The nature of the perceptual representation for decision making

Jiwon Yeon and Dobromir Rahnev

School of Psychology, Georgia Institute of Technology, Atlanta, GA

Manuscript word count: 5,954 words (38,305 characters)

Keywords: Perceptual decision making; population code; sensory representation; decision representation

Author contributions

The two authors designed the study. J. Yeon collected and analyzed data. Both authors worked together to interpret the result and write the manuscript and approved the final version of the manuscript for submission.

Acknowledgements

We thank Tristan Hackman and Megan Kelley for help with the data collection. This work was funded by a startup grant to D.R. from the Georgia Institute of Technology.

Corresponding authors

Jiwon Yeon

Georgia Institute of Technology

654 Cherry Str. NW

Atlanta, GA 30332

E-mail: j.yeon@gatech.edu

Dobromir Rahnev

Georgia Institute of Technology

654 Cherry Str. NW

Atlanta, GA 30332

E-mail: rahnev@psych.gatech.edu

Abstract

Electrophysiological recordings show that any visual stimulus produces a response in a large population of differently tuned sensory neurons. However, it remains unclear how this population response is used to make perceptual decisions. Major theories, such as probabilistic population codes and neural sampling, assume that the population response can be flexibly used for decision making. However, another possibility is that decision-making circuits do not have access to the sensory population code but only to a summary of this code. These possibilities can be disentangled for choices among n>2 discrete stimulus categories. In two experiments, we asked subjects to choose between n=4 colors or n=6 symbols. We then used the n-alternative condition to predict the performance on a different, 2-alternative condition where only two colors or symbols were given as available alternatives on each trial. A population model assuming that decision-making circuits have access to the whole distribution of the sensory activation levels across all colors or symbols significantly overestimated the performance in the 2-alternative condition. Instead, a summary model assuming that decision-making circuits only have access to the sensory activity associated with the dominant color predicted the 2-alternative condition very well and was preferred in Bayesian model comparison. Finally, a third experiment asked subjects to make two decisions in a row and also found the summary model to provide a better description of the data. These results show that the full population code in sensory cortex may not be available for deliberate decision making, at least within the context of decisions between discrete alternatives.

Introduction

Perception has been conceptualized as a decision process for over a century and a half (Helmholtz, 1856). According to this view, the outside world is encoded in a pattern of neural firing and the brain needs to decide what these patterns signify. Although this view is widely accepted, it remains unclear what type of information perceptual decisions are actually based on.

Electrophysiological studies starting with Hubel and Wiesel (1970) have revealed how individual neurons in various sensory areas of cortex respond to different stimuli. Such studies have demonstrated that a single stimulus gives a rise to activity in a whole population of differently-tuned neurons in sensory cortex. This population response must therefore form the foundation of the information representations on which the brain makes decisions.

In fact, two prominent theories – probabilistic population codes and neural sampling – postulate that deliberate decision making is performed directly on these sensory distributions. Probabilistic population codes (Zemel et al., 1998; Pouget et al., 2000, 2003) treat the population activity as representing uncertainty about stimuli in the form of probability distributions. These probability distributions can then be used to perform various computations such as cue combination (Ma et al., 2006). Neural sampling conceptualizes of neural activity in sensory cortex as samples from a posterior distribution over external variables (Fiser et al., 2010; Berkes et al., 2011; Haefner et al., 2016; Orbán et al., 2016). The posterior distribution is then used directly to make perceptual decisions.

However, an alternative possibility is that only a summary of the population code in sensory cortex is available for deliberate decision making. According to this view, the whole population code may be used in certain circumstances such as mandatory cue combination (Hillis et al., 2002) but not be available for novel, deliberate decisions. Such views are implicitly endorsed by traditional theories such as signal detection theory (Green and Swets, 1966) and the drift diffusion model (Ratcliff, 1978), which assume that, in 2-alternative tasks, the population code is summarized down to a single number that is subsequently used for decision making.

These two views – decision making being based on the whole population code in sensory cortex vs. only on a summary of this population code – are difficult to disentangle experimentally. The challenge comes both because it is unclear what using the “whole” population code means in practice and because there are many ways to extract a “summary” of the population code. An even bigger challenge is that most research to date has focused on judging features represented on a continuous scale such as orientation or motion direction (Ma et al., 2006; Haefner et al., 2016), which are typically assumed to give rise to a Gaussian distribution of neural activity. Such Gaussian distribution can be represented with no loss of information with just two numbers – its mean and variability – thus blurring the distinction between making decisions based on the whole population code or a summary of it **(Figure 1A)**.

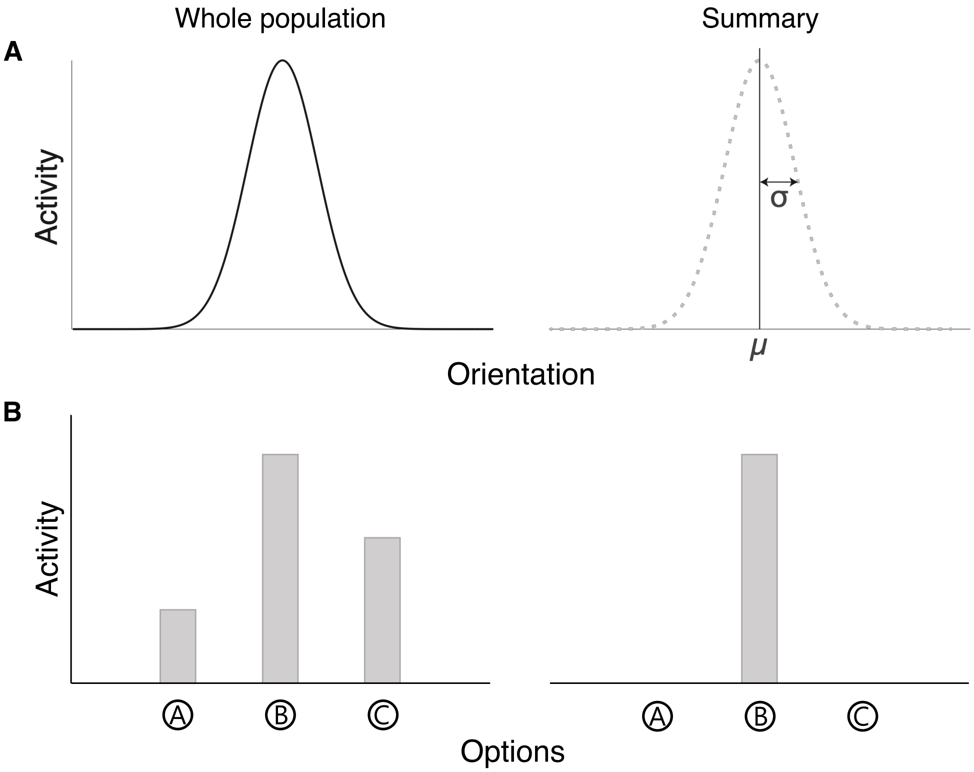

Figure 1. Decision making about continuous vs. discrete stimuli. (A) Decision making for continuous-scale stimuli. Stimulus features represented on a continuous scale, such as orientation, are typically assumed to give rise to a Gaussian distribution of neural activity in sensory cortex. For such stimuli, the sensory population response (left) can be summarized without any loss of information by only considering the distribution's mean, μ, and standard deviation, σ (right). These differing but equivalent in their informational content representations make it difficult to determine whether decision-making circuits have access to the full sensory population code vs. a summary of it. (B) Decision making for discrete alternatives. In cases where a stimulus can come from several discrete alternatives (e.g., options A, B, and C), a stimulus can be assumed to give rise to a different amount of sensory cortex activity associated with each alternative (left panel). A summary of this distribution will likely involve information

loss. One possible summary representation consists of passing only the highest activity onto decision-making circuits (right panel). This information loss will become apparent if subjects have to choose between the other alternatives (e.g., alternatives A and C). Therefore, it is possible to adjudicate between decision-making circuits having access to the whole sensory population code vs. a summary of that code.

However, determining whether decisions are based on the whole sensory distribution or a summary of it becomes tractable in cases where subjects decide between discrete stimulus categories. In such cases, the underlying sensory response can be simplified to a distribution of the evidence available for each discrete stimulus category. At the same time, any summary of this distribution will involve significant information loss, thus making it possible in principle to adjudicate between the two possibilities (**Figure 1B**).

Here we used such discrete stimulus categories in three different experiments. All experiments featured a condition where subjects picked the most frequently presented stimulus among all of the possible stimulus categories (four different colors in Experiment 1 and six different symbols in Experiments 2 and 3). Based on these responses, we estimated the parameters of a model describing subjects’ internal distribution of sensory responses. We then included conditions where subjects were told to pick between only two alternatives after the offset of the stimulus (Experiments 1 and 2) or to make a second choice if the first one was incorrect (Experiment 3). These conditions allowed us to compare different models of how the sensory population code was actually used for decision making. To anticipate, we found robust evidence that decisions in our experiments were based on a summary of the population code rather than the whole distribution of activity for each stimulus category. These results indicate that perceptual decision-making circuits may not have access to the full population code and that significant amount of simplification is likely to occur before information is used for deliberate decisions.

Results

In order to determine whether perceptual decision making uses the whole sensory population code or only a summary of it, we performed three experiments in which subjects made choices about discrete stimulus categories.

Experiment 1 required subjects to pick which of four possible colors – blue, red, green, and white – was most frequently presented (**Figure 2**). The stimulus consisted of 49 colored circles arranged in a 7x7 square presented for 500 ms. On each trial, one color was randomly chosen to be "dominant" and 16 circles were painted in that color, whereas the remaining three colors were "non-dominant" and 11 circles were painted in each of those colors. The experiment featured two different condition. In the 4-alternative condition, subjects picked the dominant color among the four possible colors. In the 2-alternative condition, after the stimulus presentation, subjects were asked to choose between the dominant and one randomly-chosen non-dominant color. Note that in both conditions, the subjects' task was always to correctly identify the dominant color.

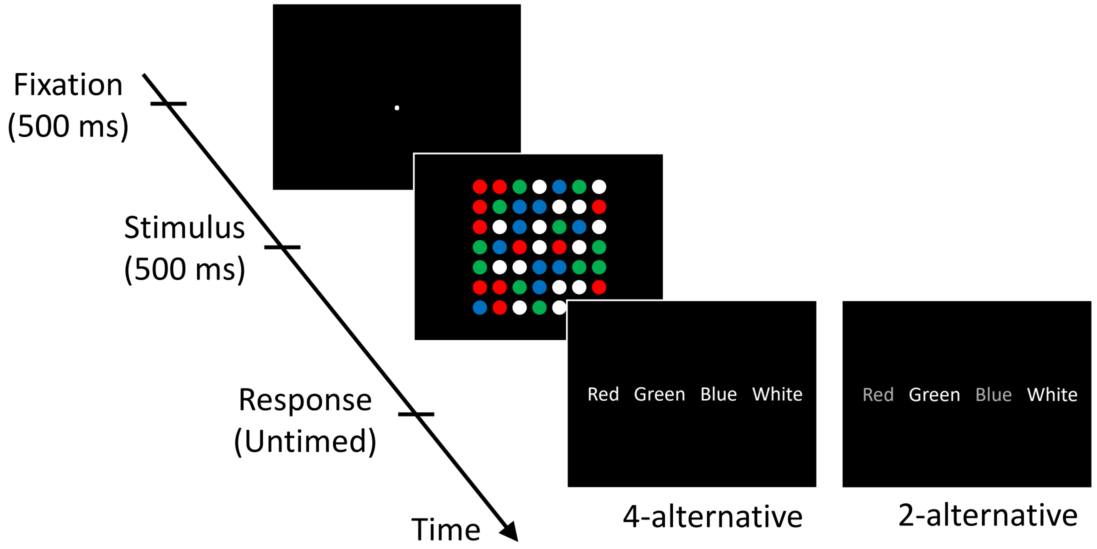

Figure 2. Task for Experiment 1. Each trial consisted of a fixation period (500 ms), stimulus presentation (500 ms), and untimed response period. The stimulus comprised of four different colored circles (red, green, blue, and white). One of the colors (white in this example) was presented more frequently (16 circles; dominant color) than the other colors (11 circles each; non-dominant colors). Subjects' task was to indicate the dominant color. Two conditions were presented in different blocks. In the 4-alternative condition, subjects chose between all four colors. In a separate 2-alternative condition, on each trial subjects were given a choice between the dominant and one randomly chosen non-dominant color.

Using subjects' responses in the 4-alternative condition, we estimated the parameters of the sensory distribution representing the activity level for each color. For these computations, we made only basic assumptions regarding the between-trial variability of the activity level, which was drawn from a Gaussian distribution with a standard deviation of 1. The estimated parameters could then be used to simulate the single-trial activation levels for each color.

Using the parameters estimated from the 4-alternative condition, we considered the predictions for the 2-alternative condition for two different models: (1) a "population model," according to which perceptual decisions are based on the whole distribution of activities over the four colors, and (2) a "summary model," according to which perceptual decisions are based on a summary of the whole distribution. There are a number of ways to create a summary of the distribution. However, in the context of this task, the only relevant information is the order of activation levels from highest to lowest (this order determines how a subject would pick different colors as the dominant color in the 2-alternative condition). Other information, such as average activity level, is irrelevant to the task here. We first considered an extreme summary model that consists of the activity level for the one color with highest level of activity. Other summary models in which decision-making circuits have access to the activity levels of the first n>1 largest activations are examined later.

The population and summary models could be easily compared because they make different predictions about a subject's performance in the 2-alternative task. To explain the reason behind the differing model predictions, it is helpful to first consider when the two models make the same predictions. First, the models make the same prediction when the dominant color gives rise to the highest activity level. In such cases, that color would be selected according to both the population and summary models resulting in a correct choice (**Figure 3A**). Second, the models also make the same prediction when the alternative option given to the subject happens to have the highest activity. In such cases, the alternative option would be selected according to both the population and summary models resulting in an incorrect choice (**Figure 3B**).

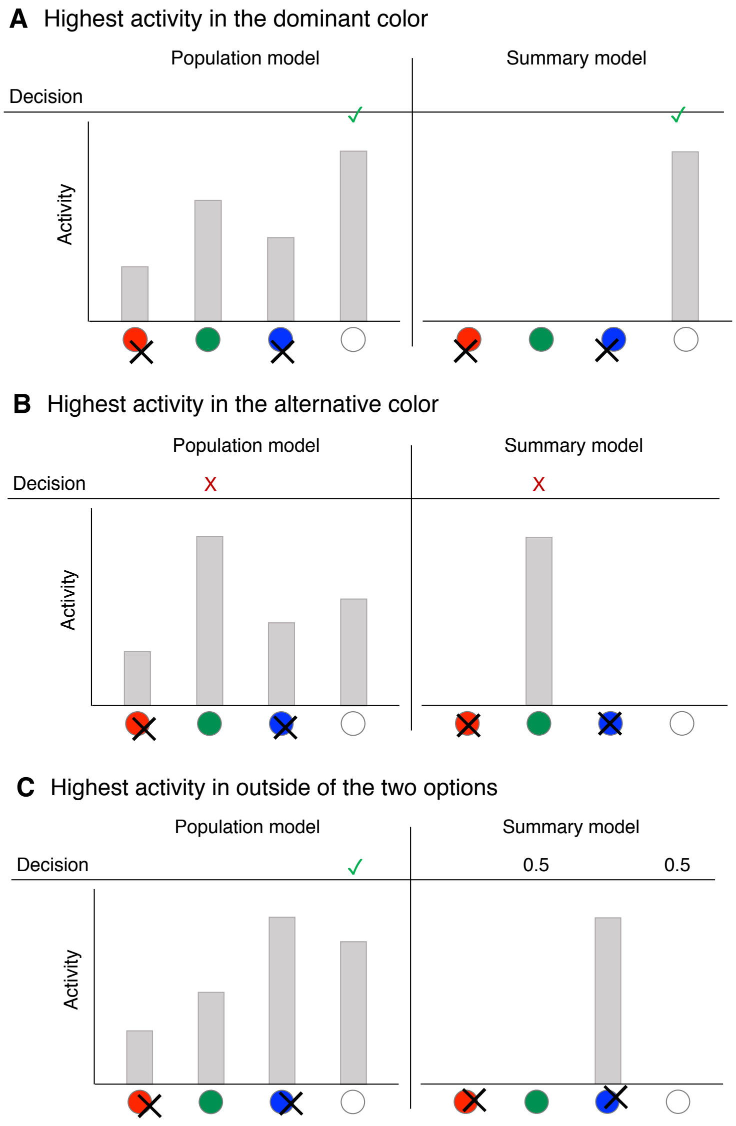

Figure 3. Predictions of the population and summary models for subjects' choices in the 2-alternative condition. The population model (left panels) assumes that decision-making circuits

have access to the activity levels associated with each of the four colors (four gray bars), whereas the summary model (right panels) assumes that decision-making circuits only have access to the highest activity level (single gray bar). In all examples, the dominant circle is white and subjects are given a choice between white and green. (A) When the highest activity happens to be at the dominant color, both models predict that the subject would correctly choose the dominant color. (B) When the highest activity happens to be for the alternative color, both models predict that the subject would incorrectly choose the alternative color. (C) The two models' prediction diverge when the highest activity is associated with a color other than the two presented alternatives. In such cases, the activation for the dominant color is likely to be higher than for the alternative color, so according to the population model, subjects would ignore the color with the highest activity (red color in the example here) and correctly pick the dominant color in the majority of the trials. However, according to the summary model, subjects have no information about the activation levels for the dominant and alternative colors and would thus correctly pick the dominant color on only 50% of such trials.

The population and summary models diverge in their behavior when the highest activity is associated with a color that is not among the two options (Figure 3C). In such cases, the summary model postulates that subject have no activity-related information on which to base their decision and thus must randomly select between the two colors options. It is clear that such cases would result in the summary model having 50% accuracy rate (chance level). At the same time, the population model postulates that subjects would select the color with higher activity between the two options. The accuracy rate of the population model would thus depend on the probability of having higher activity in the dominant color than in the alternative color. Given that the dominant color typically produces higher activity level than the alternative option, the accuracy rate in such trials would be significantly higher than 50%. Therefore, for these types of trials, the population and summary models make different predictions with the population model predicting a higher performance level.

The intuition above was borne out in the actual model predictions produced by the population and summary models. Indeed, based on the performance in the 4-alternative condition (average accuracy = 69.2%), the population and summary models predicted an average accuracy of 84.2% and 79.7% in the 2-alternative condition, respectively. Compared to the actual subject performance (average accuracy = 78%), the population model overestimated the accuracy in the 2-alternative conditions for 29 of the 32 subjects (average difference = 6.21%, t(31) = 8.19, $p = 3.02 \times 10^{-9}$). Surprisingly, the summary model also overestimated the accuracy in the 2-alternative condition but the misprediction was much smaller (average difference = 1.72%, t(31) = 2.35, $p = .025$) (**Figure 4A**). Indeed, the absolute error of the predictions of the population model (average = 6.24%) was significantly larger than for the summary model (average = 3.61%; t(31) = 5.65, $p = 3.36 \times 10^{-6}$). Overall, the summary model predicted the accuracy in the 2-alternative condition better than the population model for 26 of the 32 subjects (**Figure 4B**).

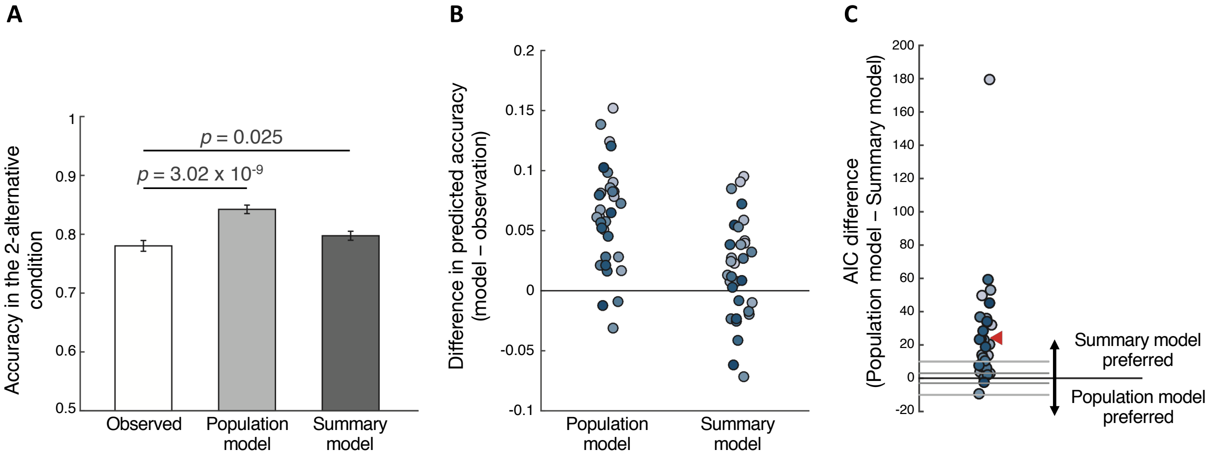

Figure 4. Comparisons between the population and summary models in Experiment 1. (A) Task accuracy in the 2-alternative condition observed in the actual data (white bar), and predicted by the population (light gray bar) and summary (dark gray bar) models. The predictions for both models were derived based on the data in the 4-alternative condition. (B) Individual subjects' differences in the accuracy in the 2-alternative condition between the two models and the observed data. (C) Difference in Akaike Information Criterion (AIC) between the population and the summary models. Positive AIC values indicate that the summary model provides a better fit to the data. Each dot represents one subject. The gray horizontal lines at ±3 and ±10 indicate common thresholds for suggestive and strong evidence for one model over another. The red triangle indicates the average AIC difference. The summary model provided a better fit than the population model for 30 of the 32 subjects.

Beyond simply checking accuracy levels, we further compared the model fits to the whole distribution of responses. We found that the Akaike Information Criterion (AIC) favored the summary model by on average 24.30 points (**Figure 4C**), which corresponds to the summary model being 1.86 x 10^5 times more likely than the population model for the average subject. Across the whole group of 32 subjects, the total AIC difference was thus 777.63 points, corresponding to the summary model being 7.26 x 10^{168} times more likely in the group. Note that since the population and summary models had the same number of parameters, the same results would be obtained regardless of the exact metric employed (e.g., the BIC differences would be exactly the same).

Finally, we constructed another summary model, according to which deliberate decisions had access to highest two activations of the sensory distribution (the 2-Highest model). We found that the 2-Highest model (average predicted accuracy = 83.5%) significantly overestimated the observed accuracy level for the 2-alternative condition (average difference = 5.46%, t(31) = 7.49, p = 1.94 x 10^{-8}) (**Supplementary Figure 1A**). Correspondingly, the absolute errors in the prediction of the 2-Highest model for the 2-alternative condition (average = 5.86%) is larger compared to the summary model (t(31) = 2.24, p = 4.07 x 10^{-5}). Model comparison favored the summary model over the 2-Highest by an average 11.86 AIC points (corresponding to the summary model being 357.41 times more likely for the average subject) and by 379.39 AIC points in the group as a whole (corresponding to the summary model being 2.42 x 10^{82} times more likely in the group) (**Supplementary Figure 1B and C**). Note that within the context of this experiment, a model assuming that decision-making circuits have access to the three highest sensory activation levels (3-Highest model) is functionally equivalent to the population model. Indeed, for such 3-Highest model, the activity level which is not represented is always the lowest; therefore, the 3-Highest model allows one to still order all four activity levels in descending order making it equivalent to the population model.

These results strongly suggest that within the context of our experiment, decision-making circuits do not have access to the whole sensory population distribution. We sought to confirm and generalize these findings in two additional, pre-registered experiments. For Experiment 2, we made several modifications: (1) we changed the stimulus from color to symbols, (2) we raised the number of stimulus categories from four to six, and (3) we significantly increased the number of trials per subject in order to obtain stronger results on the individual subject level. Specifically, we presented the six symbols '?', '#', '$', '%', '+', and '>' such that the dominant symbol was presented 14 times and each non-dominant symbol was presented 7 times (**Figure 5A**). The 49 total symbols were again arranged in a 7x7 grid. Each subject completed a total of 3,000 trials that included equal number of trials in 6-alternative and 2-alternative conditions.

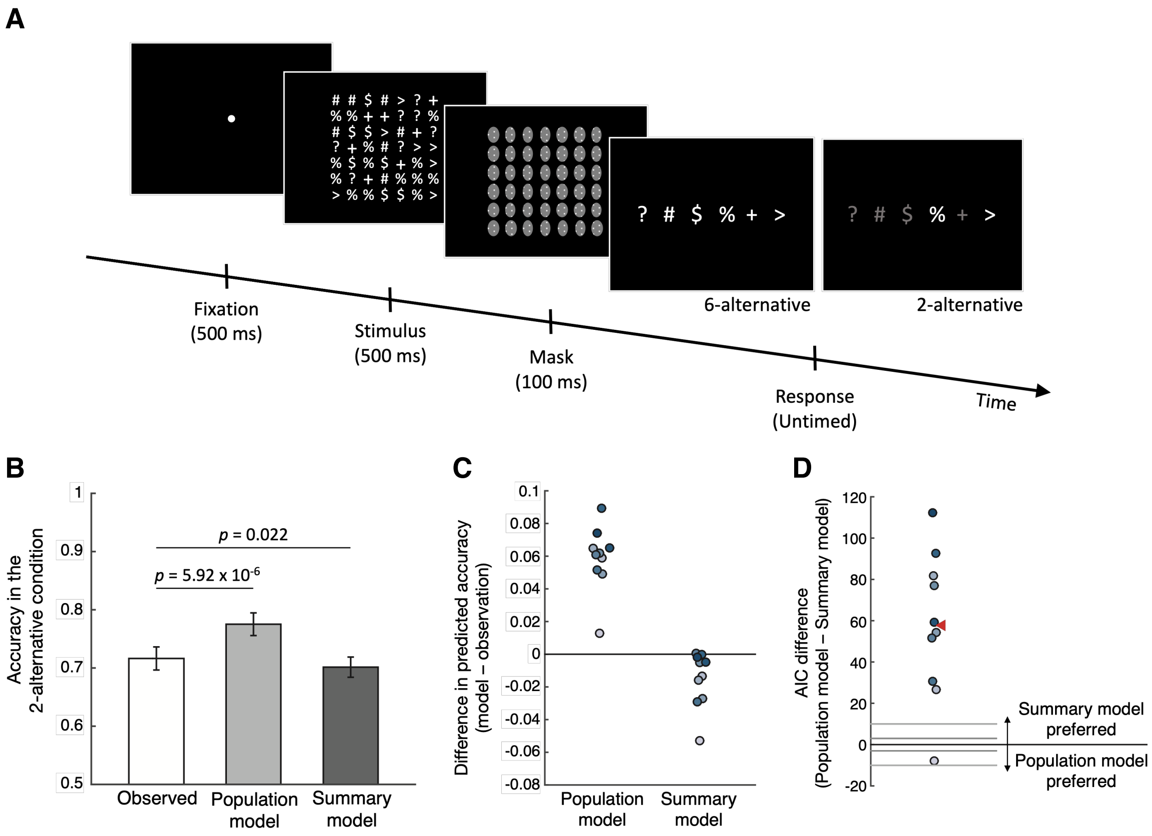

Figure 5. Task and results for Experiment 2. (A) The task in Experiment 2 was similar to Experiment 1 with the exception of using six different symbols ('?', '#', '$', '%', '+', and '>') instead of four different colors. One of the symbols was presented more frequently (14 times, dominant symbol) than the others (7 times each, non-dominant symbols) and subjects' task was to indicate the dominant symbol. Two conditions were presented in different blocks: a 6-alternative condition where subject chose between all six symbols and a 2-alternative condition where subjects were given a choice between the dominant and one randomly chosen non-dominant symbol. (B) Task accuracy in the 2-alternative condition observed in the actual data (white bar), and predicted by the population (light gray bar) and summary (dark gray bar) models. The predictions for both models were derived based on the data in the 6-alternative condition. (C) Individual subjects' differences in the accuracy of the 2-alternative condition

between the two models and the observed data. (D) Difference in Akaike Information Criterion (AIC) between the population and the summary models. Positive AIC values indicate that the summary model provides a better fit to the data. Each dot represents one subject. The red triangle indicated the average AIC difference. The summary model provided a better fit than the population model for nine out of 10 subjects.

Just as in Experiment 1, we computed the parameters of the sensory population code using the trials from the 6-alternative condition (average accuracy = 50.5%, chance level = 16.7%) and used these parameters to compare the population and summary models' predictions for the 2-alternative condition. We found that the average accuracy in the 2-alternative condition (71.6%) was slightly underestimated by the summary model (predicted accuracy = 70.1%, t(9) = 2.76, *p* = .022) but was again significantly overestimated by the population model (predicted accuracy = 77.5%, t(9) = 9.41, *p* = 5.92 x 10^{-6}) (**Figure 5B**). Individually, the summary model provided better prediction of the accuracy in the 2-alternative condition for nine out of the 10 subjects (**Figure 5C**).

Further, we compared the population and summary models' fits to the whole distribution of responses. We found that the summary model was preferred nine of our 10 subjects and the difference in AIC values in all these nine subjects were larger than 25 points. The AIC values of the one subject for whom the population model was favored over the summary model differed only by 7.8 points. On average, the summary model had an AIC value that was 57.79 points lower than the population model corresponding to the summary model being 3.55 x 10^{12} times more likely for the average subject (**Figure 5D**). Across the whole group of 10 subjects, the total AIC difference was thus 577.94 points, corresponding to the summary model being 3.14 x 10^{125} times more likely in the group.

Finally, we again investigated the predictions of the additional 2-Highest and 3-Highest models, according to which the highest two or three (respectively) activations were available for deliberate decisions. We found that the 2-Highest model overestimated the observed accuracy in the 2-alternative condition (74.9%; t(9) = 5.65, p = 3.13 x 10^{-4}) and provided worse fit to the data compared to the summary model (average AIC difference = 18.81 points, total AIC difference = 188.07 points) (**Supplementary Figure 2**). The 3-Highest model fared even worse. It overestimated the accuracy in the 2-alternative condition even more severely (76.6%; t(9) = 8.27, p = 1.69 x 10^{-5}) and provided much worse fit to the data compared to the summary model (average AIC difference = 39.87 points, total AIC difference = 398.65 points) (**Supplementary Figure 2**).

Taken together, Experiments 1 and 2 suggest that the system for deliberate decision making may not have access to the whole population of sensory activations. This conclusion is based on experiments that differed in the nature of the stimulus, the number of stimulus categories, and the amount of trials that subjects performed. Nevertheless, both Experiments 1 and 2 relied on the same design of comparing 4- (or 6-) and 2-alternative conditions.

Therefore, in order to further establish the generality of our results, in Experiment 3 we employed a different experimental design. We used the same stimulus as in Experiment 2, presented all 6 alternatives on every trial, but additionally gave subjects the opportunity to provide a second answer on about 40% of error trials (**Figure 6A**). We only allowed a second answer on a fraction of the trials in order to discourage subjects from preparing two responses while still viewing the stimuli. Using the performance on the first answer, we could then compare the predictions of the population and summary models for the second answers.

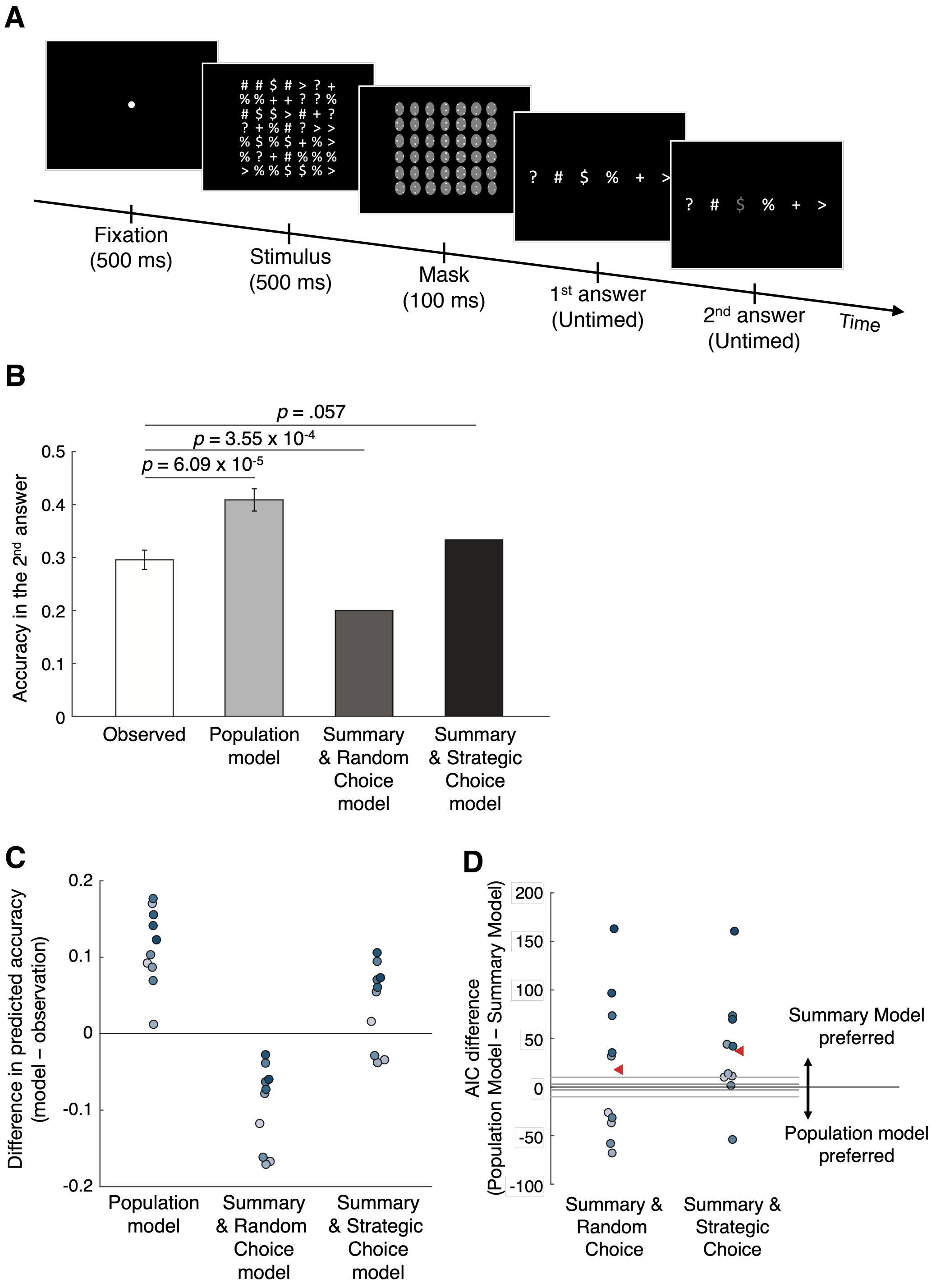

Figure 6. Task and results for Experiment 3. (A) The same stimuli as in Experiment 2 were used in Experiment 3 but the task was slightly different. Subjects always reported the dominant symbol among all six alternatives. However, on 40% of the trials in which they gave a wrong answer,

subjects were given the opportunity to make a second guess. (B) Task accuracy for the second answer observed in the actual data (white bar), predicted by the population model (light gray bar), predicted by the Summary & Random Choice model (dark gray bar), and predicted by the Summary & Strategic Choice model (black bar). The predictions of the three models were derived based on subjects' first answers. (C) Individual subjects' differences in the accuracy of the second answer between each model's prediction and the observed data. (D) Difference in Akaike Information Criterion (AIC) between the population and the two summary models. Positive AIC values indicate that the summary model provides a better fit to the data. Each dot represents one subject. The red triangle indicates the average AIC difference.

The population model makes a clear prediction about the second answer – subjects should choose the stimulus category with the highest activation from among the remaining five options. The second answer will thus have relatively high accuracy because the presented stimulus category is likely to produce one of the highest activity levels (**Supplementary Figure 3A**). Note that in the context of this experiment, the 2-Highest and 3-Highest models are functionally equivalent to the population model since they both represent the category with the second highest activity and both allow that this stimulus category be chosen with the second answer.

On the other hand, the summary model is consistent with at least two response strategies for the second answer. The summary model only features information about the stimulus category with highest activity, and thus once that stimulus category is chosen as the first answer, the model postulates that the subject does not have access to the activations associated with the other stimulus categories. One possible scenario, therefore, is for the subject to make their second answer at random. This model would result in chance level (20%) performance. We call this the "Summary & Random Choice" model (**Supplementary Figure 3B**). However, another possibility is for the subject to make the second answer strategically. One available strategy is for subjects to pick the stimulus category of a randomly recalled symbol from the 7x7 grid. Given that subjects inspected the stimuli for 500 ms, they could easily remember one location with a symbol other than the one they picked for their first answer. We call this the "Summary & Strategic Choice" model (**Supplementary Figure 3C**). According to this model, the second answer will be correct on $\frac{14 \text{ (\# locations of the dominant symbol)}}{42 \text{ (total \# of remaining locations)}} = \frac{1}{3}$ or 33.3% of the time. Conversely, each of the four remaining incorrect categories will be chosen on $\frac{7 \text{ (\# locations of each non-dominant symbol)}}{42 \text{ (total \# of remaining locations)}} = \frac{1}{6}$ or 16.7% of the time (**Supplementary Figure 4**).

To adjudicate between these three models, we first examined subjects' accuracy on the first answer. Subjects responded correctly in their first answer on 50.7% of the trials (chance level = 16.7%). Using this performance, we computed the parameters of the population code as in Experiments 1 and 2 in order to generate the models' predictions for the second answer. We found that subjects' accuracy for the second answers was 29.6%. This value was greatly overestimated by the population model, which predicted accuracy of 40.89% (t(9) = 7.04, p = 6.09×10^{-5}; **Figure 6B**). On the other hand, the Summary & Random Choice model greatly underestimated the observed accuracy (predicted accuracy = 20%, t(9) = 5.55, $p = 3.55 \times 10^{-4}$).

Finally, the Summary & Strategic Choice model produced the most accurate prediction (predicted accuracy = 33.3%, t(9) = 2.18, *p* = .057). On an individual subject level, the population model overestimated the accuracy of the second answer for all 10 subjects, the Summary & Random Choice model underestimated the accuracy of the second answer for all of the 10 subjects, whereas the Summary & Strategic Choice model was best calibrated overestimating the accuracy of the second answer for 7 subjects and underestimating it for the remaining 3 subjects (**Figure 6C**).

Formal comparisons of the models' ability to fit the full distribution of responses for the second answers demonstrated that the population model provided the worst overall fits (**Figure 6D**). Indeed, the population model resulted in AIC values that were higher than the Summary & Random Choice model by an average of 18.05 points (corresponding to 8.29 x 10^3-fold difference in likelihood in the average subject) and a total of 180.46 points (corresponding to 1.53 x 10^{39}-fold difference in likelihood in the group). The population model underperformed the Summary & Strategic Choice model even more severely (average AIC difference = 37.29 points, corresponding to 1.25 x 10^8-fold difference in likelihood in the average subject; total AIC difference = 372.93 points, corresponding to 9.57 x 10^8-fold difference in likelihood in the group). Thus, just as Experiments 1 and 2, Experiment 3 provided strong evidence that decision-making circuits do not have access to the whole distribution of sensory responses.

Discussion

We investigated whether decision-making brain circuits have access to the complete population codes in sensory cortex or only to a summary form of these codes. We designed three experiments with multiple (either four or six) discrete stimulus categories where the population and summary models made clearly different predictions. The results across all experiments showed that the population model did not provide a good fit to the data. Instead, the summary model, which assumes that decision-making circuits only have access to information about the highest-activity stimulus category, consistently provided better fit. These results strongly suggest that deliberate decision making, at least when performed on discrete stimulus categories, only has access to a summary form of the sensory population code.

What do the current results imply for decision-making circuits' access to the population codes produced by features represented on a continuous dimension such as stimulus orientation or motion direction? Theories like probabilistic population codes (Zemel et al., 1998; Pouget et al., 2000, 2003; Ma et al., 2006) and neural sampling (Fiser et al., 2010; Berkes et al., 2011; Haefner et al., 2016; Orbán et al., 2016) have typically been applied to such continuous situations. Importantly, in many cases the population and summary models may actually be functionally equivalent for stimuli represented on a continuous dimension. Indeed, electrophysiological recordings suggest that the population code produced by a particular stimulus orientation or motion direction follows a roughly Gaussian distribution. If the population code is Gaussian, it can be summarized without loss of information with only its mean and variance. If normalization (Reynolds and Heeger, 2009; Carandini and Heeger, 2012) had been applied so that the total firing can be expected to sum to a constant, then the population code can be summarized by the level of activity of the most active neuron. This is possible because the identity of this neuron signals the mean of the distribution, and the level of its activity is proportionate to the standard deviation of the distribution. This last representation is equivalent to the summary model used here. These considerations demonstrate that the population and summary models are functionally equivalent in the context of most existing experiments where Gaussian sensory distributions can be assumed. Nevertheless, the population and summary models may be distinguishable for non-Gaussian (e.g., multimodal) distributions even for stimuli represented on a continuous scale.

If decision-making circuits indeed only have access to a summary form of the sensory population, does that mean that absolutely no computations can be based on the complete population codes? At present, we are agnostic about this issue. It is possible that some forms of automatic multisensory integration are accomplished by computing with the whole population codes in two different parts of sensory cortex (Hillis et al., 2002; Prsa et al., 2012; Saarela and Landy, 2015). Other automated computations may also use the whole sensory population codes. Nevertheless, it appears that non-automatic, flexible, and deliberate decision making only has access to a summary of the sensory population code.

Our findings can be interpreted as suggesting that complex visual displays are represented as a point estimate: that is, the decision-level representation features only the best guess of the system (e.g., “60°” orientation, “red” color, or “+” symbol). The possibility of a decision-level representation consisting of a single point estimate has been thoroughly debunked (Pouget et al., 2003; Knill and Pouget, 2004; Fiser et al., 2010; Ma, 2010; Ma and Jazayeri, 2014). For example, a point estimate does not allow us to rate how confident we are in our decision because we lack a sense of how uncertain our point estimate is. Given that humans and animals can use confidence ratings to judge the likely accuracy of their decisions (Metcalfe and Shimamura, 1994; Kiani and Shadlen, 2009; Mamassian, 2016), decision-making circuits must have access to more than a point estimate of the stimulus.

It should therefore be clarified that our summary model does not imply that decision making operates on point estimates. Indeed, as conceptualized in Figure 1, the summary model assumes that subjects have access to both the identity of the most likely stimulus category (e.g., the color "white") and the level of activity associated with that stimulus category. The level of activity can then be used as a measure of uncertainty, and confidence levels can be based on this level. Such confidence ratings will be less informative than the perceptual decision, which is exactly what has been observed in a number of studies (Maniscalco and Lau, 2012; Rahnev et al., 2015; Rahnev and Denison, 2018). In addition, this type of confidence generation may explain findings that confidence tends to be biased towards the level of the evidence for the chosen stimulus category and tends to ignore the level of evidence against that stimulus category (Zylberberg et al., 2012; Koizumi et al., 2015; Maniscalco et al., 2016; Samaha et al., 2016; Peters et al., 2017; Talluri et al., 2018). Thus, a summary model, consisting of the identity of the most likely stimulus and the level of activity associated with this stimulus, appears to be broadly consistent with findings related to how people compute uncertainty and is qualitatively different than a decision-level representation consisting of a point estimate.

Another important question concerns whether any additional information is extracted from the sensory population code beyond what is assumed by the summary model. It is well known that humans can quickly and accurately extract a high-order “gist” of a scene (Potter, 1975, 1976; Greene and Oliva, 2009), as well as the statistical structure of an image (Fiser and Aslin, 2001). Therefore, it appears that rich information is extracted during the time when the stimulus is being viewed. For example, the subjects in our Experiment 1 were certainly aware that four different colors were presented in each display and would had noticed if we ever have presented additional colors. Nevertheless, within the context of our experiments, both the gist and the statistics of the scene remained unchanged between displays and therefore could not be used to improve performance on the task. In any case, we do not claim that rich information about the visual scene cannot be quickly and efficiently extracted (it can). What our results do suggest, however, is that decision-making circuits do not create a copy of the detailed sensory population code that can be used after the disappearance of the stimulus.

A related question concerns whether the decision-making circuits could ever access the whole population code. We believe that they can as long as the stimulus remains on the screen. However, the representation within the decision-making circuit itself may be of severely limited bandwidth. This is why, once the stimulus disappears, the richness of the sensory population code may be gone and only a limited-bandwidth summary representation within decision-making circuits may remain.

In conclusion, we found evidence from one exploratory (Experiment 1) and two preregistered (Experiments 2 and 3) studies that deliberate decision making for discrete stimulus categories is performed based on a summary of, rather than the whole, sensory population code. These results demonstrate that, at least within the context of discrete stimulus categories, flexible computations may not be performed using the sensory activity itself but only a summary form of that activity.

Methods

Subjects

A total of 52 subjects participated in the three experiments (32 in Experiment 1, 10 in Experiment 2, and 10 in Experiment 3). Each subject participated in only one experiment. All subjects provided informed consent and had normal or corrected-to-normal vision. The study was approved by the Georgia Tech Institutional Review Board.

Apparatus and experiment environment

The experiments stimuli were presented on a 21.5-inch iMac monitor in a dark room. The distance between the monitor and the subjects was 60 cm. The stimuli were created in MATLAB, using Psychtoolbox 3 (Pelli, 1997).

Experiment 1

The stimulus consisted of 49 circles colored in four different colors – red, blue, green, and white – presented in a 7x7 grid on black background. The diameter of each colored circle was .24 degrees and the distance between the centers of two adjacent circles was .6 degrees. The grid was located at the center of the screen. On each trial, one of the four colors was “dominant” – it was featured in 16 different locations – whereas the other three colors were non-dominant and were featured in 11 locations each. The exact locations of each color were pseudo-randomly chosen so that each color was presented the desired number of times.

A trial began with a 500-ms fixation followed by 500-ms stimulus presentation. Subjects then indicated the dominant color in the display and provided a confidence rating without time pressure. There were three different conditions in the experiment. In the first condition, subjects could choose any of the four colors (4-alternative condition). In the second condition, after the stimulus offset subjects were asked to choose between only two options that were not announced in advance – one was always the correct dominant color and the other was a randomly selected non-dominant color (2-alternative condition). Finally, in the third condition, subjects were told in advance which two colors will be queried at the end of the trial (advance warning condition).

Subjects completed six runs, each consisting of three 35-trial blocks (for a total of 630 trials). The three conditions used in the experiment were blocked such that one block in each run consisted entirely of trials from one condition and each run included one block from each condition. Subjects were given 15-second breaks between blocks and untimed breaks between runs. Before the start of the main experiment, subjects completed a training session where they completed 15 trials per condition with trial-to-trial feedback, and another 15 trials per condition without trial-to-trial feedback. No feedback was provided during the main experiment.

For the purposes of the current analyses, we focused on subjects' choices in the 4-alternative and 2-alternative conditions. The advanced warning condition and the confidence ratings were not analyzed.

Experiment 2

Following our exploratory analyses on the data from Experiment 1, we preregistered two additional experiments (Experiment 2 and 3) (https://osf.io/dr89k/). These experiments were designed to generalize the results from Experiment 1 and to obtain stronger evidence for our model comparison results on the individual subject level. Consequently, we had fewer number of subjects in Experiments 2 and 3 but each subject completed many more trials.

The stimulus in Experiments 2 and 3 consisted of 49 characters from among 6 possible symbols – '?', '#', '$', '%', '+', and '>' – presented in a 7x7 grid. The symbols were chosen to be maximally different from each other. The symbols' width was .382 degrees on average and height was .66 degrees on average. The distance between two centers of adjacent symbols was 1.1 degrees. The symbols were presented in white on black background. On each trial, one of the six symbols was "dominant" – it was featured in 14 different locations – whereas the other five were non-dominant and were featured in 7 locations each. The exact locations in the 7x7 grid where each symbol was displayed were pseudo-randomly chosen so that each symbol was presented the desired number of times.

Each trial began with a 500-ms fixation, followed by a 500-ms stimulus presentation. The stimuli were then masked for 100 ms with a 7x7 grid of ellipsoid-shaped images consisting of uniformly distributed noise pixels. Each ellipsoid had width of .54 degrees and height of .95 degrees, ensuring that it entirely covered each symbol. After the offset of the mask, subjects indicated the dominant symbol in the display without time pressure. No confidence ratings were obtained. The experiment had two conditions equivalent to the first two conditions in Experiment 1. In the first condition, subjects had to choose the dominant symbol among all six alternatives (6-alternative condition). In the second condition, subjects had to choose between two alternatives that were not announced in advance: the correct dominant symbol and a randomly selected non-dominant symbol (2-alternative condition). No feedback was provided.

To obtain clear individual-level results, we collected data from each subject over the course of three different days. On each day, subjects completed 5 runs, each consisting of 4 blocks of 50 trials (for a total of 3,000 trials per subject). The 6- and 2-alternative condition blocks were presented alternately, so that there were two blocks of each condition in a run. Subjects were given 15-second breaks between blocks and untimed breaks between runs. Before the start of the main experiment, subjects were given a short training on each day of the experiment.

Experiment 3

Experiment 3 used the same stimuli as in Experiment 2. Similar to Experiment 2, we presented a 500-ms fixation, a 500-ms stimulus, a 100-ms mask, and finally a response screen. Experiment 3 consisted of a single condition – subjects always chose the dominant symbol among all six alternatives. However, on 40% of trials in which subjects gave a wrong answer, they were asked to provide a second answer by choosing among the remaining five symbols. The second response prompt was presented infrequently in order to discourage subjects from preparing two responses from the very beginning. Subjects could take as much time as they wanted for both responses. Subjects again completed 3,000 trials over the course of three different days in a manner equivalent to Experiment 2.

Model development

To investigate how the sensory population code is represented and used for decision making, we developed and compared two main models. According to the "population" model, decision-making circuits have access to the whole sensory population code. On the other hand, according to the "summary" model, decision-making circuits only have the access to a summary of the sensory population code but not to the whole sensory distribution.

Although there are many ways to summarize a distribution, we considered an extreme case where the summary only contains the activity level of the stimulus category with the highest activity. In the case of making judgments on a continuous quantity (e.g., motion direction or Gabor orientation), the sensory population code can be assumed to follow a continuous distribution. Such distributions (e.g., a Gaussian) can often be represented without loss of information with just a few parameters (e.g., mean and variance for a Gaussian). However, in the context of deciding between discrete options, the activity levels in the sensory cortex corresponding to each stimulus category does not follow a distribution that can be easily summarized. In particular, our summary model, which only contains information about the highest activity level, loses significant amount of information compared to the population model. This feature of decisions about discrete stimuli enables us to distinguish the two models (**Figure 3**).

In order to compare the population and summary models, we first had to develop a model of the sensory population responses. We created this model using the 4- and 6-alternative conditions in Experiments 1 and 2, and the first answer in Experiment 3. The population and summary models were then used to make predictions about the 2-alternative condition in Experiments 1 and 2, and the second answer in Experiment 3. These predictions were made without the use of any extra parameters.

We created a model of the sensory population response for Experiment 1 as follows. We assumed that each of the four types of stimuli (red, blue, green, or white being the dominant color) produced variable across-trial activity corresponding to each of the four colors. We modeled this activities as Gaussian distributions whose mean (μ) is a free parameter and variance is set to one. However, perceptual decisions about the presented color only depend on the relative values of the activity levels, and therefore adding a constant to all μ's for a given stimulus results in equivalent decisions. Therefore, without loss of generality, we set the mean for the activity corresponding to each dominant color as 0. This procedure resulted in 12 different free parameters (4 types of stimuli $\times$ 3 μ's modeled for each stimulus type). Finally, we included an additional parameter modeling subjects' lapse rate. Note that the inclusion of lapse rate decreases performance more in the 2-alternative compared to the 4-alternative condition (because the overall performance is higher in the 2-alternative condition), and therefore favors the population model, which tends to overestimate the performance in the 2-alternative condition.

The sensory population response was modeled in a similar fashion in Experiments 2 and 3. In both cases, the model was created based on subjects choosing between all available options (i.e., the 6-alternative condition in Experiment 2 and the first answer in Experiment 3). The model of the sensory population response in Experiments 2 and 3 thus had 30 free parameters related to the sensory activations (6 types of stimuli × 5 μ's modeled for each stimulus type) and an additional free parameter for the lapse rate.

We modeled the activations produced by each stimulus type separately to capture potential relationships between different colors or symbols (e.g., some color pairs may be perceptually more similar than others). However, we re-did all analyses using the simplifying assumption of independence, which allowed us to significantly reduce the number of parameters in our model of the sensory population response. In this alternative version of the sensory population response model, the mean activity for each color/symbol was determined only based on whether that color/symbol was dominant or not. This modeling approach reduced the total number of free parameters to eight in Experiment 1 and 12 in Experiments 2 and 3. This modeling approach produced virtually the same results (Supplementary Figure 5).

We fit the models to the data as previously (Bang et al., (in press).; Rahnev et al., 2011, 2012, 2013) using a maximum likelihood estimation approach. The models were fit to the full distribution of probabilities of each response type contingent on each stimulus type. Model fitting was done by finding the maximum-likelihood parameter values using a simulated annealing (Kirkpatrick et al., 1983) by simulating the results of each parameter set 100,000 times. Fitting was conducted separately for each subject's data.

Based on the parameters of the sensory population response model, we generated predictions for the population and summary models for the 2-alternative condition in Experiments 1 and 2 and the second answer in Experiment 3. These predictions contained no free parameters. To compare the models, we calculated the log-likelihood ratio ($\log(\mathcal{L})$) of each model. However, for ease of comparison, we additionally computed Akaike Information Criterion (AIC). The value is computed by $-2 \times \log(\mathcal{L}) + 2 \times k$, where k is the number of parameters of a model. Because of the lack of free parameters, other measures, such as the AIC corrected for small sample sizes (AICc) or the Bayesian Information Criterion (BIC), would result in the exact same pattern of results. Note that lower AIC values correspond to better model fits.

In addition to the population and summary models, we considered two additional models that contained information about the sensory population that was more detailed than the summary model but less detailed than the population model. Specifically, we created models according to which decision-making circuits have access to the highest two or three levels of activation (2-Highest and 3-Highest models, respectively). Note that the 3-Highest model is functionally equivalent to the population model in the context of Experiment 1 and both the 2- and 3-Highest models are functionally equivalent to the population model in the context of Experiment 3.

Data and code

The data from the three experiments, together with all of the analysis codes are freely available online at https://github.com/wiseriver531/Discrete-representation.

References

Bang JW, Shekhar, Medha, Rahnev D (n.d.) Sensory noise increases metacognitive efficiency. J Exp Psychol Gen.

Bang JW, Shekhar M, Rahnev D (2018) Sensory noise increases metacognitive efficiency. J Exp Psychol Gen.

Berkes P, Orbán G, Lengyel M, Fiser J (2011) Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. Science (80-) 331:83–87.

Carandini M, Heeger DJ (2012) Normalization as a canonical neural computation. Nat Rev Neurosci 13:51–62.

Fiser J, Aslin RN (2001) Unsupervised Statistical Learning of Higher-Order Spatial Structures from Visual Scenes. Psychol Sci 12:499–504.

Fiser J, Berkes P, Orbán G, Lengyel M (2010) Statistically optimal perception and learning: from behavior to neural representations. Trends Cogn Sci 14:119–130.

Green DM, Swets JA (1966) Signal detection theory and Psychophysics. New York: John Wiley & Sons Ltd.

Greene MR, Oliva A (2009) The briefest of glances: the time course of natural scene understanding. Psychol Sci 20:464–472.

Haefner RM, Berkes P, Fiser J (2016) Perceptual Decision-Making as Probabilistic Inference by Neural Sampling. Neuron 90:649–660.

Helmholtz HLF (1856) Treatise on physiological optics. Bristol: Thoemmes Continuum.

Hillis JM, Ernst MO, Banks MS, Landy MS (2002) Combining Sensory Information: Mandatory Fusion Within, but Not Between, Senses. Science (80-) 298:1627–1630.

Hubel DH, Wiesel TN (1970) The period of susceptibility to the physiological effects of unilateral eye closure in kittens. J Physiol 206:419–436.

Kiani R, Shadlen MN (2009) Representation of confidence associated with a decision by neurons in the parietal cortex. Science (80-) 324:759–764 Available at: http://www.ncbi.nlm.nih.gov/pubmed/19423820.

Kirkpatrick S, Gelatt CD, Vecchi MP (1983) Optimization by simulated annealing. Science 220:671–680.

Knill DC, Pouget A (2004) The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci 27:712–719.

Koizumi A, Maniscalco B, Lau H (2015) Does perceptual confidence facilitate cognitive control? Atten Percept Psychophys.

Ma WJ (2010) Signal detection theory, uncertainty, and Poisson-like population codes. Vision Res 50:2308–2319.

Ma WJ, Beck JM, Latham PE, Pouget A (2006) Bayesian inference with probabilistic population codes. Nat Neurosci 9:1432–1438.

Ma WJ, Jazayeri M (2014) Neural Coding of Uncertainty and Probability. Annu Rev Neurosci 37:205–220.

Mamassian P (2016) Visual Confidence. Annu Rev Vis Sci 2:annurev-vision-111815-114630.

Maniscalco B, Lau H (2012) A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. Conscious Cogn 21:422–430 Available at: https://www.sciencedirect.com/science/article/pii/S1053810011002303.

Maniscalco B, Peters MAK, Lau H (2016) Heuristic use of perceptual evidence leads to dissociation between performance and metacognitive sensitivity. Atten Percept Psychophys 78:923–937.

Metcalfe J, Shimamura AP eds. (1994) Metacognition: Knowing about knowing. Cambridge, MA, US: The MIT Press.

Orbán G, Erkes P, Fiser J, Lengyel M (2016) Neural Variability and Sampling-Based Probabilistic Representations in the Visual Cortex. Neuron 92:530–543.

Pelli DG (1997) The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spat Vis 10:437–442.

Peters MAKK, Thesen T, Ko YD, Maniscalco B, Carlson C, Davidson M, Doyle W, Kuzniecky R, Devinsky O, Halgren E, Lau H (2017) Perceptual confidence neglects decision-incongruent evidence in the brain. Nat Hum Behav 1:0139.

Potter MC (1975) Meaning in visual search. Science (80-) 187:965–966.

Potter MC (1976) Short-term conceptual memory for pictures. J Exp Psychol Hum Learn 2:509–522.

Pouget A, Dayan P, Zemel R (2000) Information processing with population codes. Nat Rev Neurosci 1:125–132.

Pouget A, Dayan P, Zemel RS (2003) Inference and Computation with Population Codes. Annu Rev Neurosci 26:381–410.

Prsa M, Gale S, Blanke O (2012) Self-motion leads to mandatory cue fusion across sensory modalities. J Neurophysiol 108:2282–2291.

Rahnev D, Denison RN (2018) Suboptimality in perceptual decision making. Behav Brain Sci.

Rahnev D, Koizumi A, McCurdy LY, D’Esposito M, Lau H (2015) Confidence leak in perceptual decision making. Psychol Sci 26:1664–1680 Available at: http://journals.sagepub.com/doi/10.1177/0956797615595037.

Rahnev D, Kok P, Munneke M, Bahdo L, de Lange FP, Lau H (2013) Continuous theta burst transcranial magnetic stimulation reduces resting state connectivity between visual areas. J Neurophysiol 110:1811–1821 Available at: http://www.physiology.org/doi/10.1152/jn.00209.2013.

Rahnev D, Maniscalco B, Graves T, Huang E, de Lange FP, Lau H (2011) Attention induces conservative subjective biases in visual perception. Nat Neurosci 14:1513–1515 Available at: http://www.nature.com/articles/nn.2948.

Rahnev D, Maniscalco B, Luber B, Lau H, Lisanby SH (2012) Direct injection of noise to the visual cortex decreases accuracy but increases decision confidence. J Neurophysiol 107:1556–1563 Available at: http://www.physiology.org/doi/10.1152/jn.00985.2011.

Ratcliff R (1978) A theory of memory retrieval. Psychol Rev 85:59–108.

Reynolds JH, Heeger DJ (2009) The normalization model of attention. Neuron 61:168–185.

Saarela TP, Landy MS (2015) Integration trumps selection in object recognition. Curr Biol 25:920–927.

Samaha J, Barrett JJ, Sheldon AD, LaRocque JJ, Postle BR (2016) Dissociating perceptual confidence from discrimination accuracy reveals no influence of metacognitive awareness on working memory. Front Psychol 7:851 Available at: http://journal.frontiersin.org/Article/10.3389/fpsyg.2016.00851/abstract.

Talluri BC, Urai AE, Tsetsos K, Usher M, Donner TH (2018) Confirmation Bias through Selective Overweighting of Choice-Consistent Evidence. Curr Biol 28:3128–3135.e8.

Zemel RS, Dayan P, Pouget A (1998) Probabilistic Interpretation of Population Codes. Neural Comput 10:403–430.

Zylberberg A, Barttfeld P, Sigman M (2012) The construction of confidence in a perceptual decision. Front Integr Neurosci 6:79.