & + \frac{\lambda_{data}}{\lambda_{prior} + \lambda_{data}} \cdot mean_{data} \end{aligned}$$

For the extreme values in orientation range 75°, the target-ensemble distance is $75^\circ/2 = 37.5^\circ$, and the bias is 12.4°. Therefore, we can estimate that the ratio between λ_{data} and λ_{prior} is approximately 2:1 (i.e., $sd_{prior}/sd_{data} = 1.44$). We do not know exactly what λ_{prior} is. But by any reasonable assumption, it has to be significantly worse than λ_{data} . Therefore, it is clear that in this case, the integration has fully used the prior information. In other words, for the case of orientation range 75°, the magnitude of target bias is as large as it theoretically can be.

As discussed above, many previous studies (e.g., Bae et al., 2015; Brady & Alvarez, 2011; Huang & Sekuler, 2010; Huttenlocher et al., 1991; Orhan & Jacobs, 2013; Sims et al., 2016) have referred to the notion of Bayesian inferences to account for biases in responses. Some of them have addressed the possibility of strategic guess, and they usually found that target bias is a dominant factor, whereas the strategic guess plays a negligible role (e.g., Brady & Alvarez, 2011; Orhan & Jacobs, 2013). It is important to consider the possible reasons why the results differ so much between the present study and these previous studies.

Factors affecting the target biases and strategic guesses

The relative contributions of target biases and strategic guesses apparently depend on the specific tasks and stimuli. Next, I will visit a few potential factors.

Memory load

An obviously critical factor is memory load. For example, in the classic prototype effect (Huttenlocher et al., 1991), observers only need to memorize one single location, and it seems reasonable to expect that all responses will be target based, and the prototype effect will depend mainly on target biases. Similarly, in other studies in which the memory load is limited to two or three items (Bae et al., 2015; Huang & Sekuler, 2010; Orhan & Jacobs, 2013; Sims et al., 2016), there was probably little or no need for making strategic guesses. Hypothetically, in orientation range 75° of the present study, if the memory load is reduced to two or three, then there will be no need of making strategic guesses, and the contribution of target biases will probably also increase dramatically from 20% to full dominance. In other words, the present finding is *not* incompatible with the results of these previous studies, but it does give a clear warning that the lack of strategic guesses cannot be generalized beyond the memory load of two or three.

Color versus orientation

In the present results, there is a clear target bias in some of the orientation ranges, but no target bias in any of the color ranges. Of course, it is always possible that there will be target biases in some other types of color ensemble.¹⁰ Nevertheless, previous studies showed color and orientation work differently for visual work memory in several other important ways, so it worth considering the present color/orientation distinction in this broader context.

¹⁰ Bae et al. (2015) demonstrated a prototype effect that is an effect of the innate color categories rather than an ensemble-induced color bias.

Huang (2015a; see also Alvarez & Cavanagh, 2008) showed that visual working memory for orientations is better than that of colors after the discriminability of stimulus items has been controlled in a perceptual discrimination task. Interestingly, this is not simply because the orientations are stronger visual features than colors are: Performance of orientations is actually much worse than that of colors in a visual search task. Huang (2015c) explained this difference by assuming that multiple orientations, but not multiple colors, can be represented together as a spatial structure that is formalized as a Boolean map (e.g., Huang, 2010a, 2010c; Huang & Pashler, 2007, 2009; Huang, Treisman, & Pashler, 2007). Huang (2015c) labeled this factor (being represented as a spatial structure) as spatial strength and examined its value for other featural dimensions. For a few examples, sizes are like orientations and have high spatial strength, whereas shapes are like colors and have low spatial strength.

Recently, Huang (2019) used this factor of spatial strength to account for a few findings in visual working memory. For one, memory for color–shape binding is largely integrated (see also Gajewski & Brockmole, 2006), whereas memory for color–orientation binding is largely independent of each other (see also Fougner, Cormiea, & Alvarez, 2013). For another, although color–color stimulus does not enjoy any same-object advantage (e.g., Xu, 2002), orientation–orientation stimulus does enjoy a substantial same-object advantage (e.g., Huang, 2019, Experiments 8–9).

In this context, it seems not entirely implausible that the root of the color/orientation distinction in the present study can also be attributed to their difference in spatial strength. In addition, the dimension size is frequently used to demonstrate the ensemble-induced biases (e.g., Brady & Alvarez, 2011; Corbett, 2017). In Huang's (2015c) measure, size also scores high on spatial strength, so this is again consistent with this hypothesis.

Then, why is it, exactly, that there are target biases only in high spatial strength features (e.g., orientation, size), but not in low spatial strength features (e.g., colors)? One straightforward account is that the colors are always represented individually, so they cannot influence the representations of each other, whereas orientations or sizes are represented together in multiple-item representations, so they will interact with each other in these multiple-item representations. Future studies will be needed to determine this and other possible mechanisms by which spatial strength affects target biases.

Effect of ranges

The feature range apparently affects the results. As shown in Fig. 5, for both colors and orientations, the $portion_{guess}$ increases with range, where the $portion_{swap}$ decreases with range. The portions of all three types of responses were calculated and presented in Fig. 7. It seems that the portion of target-based

Portions of the 3 types of responses

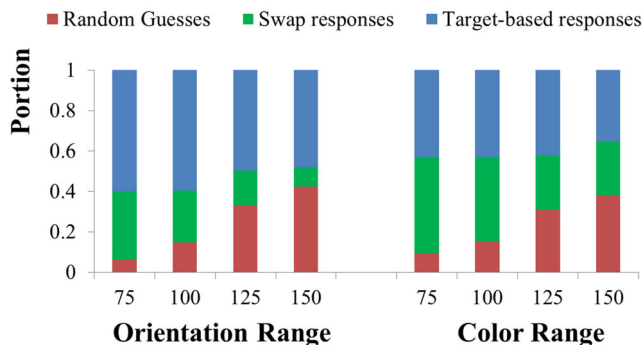


Fig. 7 The portions of the three types of responses were calculated. It seems that the portion of target-based responses is constant regardless of the range, whereas swap responses gradually turn into random responses. Perhaps when the members of a set are more different from each other, it becomes less useful to speculate the target based on another item in the ensemble, so observers are less inclined to make strategic guesses. (Color figure online)

responses is constant regardless of the range, whereas swap responses gradually turn into random responses. Perhaps, when the members of a set are more different from each other, it becomes less useful to guess the target based on another item in the ensemble, so observers are less inclined to make strategic guesses. Similarly, this perhaps also explains why target biases of orientation stimuli disappear in larger orientation ranges.

These effects of ranges are consistent with previous studies showing that stimuli variance is very important in ensemble perception (Corbett, Wurmitsch, Schwartz, & Whitney, 2012; Fouriez, Rubinfeld, & Capstick, 2008; Im & Halberda, 2013; Morgan, Chubb, & Solomon, 2008; Solomon, Morgan, & Chubb, 2011). In this specific case, these effects of ranges are also consistent with what we would have intuitively expected. For a more dissimilar ensemble, the members in the ensemble become less effective substitutes of each other. In addition, a more dissimilar ensemble spread over multiple feature categories, and it seems plausible that same-category features (e.g., approximately upward-pointing arrows) are especially likely to be used as substitutes of each other. For these reasons, the underlying mechanisms (of both Bayesian integration and strategic guesses) are less inclined to exploit ensemble information for more dissimilar ensembles.

Other potential factors

The present three-part model has not considered some of the known factors. First, the present study has not considered the prototype effects (i.e., biased toward the vertical/horizontal orientations, or biases toward the center of color categories; see Bae et al., 2015). Second, by keeping the parameters sd_{target} constant for all trials of an observer, the present three-part model has not considered variability in how accurately

observers encode the individual item or the ensemble mean (see Fougny, Suchow, & Alvarez, 2012). Although these factors certainly do exist in the present task, they are not directly relevant to the present purpose, because their effects would be averaged out across the trials and they are unlikely to account for the specific distribution of responses (shift of peaks and/or fat tails) shown in Fig. 4.

Funding The work described in this paper was fully supported by two grants from the Research Grants Council of the Hong Kong Special Administrative Region, China (CUHK 14617615, CUHK 14128016). The writing of this paper was substantially helped by the Humanities and Social Sciences Prestigious Fellowship (CUHK 34000817), also from the Research Grants Council. Data and script for data analysis are available on the Open Science Framework project page (<https://osf.io/wscpv/>). I am grateful to Tim Ricker, Sang Chul Chong, and one anonymous reviewer for very helpful comments.

References

- Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, 15, 122–131. doi:<https://doi.org/10.1016/j.tics.2011.01.003>
- Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, 15, 106–111. doi:<https://doi.org/10.1111/j.0963-7214.2004.01502006.x>
- Alvarez, G. A., & Cavanagh, P. (2008). Visual short-term memory operates more efficiently on boundary features than on surface features. *Perception & Psychophysics*, 70, 346–364. doi:<https://doi.org/10.3758/PP.70.2.346>
- Bae, G.-Y., Olkkonen, M., Allred, S. R., & Flombaum, J. I. (2015). Why some colors appear more memorable than others: A model combining categories and particulars in color working memory. *Journal of Experimental Psychology: General*, 144, 744. doi:<https://doi.org/10.1037/xge0000076>
- Bays, P. M., Catalao, R. F., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, 9, 7–7. doi:<https://doi.org/10.1167/9.10.7>
- Brady, T. F., & Alvarez, G. A. (2011). Hierarchical encoding in visual working memory: Ensemble statistics bias memory for individual items. *Psychological Science*, 22, 384–392. doi:<https://doi.org/10.1177/0956797610397956>
- Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in Visual Working Memory: Using Statistical Regularities to Form More Efficient Memory Representations. *Journal of Experimental Psychology-General*, 138, 487–502. doi:<https://doi.org/10.1037/a0016797>
- Brady, T. F., Stormer, V. S., & Alvarez, G. A. (2016). Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli. *Proceedings of the National Academy of Sciences*, 113, 7459–7464
- Chen, H., & Wyble, B. (2018). The neglected contribution of memory encoding in spatial cueing: A new theory of costs and benefits. *Psychological Review*, 125, 936. doi:<https://doi.org/10.1037/rev0000116>
- Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, 43, 393–404. doi:[https://doi.org/10.1016/s0042-6989\(02\)00596-5](https://doi.org/10.1016/s0042-6989(02)00596-5)
- Corbett, J. E. (2017). The whole warps the sum of its parts: Gestalt-defined-group mean size biases memory for individual objects.

- Psychological Science*, 28, 12–22. doi:<https://doi.org/10.1177/0956797616671524>
- Corbett, J. E., Wumitsch, N., Schwartz, A., & Whitney, D. (2012). An aftereffect of adaptation to mean size. *Visual Cognition*, 20, 211–231. doi:<https://doi.org/10.1080/13506285.2012.657261>
- Fougnie, D., Cormiea, S. M., & Alvarez, G. A. (2013). Object-based benefits without object-based representations. *Journal of Experimental Psychology: General*, 142, 621. doi:<https://doi.org/10.1037/a0030300>
- Fougnie, D., Suchow, J., & Alvarez, G. (2012). Variability in the quality of visual working memory. *Nature Communications*, 3, 1229. doi:<https://doi.org/10.1038/ncomms2237>
- Fouriez, G., Rubenfeld, S., & Capstick, G. (2008). Visual statistical decisions. *Perception & Psychophysics*, 70, 456–464. doi:<https://doi.org/10.3758/pp.70.3.456>
- Gajewski, D. A., & Brockmole, J. R. (2006). Feature bindings endure without attention: Evidence from an explicit recall task. *Psychonomic Bulletin & Review*, 13, 581–587. doi:<https://doi.org/10.3758/Bf03193966>
- Hardman, K. O., Vergauwe, E., & Ricker, T. J. (2017). Categorical working memory representations are used in delayed estimation of continuous colors. *Journal of Experimental Psychology: Human Perception and Performance*, 43, 30. doi:<https://doi.org/10.1037/xhp0000290>
- Huang, L. (2010a). Characterizing the nature of visual conscious access: The distinction between features and locations. *Journal of Vision*, 10(10): 24, 1–17. doi:<https://doi.org/10.1167/10.10.24>
- Huang, L. (2010b). Visual working memory is better characterized as a distributed resource rather than discrete slots. *Journal of Vision*, 10(14): 1–8. doi:<https://doi.org/10.1167/10.10.24>
- Huang, L. (2010c). What is the unit of visual attention? Object for selection, but Boolean map for access. *Journal of Experimental Psychology—General*, 139, 162–179. doi:<https://doi.org/10.1037/a0018034>
- Huang, L. (2015a). Color is processed less efficiently than orientation in change detection but more efficiently in visual search. *Psychological Science*, 26, 646–652. doi:<https://doi.org/10.1177/0956797615569577>
- Huang, L. (2015b). Statistical properties demand as much attention as object features. *PLOS ONE*, 10, e0131191. doi:<https://doi.org/10.1371/journal.pone.0131191>
- Huang, L. (2015c). Visual features for perception, attention, and working memory: Toward a three-factor framework. *Cognition*, 145, 43–52. doi:<https://doi.org/10.1016/j.cognition.2015.08.007>
- Huang, L. (2019). Unit of visual working memory: A Boolean map provides a better account than an object does. *Journal of Experimental Psychology—General*. Advance online publication. doi:<https://doi.org/10.1037/xge0000616>
- Huang, L., & Awh, E. (2018). Chunking in working memory via content-free labels. *Scientific Reports*, 8, 23. doi:<https://doi.org/10.1038/s41598-017-18157-5>
- Huang, L., & Pashler, H. (2007). A Boolean map theory of visual attention. *Psychological Review*, 114, 599–631. doi:<https://doi.org/10.1037/0033-295x.114.3.599>
- Huang, L., & Pashler, H. (2009). Reversing the attention effect in figure-ground perception. *Psychological Science*, 20, 1199–1201. doi:<https://doi.org/10.1111/j.1467-9280.2009.02424.x>
- Huang, J., & Sekuler, R. (2010). Distortions in recall from visual memory: Two classes of attractors at work. *Journal of Vision*, 10, 24–24. doi:<https://doi.org/10.1167/10.2.24>
- Huang, L., Treisman, A., & Pashler, H. (2007). Characterizing the limits of human visual awareness. *Science*, 317, 823–825. doi:<https://doi.org/10.1126/science.1143515>
- Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review*, 98, 352. doi:<https://doi.org/10.1037/0033-295X.98.3.352>
- Im, H. Y., & Halberda, J. (2013). The effects of sampling and internal noise on the representation of ensemble average size. *Attention, Perception, & Psychophysics*, 75, 278–286. doi:<https://doi.org/10.3758/s13414-012-0399-4>
- Jiang, Y. H., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, 26, 683–702. doi:<https://doi.org/10.1037/0278-7393.26.3.683>
- Johnston, J. C., & Pashler, H. (1990). Close binding of identity and location in visual feature perception. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 843–856. doi:<https://doi.org/10.1037/0096-1523.16.4.843>
- Kruschke, J. K. (2011). Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science*, 6, 299–312. doi:<https://doi.org/10.1177/1745691611406925>
- Morgan, M., Chubb, C., & Solomon, J. A. (2008). A ‘dipper’ function for texture discrimination based on orientation variance. *Journal of Vision*, 8, 9. doi:<https://doi.org/10.1167/8.11.9>
- Ngiam, W. X., Brissenden, J. A., & Awh, E. (2019). “Memory compression” effects in visual working memory are contingent on explicit long-term memory. *Journal of Experimental Psychology: General*, 148, 1373. doi:<https://doi.org/10.1037/xge0000649>
- Oberauer, K., & Lin, H.-Y. (2017). An interference model of visual working memory. *Psychological Review*, 124, 21. doi:<https://doi.org/10.1037/rev0000044>
- Oberauer, K., Stoneking, C., Wabersich, D., & Lin, H.-Y. (2017). Hierarchical Bayesian measurement models for continuous reproduction of visual features from working memory. *Journal of Vision*, 17, 11–11. doi:<https://doi.org/10.1167/17.5.11>
- Olson, I. R., & Jiang, Y. H. (2002). Is visual short-term memory object based? Rejection of the “strong-object” hypothesis. *Perception & Psychophysics*, 64, 1055–1067. doi:<https://doi.org/10.3758/BF03194756>
- Orhan, A. E., & Jacobs, R. A. (2013). A probabilistic clustering theory of the organization of visual short-term memory. *Psychological Review*, 120, 297. doi:<https://doi.org/10.1037/a0031541>
- Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, 44, 369–378. doi:<https://doi.org/10.3758/bf03210419>
- Pertsov, Y., & Husain, M. (2014). The privileged role of location in visual working memory. *Attention, Perception, & Psychophysics*, 76, 1914–1924. doi:<https://doi.org/10.3758/s13414-013-0541-y>
- Phillips, W. A. (1974). Distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, 16, 283–290. doi:<https://doi.org/10.3758/Bf03203943>
- Pratte, M. S. (2019). Swap errors in spatial working memory are guesses. *Psychonomic Bulletin & Review*, 26, 958–966. doi:<https://doi.org/10.3758/s13423-018-1524-8>
- Rajsic, J., & Wilson, D. E. (2014). Asymmetrical access to color and location in visual working memory. *Attention, Perception, & Psychophysics*, 76, 1902–1913. doi:<https://doi.org/10.3758/s13414-014-0723-2>
- Ricker, T. J., & Hardman, K. O. (2017). The nature of short-term consolidation in visual working memory. *Journal of Experimental Psychology: General*, 146, 1551. doi:<https://doi.org/10.1037/xge0000346>
- Schneegans, S., & Bays, P. M. (2017). Neural architecture for feature binding in visual working memory. *Journal of Neuroscience*, 37, 3913–3925. doi:<https://doi.org/10.1523/JNEUROSCI.3493-16.2017>
- Sims, C. R. (2016). Rate distortion theory and human perception. *Cognition*, 152, 181–198. doi:<https://doi.org/10.1016/j.cognition.2016.03.020>

- Sims, C. R., Jacobs, R. A., & Knill, D. C. (2012). An ideal observer analysis of visual working memory. *Psychological Review*, *119*, 807. doi:<https://doi.org/10.1037/a0029856>
- Solomon, J. A., Morgan, M., & Chubb, C. (2011). Efficiencies for the statistics of size discrimination. *Journal of Vision*, *11*, 13–13. doi:<https://doi.org/10.1167/11.12.13>
- Son, G., Oh, B.-I., Kang, M.-S., & Chong, S. C. (2019). Similarity-based clusters are representational units of visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. doi:<https://doi.org/10.1037/xlm0000722>
- Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131*, 48–64. doi:<https://doi.org/10.1037/0096-3445.131.1.48>
- Xu, Y. D. (2002). Limitations of object-based feature encoding in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 458–468. doi:<https://doi.org/10.1037/0096-1523.28.2.458>
- Zhang, W. W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*, U233–U213. doi:<https://doi.org/10.1038/nature06860>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.