GIVEN THE BIRDS, WHERE IS THE FLOCK?

Visual estimation of the location of collections of points

Keiji Ota[1,2,3,4*] Qihan Wu[5] Pascal Mamassian[6] Laurence T Maloney[3,4]

[1] Centre for Brain and Behaviour, Department of Psychology, Queen Mary University of London

[2] Institute of Cognitive Neuroscience, University College London

[3] Department of Psychology, New York University, New York, United States

[4] Center for Neural Science, New York University, New York, United States

[5] Department of Psychological and Brain Sciences, Johns Hopkins University

[6] Laboratoire des systèmes perceptifs, Département d'études cognitives, École normale supérieure, PSL University, CNRS, Paris, France

* Corresponding author

Keywords

Location estimation, aggregation, parametric estimation, perceptual organization, clustering, non-Gaussian distributions

Abstract

How does the visual system assign locations to collections of points? The perceived location of a flock of birds is influenced by the locations of individual birds but is not synonymous with them. If the collection is a statistical sample from a particular distribution, the location estimation task is analogous to parametric estimation in statistics. The statistician knows the sample and also has partial knowledge of the distribution from which the sample was drawn. We examine how observers carry out parametric estimation tasks that are also visual organization tasks.

In our experiment the collections were one-dimensional statistical samples plotted along a horizontal line. They were drawn from one of three location families – the Gaussian, the Laplacian, and the Uniform. Observers saw only the sample and had to estimate the location of the generating probability density function. We compared human performance to unbiased minimum variance estimators (UMVUEs) for each family. We tested whether observers used the same or different estimators for all three distributional families and whether their estimators were UMVUEs.

Observers used different estimators for different distributional families. The estimators chosen for samples from the Gaussian and Uniform distributions were significantly different from the corresponding UMVUEs while that for the Laplacian was not. We developed a model of human performance based on clusters in samples. Clusters convey information about the distributions from which they are drawn. We propose a Two-Stage Visual Cluster Model (VCM) of location estimation. The observer first partitions the sample into non-overlapping clusters assigning a location to each cluster and then computes a weighted average of the cluster locations with weights reflecting how well each cluster fit the entire sample. The VCM captured human performance across all three distributions. These findings suggest that the visual system is sensitive to the structural information underlying the locations of individual points and bases their location estimates on Gestalt principles.

[309 words]

Introduction

The human visual system segments scenes into visual objects (Ullman, 1984; Ullman et al., 2002) adding visual structure to variation of light intensities. For example, birds in flight form a flock (Figure 1A). Each bird has a location, as does the flock. These 'wholes made of parts' have spatial locations and other spatial attributes – shape, size, etc (Koffka, 1935/2013; Ullman, 1984).

But where is the "center" of the flock in Figure 1A? And how does the visual system combine the locations of the individual birds to estimate the center of the flock? There is not a single correct answer, but from a practical point of view, we can appeal to the observer's own judgments. For example, the landing position of saccades to bivariate point clouds reflects estimates of the location of the clouds (McGowan et al., 1998; Melcher & Kowler, 1999).

One evident interpretation of a collection of points is that it is a sample drawn from a probability distribution. Exploratory data analysis (Tukey, 1977) is based on human ability to carry out statistical judgments visually (Rensink, 2017; Tufte, 2001). Visual estimates provide a welcome help to computational data analyses. With this interpretation, location estimation becomes a statistical problem and we can compare human performance in a location task to optimal performance as measured by various statistical criteria (Lehmann, 1998; Mood et al., 1974).

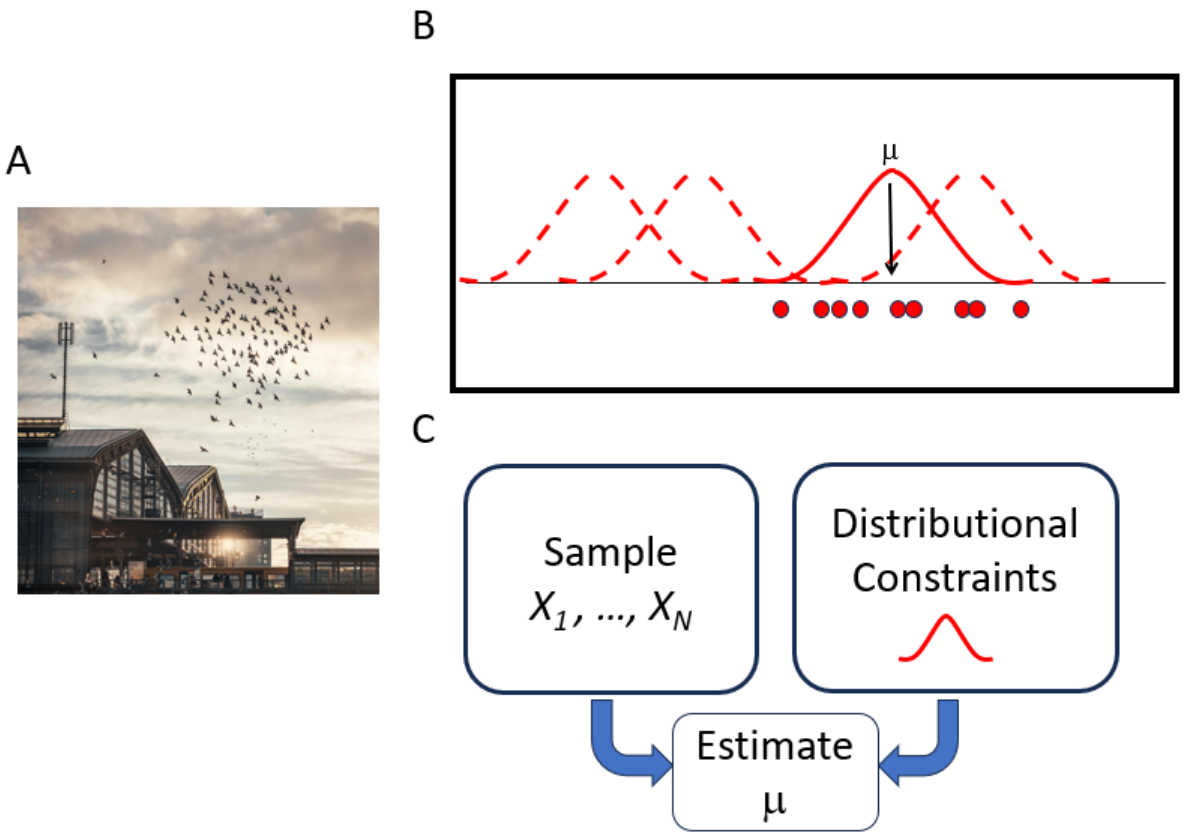


Figure 1: Estimating the Location of a Collection of Points. A. A Flock of Birds. A flock of birds in flight. **B. The 1D Gaussian Location Task.** The dashed and solid red curves are examples of Gaussian probability density functions (pdfs) with unit variance for different values of the location parameter μ. All of the red curves together comprise a *Gaussian location family*. The solid red curve is the actual pdf from which the sample was drawn. The sample is composed of *N* positions and is plotted as red points along the horizontal axis. The observer sees only the sample $X_1, \ldots, X_N$ and is asked to estimate the value of the location parameter μ corresponding to the pdf that was used to generate the sample. **C. 1D Parametric Estimation.** The observer sees the individual points and knows that the points are a sample from a pdf in the Gaussian location family with unit variance. They estimate the unknown location of the generating pdf.

Figure 1B illustrates the kind of statistical judgment we considered here, a one-dimensional analogue of the visual location task just discussed. The observer sees a series

of points $X_1, \ldots, X_N$ plotted as red dots along a horizontal line and is asked to estimate the center of the series of points.

In the present study, we consider different families of probability density functions (pdfs) from which the displayed sample is drawn. One of these families is the Gaussian pdf since it is commonly used in visual modeling, including in the important case of cue combination studies (Landy et al., 2007; Oruç et al., 2003). A second motivation for using Gaussian distributions in modeling visual judgments is that actual visual and motor error is close to Gaussian in a wide class of everyday tasks such as pointing and reaching (Ota et al., 2016; Ota et al., 2019; Ota et al., 2020; Tanae et al., 2021; Trommershauser et al., 2008). It is plausible that humans frequently encounter tasks involving Gaussian random variables and they may tend to apply estimation rules and other statistical procedures that are appropriate when distributions are Gaussian. However, it is less well known how humans behave when distributions are non-Gaussian (see (Maloney & Thomas, 1991) for discussion).

Figure 1C summarizes what the observer in our experiment knows – the locations of points and information about the distribution from which they are drawn. *Parametric point estimation* methods (Lehmann, 1998; Mood et al., 1974) allow us to compute optimal estimators based on criteria such as minimum variance for a wide range of distributions.

A *location family* is a collection of pdfs differing only in location (e.g. Figure 1B). Assuming for now that the pdf is Gaussian with fixed variance 1 and with unknown *location* parameter μ, the probability that position *x* is part of the sample is

$$\phi_\mu(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}} . \quad (1)$$

Several examples of pdfs from a *Gaussian location family* (Eq. 1) are plotted as dashed red curves in Figure 1B. As μ is varied, the pdf moves left and right (hence the term “location parameter”). The “true” pdf, the one from which the sample was actually drawn, is plotted as a solid red curve.

The observer’s task is – given only the sample and the Gaussian constraint – to estimate the unknown *location* μ of the true pdf, a key problem in parametric point estimation (Lehmann, 1998). The observer starts with the visible sample *and* knowledge that the points are a sample from a Gaussian to estimate the location of the distribution from which they were drawn.

The estimator that the observer uses is a function of the sample and may depend as well on the distributional assumption (Figure 1C). The observer’s estimate can be denoted $T_G^O(X_1, ..., X_n)$ or, if there is no ambiguity, just T_G^O. "O” stands for the Observer, “G” for the Gaussian assumption.

While the observer has two sources of information (*distributional constraints* and the *sample*) that inform their estimate $T_G^O(X_1, ..., X_n)$, it is usually the sample values that take center stage, with the Gaussian assumption in the background. But what use – if any – does the observer make of the distributional constraints? How are the parts combined to estimate the location of the whole, when the distribution generating the parts is not Gaussian?

In Figure 2 we illustrate the Location Task for different choices of distribution. Figure 2A corresponds to the Gaussian that we discussed above (Eq. 1). Figure 2B illustrates a location task based on the Laplacian distribution[1], where the pdf is now

[1] The actual pdfs used were always scaled to have the same variance. We present them in simplified form in Eqs. 2 and 3. See Methods.

$$\varphi_\mu(x) = \frac{1}{2} e^{-|x-\mu|} \tag{2}$$

The pdf is symmetric around the location parameter unknown μ. The observer's task is still to estimate μ. Finally, the Uniform Location Task is based on a uniform distribution (Figure 2C)

$$\begin{aligned} \varphi_\mu(x) &= 1 \qquad |x-\mu| < \frac{1}{2} \\ \varphi_\mu(x) &= 0 \qquad |x-\mu| \geq \frac{1}{2} \end{aligned} \quad . \tag{3}$$

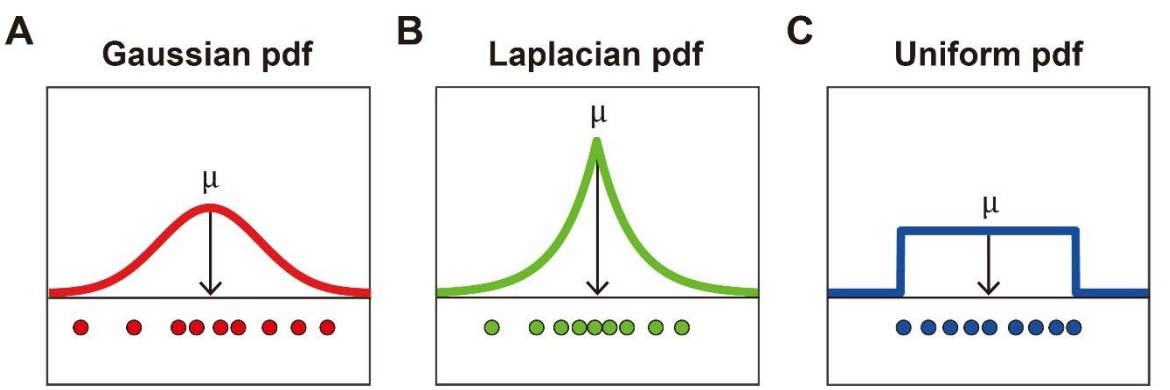


Figure 2. Location Tasks. A. Gaussian Location Task. B. Laplacian Location Task. C. Uniform Location Task. Each plot shows a probability function and a sample, color coded. The pdfs used in the experiment all had the same variance. See **Methods and Materials**.

We first assessed human performance in the three location tasks for the distributions shown in Figure 2. In our experiment, the observer will receive training to recognize and discriminate samples from each of the three distributions before attempting any of the three location tasks. On every trial, they will be told the identity of the location family used to generate the sample on that trial by a color code (as in Fig 2).

Uniformly Minimum Variance Unbiased Estimators (UMVUE). An estimator is a function of the sample, $T(X_1, \ldots, X_n)$ and, since an estimator is a function of random variables, it is itself a random variable with its own expected value $E[T]$ and variance $Var(T)$.

Common criteria for a "good" estimator are that it is *unbiased*, i.e. $E[T] = \mu$. Intuitively, the estimator of location in 1-D is unbiased if it is not – on average – larger or smaller than the true location. A second common criterion is that the estimator should have the lowest possible variance. If an estimator satisfies these two criteria for all values of μ then it is a *Uniformly Minimum Variance Unbiased Estimator (UMVUE)* or, less formally, the *minimum variance estimator.* Cue combination models (Landy et al., 1995; Oruç et al., 2003; Trommershauser et al., 2011) use UMVUEs and there is evidence that human observers do as well (Ernst & Banks, 2002; Juni et al., 2010).

We can compute a measure of the *efficiency* of any unbiased estimator T for a task as the ratio of the variance of the UMVUE to that of the estimator. Since $Var(T^{UMVUE}) \leq Var(T)$, the *efficiency* of T

$$H = Var(T^{UMVUE}) / Var(T) \tag{4}$$

is a value between 0 and 1.

The first issue we consider is, how does the human visual system adapt its estimation rule $T_G^O(X_1, \ldots, X_n)$ to location tasks based on each of the distributional families in Figure 2 – if it adapts it at all? Next, we will compare human performance to optimal unbiased estimators minimizing variance for each of the three location tasks in Figure 2.

What will be the outcome of the experiment? At one extreme, observers may simply adopt the same estimator for all three location tasks, effectively ignoring the distributional

information. At the other, the observer might adopt the correct UMVUE estimator for each of the three tasks.

L-Estimators. For convenience, we will focus on a set of estimators referred to as L-estimators (Lehmann, 1998). They are defined as weighted combinations of order statistics

$$T^{w}(X_1, \ldots, X_n) = w_1 X_{(1)} + \ldots + w_n X_{(n)} \tag{5}$$

where the sum of the weights is constrained to be 1. The order statistics (Mood et al., 1974) are just the sample values renumbered from left to right. We can represent any L-estimator as a plot of its weights (Figure 3).

The UMVUEs for the Three Location Tasks. Parametric estimation comprises a collection of mathematical techniques that allow us to identify optimal estimators including UMVUEs for different estimation problems with different distributional families (Lehmann, 1998; Mood et al., 1974). The unbiased estimator T^{+} minimizing variance for the Gaussian task illustrated in (Figure 1A) is just the arithmetic mean of the sample (Lehmann, 1998; Mood et al., 1974). It is repeated in Eq. 6, written in the form of an L-estimator

$$T^{+} = T^{+}(X_1, \ldots X_n) = \sum_{i=1}^{n} X_i \Big/ n = \sum_{i=1}^{n} \frac{1}{n} X_{(i)} . \tag{6}$$

This estimator is unbiased $E\left[T^{+}\right] = \mu$ and it minimizes variance $E\left[(T^{+} - \mu)^2\right]$.

Not every location task has an *UMVUE*. While the *UMVUE* for the Gaussian is the sample *mean* (Eq. 6), the Laplacian has no closed-form solution for its *UMVUE*. We computed the unbiased L-estimator that minimized variance by numerical methods (exhaustive search) and plot it in Figure 3B, the *UMVUE* among L-estimators denoted $UMVUE_L$. Whenever we

refer to the UMVUE of the Laplacian we intend the $UMVUE_L$ which will not affect our conclusions.

The *UMVUE* for the uniform family is

$$T^u(X_1, ..., X_n) = \frac{X_{(1)} + X_{(n)}}{2}, \qquad (7)$$

the average of the minimum and maximum values in the sample (the *average of extremes*). We have three different *UMVUEs* each matched to the distribution family of a location estimation task (Fig 3). Each is unbiased and each has the minimum possible variance among unbiased estimators when applied to its matching family.

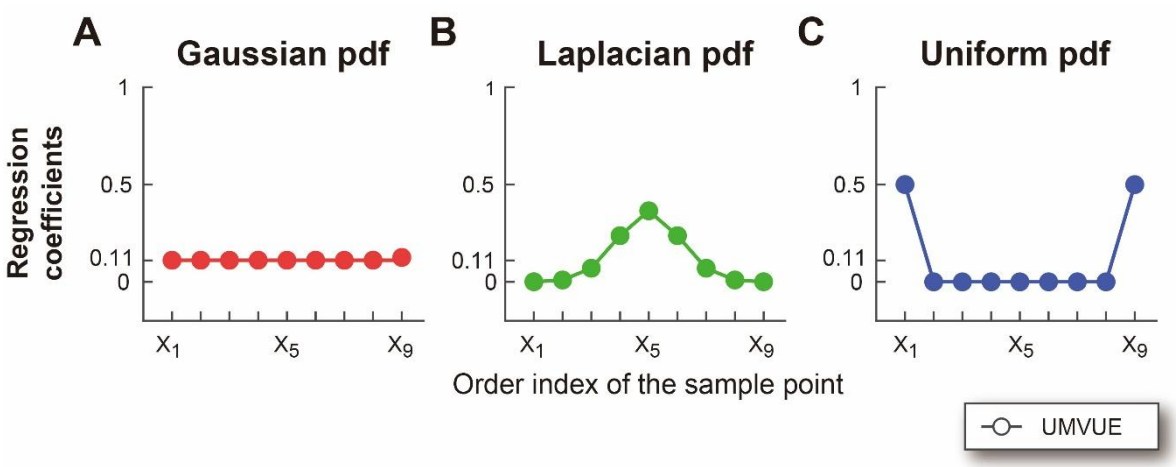


Figure 3: Uniformly Minimum Variance Unbiased Estimators (UMVUEs). A. Gaussian Location Task. B. Laplacian Location Task. The $UMVUE_L$ L-estimator estimated numerically (the Laplacian has no *UMVUE*). **C. Uniform Location Task.** Three UMVUE estimators are represented as L-estimators of weighted average of order statistics. Y-axis denotes weights (i.e., regression coefficients) on each sample point ordered from left to right. UMVUE estimator for the Gaussian location task is the arithmetic mean of the sample. The UMVUE for the Laplacian location task has the largest weight on the median. The UMVUE for the Uniform location task is the average of the two extremes.

Influence. We can interpret the weights in an L-estimator as measures of the *influence* of each point (order statistic), a useful characterization of a statistical task (Juni et al., 2010;

Maloney & Landy, 1989; Ota & Maloney, 2024) important for research in robust statistical estimation (Hampel et al., 1986; Huber, 1981).

In our experiment, we are presenting an observer with a sample whose order statistics — left to right – are $X_{(1)}$, ... $X_{(i)}$, ... , $X_{(n)}$ and recording the observer's response $T^O\left(X_{(1)} , \ldots X_{(i)} , \ldots , X_{(n)} \right)$, the location estimate. The influence of the point at position *i* on the location estimate is the partial derivative

$$I_i = \frac{dT^O}{dX_i} \quad . \tag{8}$$

which allows us to measure the contribution of each point to the overall estimate. If the location estimator is an L-estimator, then the influence estimates are just the weights in Eq. (5).

To summarize, in this article, we examine how observers combine locations of individual points to estimate the location of the generating pdf. To anticipate our results, we found that the estimators chosen for samples from the Gaussian and Uniform distributions were significantly different from the corresponding UMVUEs while that for the Laplacian was not. We developed a model of the observers' location estimates based on clusters in samples. We propose a Two-Stage Visual Cluster Model of location estimation. We test these two stages and possible alternatives in detail.

Results

First we tested whether the estimators that observers employed in the three location tasks were different or the same. That is, did the observer make any use of the distributional information? Second, we tested whether the participant's visual estimates of the position of the center of the population distribution were significantly different from optimal (UMVUE). Before doing so, we confirmed that total score did not significantly change across blocks (Supplementary Figure 2A) and that squared error between visual estimates and the true position did not significantly change across trials (Supplementary Figure 2B).

We know the UMVUE weights (influences) for each task (Figure 3) and we can estimate the observers' weights by linear regression (see Methods and Materials). Figure 4 (and Supplementary Figure 3A) illustrates the estimated weights for each participant and the average across participants. Observers were sensitive to the distributional constraints. Observers' influence measure markedly changed with changes in distributional family (interaction between the order of points and the distributional families by ANOVA on the estimated weights: F [16, 304] = 44.53, $p < 0.0001$, partial $\eta^2 = 0.701$).

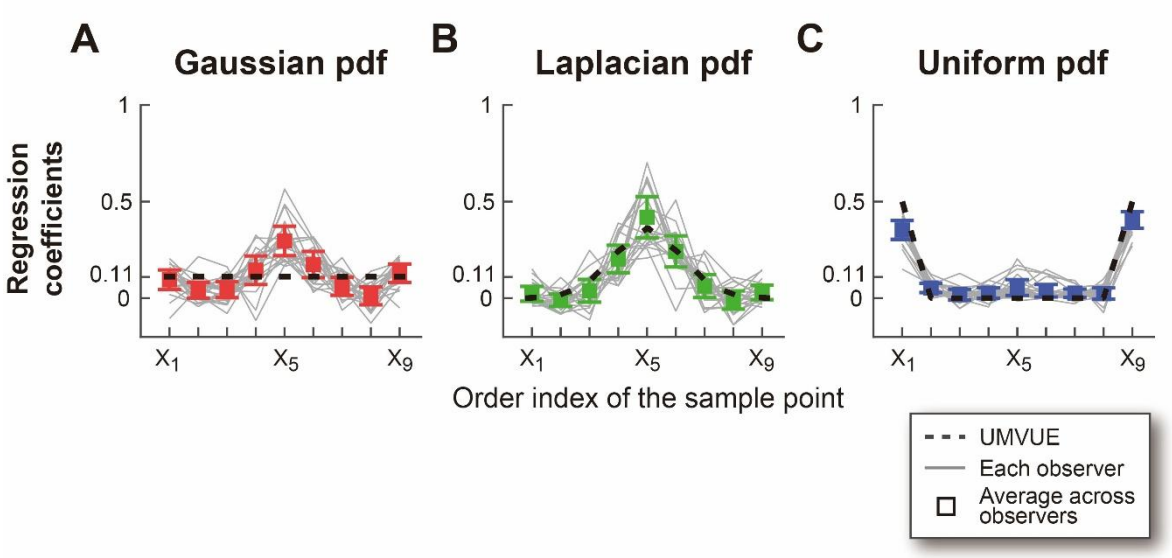


Figure 4: Observers' estimated influence measures compared to UMVUE influence measures. A. Gaussian Location Task. B. Laplacian Location Task. C. Uniform Location Task. Influence (i.e. regression coefficients) of each point in samples (ordered from left to right) on observer's visual estimates. Results for each observer are plotted as faint grey lines. Mean results across observers are plotted as color-coded squares**.** Error bar denotes ± 3 s.e.m. Weights employed by UMVUE are taken from Figure 3 and replotted as dashed lines. For the Laplacian distribution, observers' estimated influence measures are not significantly different from the UMVUE. The observers' influence measures for the other two distributions deviate from those of their respective UMVUE.

Do participants' choice of estimators approximate the UMVUEs for each location task? We compared measured influence with normative (UMVUE) influence for each location task separately. For the Gaussian task, the normative estimator is just the arithmetic mean of the sample. In contrast, the observer's estimated influences look w-shaped and are significantly different from the normative influences (F [8, 152] = 19.56, p < 0.0001, partial η^2 = 0.507; Supplementary Figure 3B), meaning that observers do something other than averaging points.

For the Uniform task, the measured influence is also different from normative influence (F [8, 152] = 35.27, p < 0.0001, partial η^2 = 0.650). Even though observers strongly relied on the two extreme sample points, they placed significantly smaller weights on them

than UMVUE (Supplementary Figure 3B). For the Laplacian task, observers employed approximated the UMVUE (F [8, 152] = 1.83, p = 0.076, partial η^2 = 0.088). We note that there was no significant change in the measured influences from the beginning to the end of the experiment (Supplementary Figure 2C). In summary, the observers are sensitive to the distributional constraints in the three location tasks – choosing significantly different estimators for the three location tasks. However, the observers' estimators significantly differed from the UMVUE for the Gaussian and Uniform location tasks.

The Visual Cluster Model (VCM)

Given a sample of points, which estimator will observers choose? We hypothesize that the visual system aggregates a collection of visual items as a cluster to define its overall location. Clusters convey information about the distributions from which they are drawn. Here, we considered the possibility that the visual system's choice of estimator is based on the cluster structure of the samples (Sun et al., 2019) that we first illustrate with an example.

In Figure 5A we plot a 9-point sample and in Figure 5B we partition the sample into K = 5 non-overlapping clusters denoted left to right as $C_1, C_2, \ldots, C_5$ using a standard clustering algorithm described below. Clusters have a minimum size of 1 element, and a maximum size of 9 (in which case, there is only one cluster). We assign a location c_j to the j'th cluster C_j, the midpoint of the endpoints of the cluster (representing the boundaries of the convex hull around the cluster).

In our example, the largest cluster, C_4 contains three points. If we had to base our estimate of location on just one of the clusters in the cluster structure of the sample, it would plausibly be the 4th cluster. But there are competing clusters with either 1 or 2 points, and we attempt to improve our estimator by weighting together the locations of the clusters with weights that depend on the clusters.

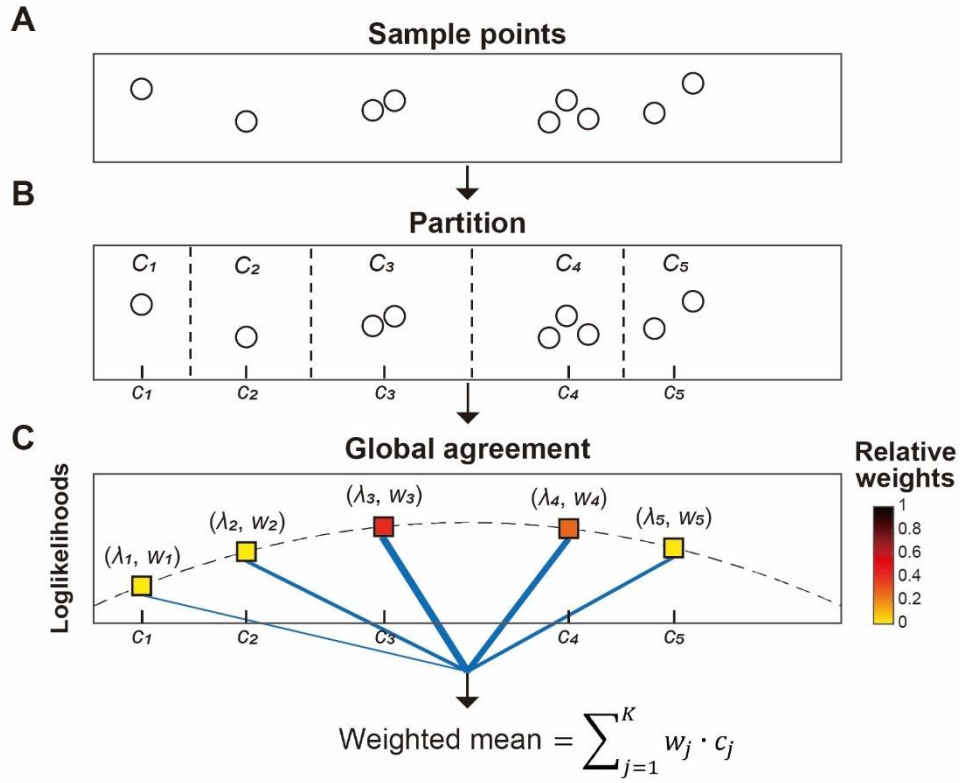


Figure 5: Visual Cluster Model. A. Gaussian Location task. A sample of 9 points is drawn from a Gaussian distribution. Observers were asked to estimate the location of the invisible center of the underlying distribution given only sample points. **B. Partition.** Visual cluster model partition points (partitions shown as dashed vertical lines) into K clusters of the sample, $C_1,...,C_K$. Each cluster is assigned a location $c_1,...,c_K$ defined as the average of its largest and smallest points if the cluster contains more than one point. Otherwise, the location of the single point is the location of the cluster. **C. Global agreement.** Visual cluster model assigns the log likelihood λ_j of the center μ of the pdf to the location c_j of the jth cluster. A dashed curve represents the log likelihood values across the entire space of the task. The log likelihood values $\lambda_1,\ldots,\lambda_K$ are converted to the relative weights $w_1,\ldots,w_K$ via a softmax transformation. The visual cluster model estimates the location of the center of the underlying distribution as a weighted average of cluster locations $c_1,...,c_K$ with weights $w_1,\ldots,w_K$.

We considered variations on the model at each step but – to simplify the presentation of the model – we first describe one version of the model and discuss possible variations only after presentation of this base version of the model. One goal of the model is to reduce

the complexity of the computation (Lieder & Griffiths, 2019) by relying on the cluster structure of the model (five clusters in our example) (Sun et al., 2019) rather than the nine points.

Step 1: Partition. The visual system partitions points in the sample into a small set of non-overlapping clusters. In Figures 5A and 5B, we apply a standard clustering algorithm, K-means clustering (Hartigan & Wong, 1979), to a sample $X_{(1)}, \ldots, X_{(N)}$ from the Gaussian distribution (Eq. 1). The clustering algorithm has one free parameter K (the number of clusters) which is estimated from the data. In Figure 5B, we obtained $K = 5$. The result is a partition into non-overlapping clusters $C_1,\ldots,C_K$ numbered from left to right, possibly varying in number of points. We assign a location to each cluster $c_1,\ldots,c_K$ as the average of its largest and smallest points (average of extremes) if the cluster contains more than one point, otherwise the location of the single point if the cluster contains only one. The first step is analogous to the cluster-of-samples representation developed by (Sun et al., 2019). We extend their model by introducing *global agreement*.

Step 2: Global agreement. How well does each cluster C_j at location c_j capture not just the points in its own cluster but rather the entire sample of points $X_{(1)}, \ldots, X_{(N)}$? We hypothesize that observers make use of the distributional constraint to assess the likelihood of the location of the distribution generating the points. In Figure 5C, for each cluster, we compute the log likelihood λ_j of the location parameter μ of the population pdf with respect to the location c_j of the j th cluster

$$\lambda_j = \sum_{i=1}^{9} \log f_\mu(X_i \mid \mu = c_j) \quad . \tag{8}$$

where f_μ is the likelihood function of the distribution chosen to generate the sample of points. This step results in an index of the likelihood that the location c_j of cluster *j* reflects the center of the distribution, taking into account all the points of the entire sample. Therefore, we refer to it as *global agreement*. A dashed curve in Figure 5C shows the log likelihood computed across the entire space of the task. In the Gaussian location task, finding the location of the maximum log likelihood value will correspond to averaging sample points. Instead of the performing the log likelihood computation over the entire space, we reduced the complexity of computation by computing the log likelihood value with respect to the locations c_j of the jth clusters.

We use a softmax transformation with one parameter β, i.e., known as inverse temperature (Reverdy & Leonard, 2016) to convert the relative log likelihood λ_j to weights

$$w_j = \frac{\exp(\beta\lambda_j)}{\sum_{j=1}^{K}\exp(\beta\lambda_j)} \tag{9}$$

The parameter β is a free parameter used to scale measures of global agreement. When β is fixed at 0, the relative weights w_j are all 1/K: all clusters get equal weight. When β increases, clusters having a large log-likelihood value gets relatively larger weights relative to those with smaller log-likelihood values. In the limit as β grows, the weight of the cluster with the largest relative log-likelihood converges to 1 while the other weights drop to 0.

On any given trial, the visual system's estimate of the position of the center of the distribution is modelled as the weighted mean of the locations of the clusters with softmax weights that depend on the relative log-likelihoods assigned to clusters (Figure 5C).

$$T^{VCM} = \sum_{j=1}^{K} w_j c_j + b \quad (10)$$

A bias term b is added in order to allow for the observer's constant offset from the population mean of the pdf reflecting a visuo-motor error, to account for the small but constant bias that we found in the participants' visual estimates (on average, 2.07 mm, SD = 1.41 mm).

Factorial model comparison. We test two levels of the grouping parameter K. At the first level, we will fix K at 9. At the second level, the grouping parameter K will be treated as a free parameter and estimated from the data (maximum likelihood). Given that samples from the three possible distributions are color-coded, the observer can select different values K_G, K_L, K_U of the grouping parameter K for the three different distributional families. These values will be estimated from the data.

For global agreement (β) we test three levels in the factorial design. At the first level, we will fix β at 0 so all clusters get equal weights. At the second level, the parameter β will be treated as a free parameter constrained to be the same value across the three distributional families. Last, at the third level, we will fix it at 1. By this constraint, global agreement is applied to the model fit but the β value is not estimated from each observer's data. We choose b to maximize log likelihood.

We built six models by combining two levels of partition and three levels of global agreement and compare the models by a factorial model comparison (van den Berg et al., 2014). Figure 6A shows the summed AICc score across participants for the six models. The full VCM model (partition and global agreement) outperformed the other five variants of the model. A group-level Bayesian model selection (Rigoux et al., 2014; Stephan et al., 2009) showed that the protected exceedance probability of the full VCM model, i.e. the probability for the model to outperform all the other models was 99.9%. This model was also better than

the normative model where UMVUE was fitted to the observer's data. Furthermore, models based on partitioning points outperformed the models without partitioning (the protected exceedance probability = 99.8%). A model where the assumption of partition was applied and global agreement was partially applied (β fixed at 1) outperformed a model where the assumption of partition was applied but global agreement was not (β fixed at 0), though the β value was not fitted to the data in these two models (the protected exceedance probability = 87.1%). These results support the claim that both partitioning and global agreement are key computations that underpin the observers' visual judgements.

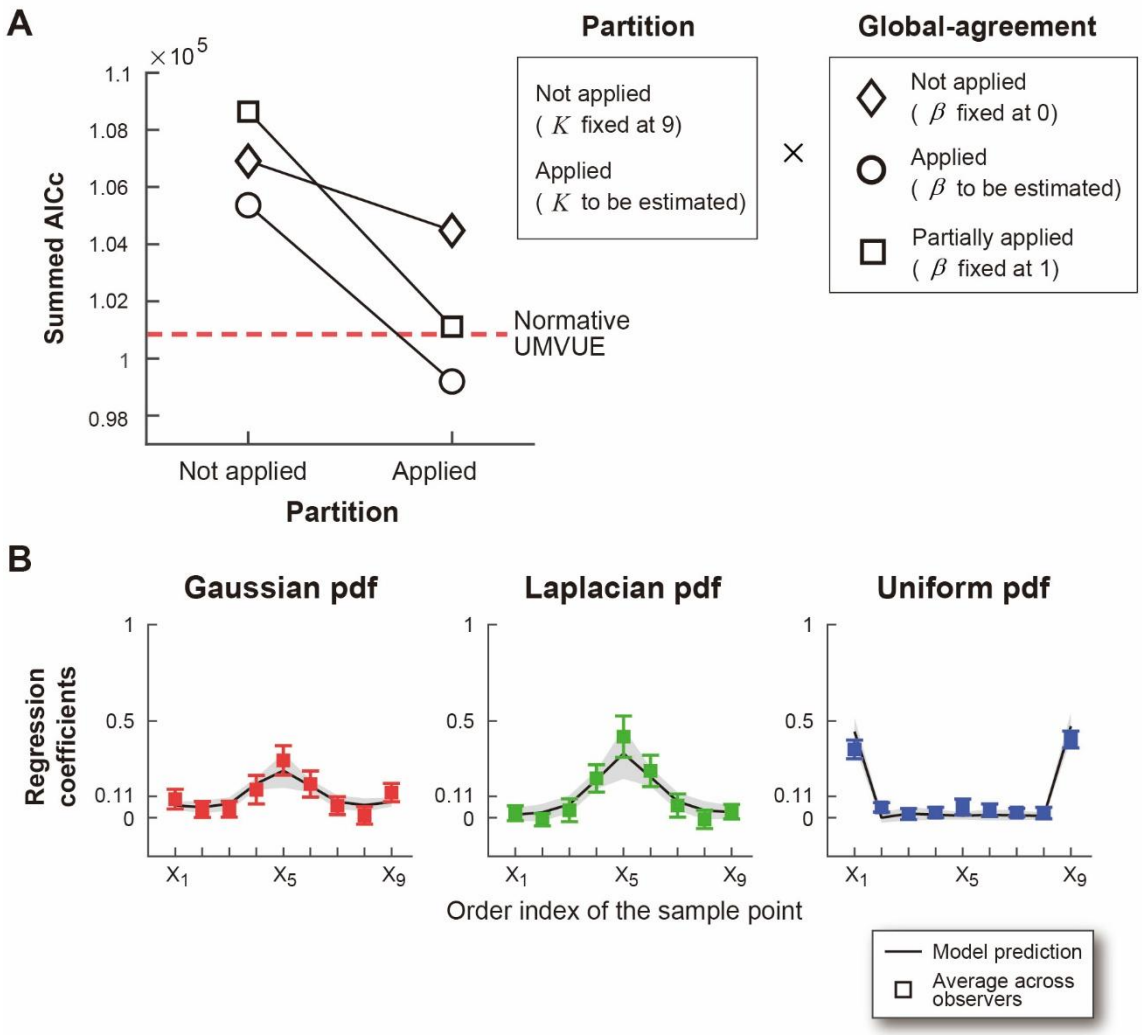


Figure 6: Visual Cluster Model compared to human performance. A. Factorial model comparison. We built six models by combining two levels of partitioning (Step 1) and three levels of global agreement (Step 2). The y-axis denotes the AICc score summed across participants. The lower the AICc, the better the model. A dashed red line plots the AICc score for the normative UMVUE. A grouping parameter K is the cluster size in K-means clustering we used to fit to the data. A parameter β is inverse temperature in the softmax transformation which converts the log likelihood values to the relative weights (Figure 5). **B. Prediction of the estimated weights (influence).** Regression weights estimated from the data are taken from Figure 4 and replotted as colored squares. Mean model fits across observers are plotted as black solid lines while shaded gray areas denote ± 3 s.e.m. The model prediction of the estimated weights was estimated using the visual estimates in the location tasks recovered from the best fit visual cluster model (Partition applied × Global agreement applied).

Using the visual estimates generated by the full VCM model, we estimated the influence of each point in the sample on their estimates. The pattern of the influence measure predicted by the model captured the human data well in all three distributional families (Figure 6B). Specifically, the model produced a w-shaped influence pattern in the Gaussian task, a median-biased influence in the Laplacian task, and the average of extremes in the Uniform task.

Patterns of the predicted influence by other variants of the model are summarized in Figure 7. When we applied the global agreement partially (β fixed at 1) and had grouping parameters free, the model still predicted patterns of measured influence in the three location tasks (Figure 7F), suggesting that we are not overfitting the data by including the global agreement parameter. However, without applying global agreement (β fixed at 0), we failed to properly fit the data (Figure 7A&D). The models without partition (K fixed at 9) also failed to fit the data (Figure 7A-C).

Finally, we tested whether our results do not depend on a choice of the clustering algorithm. We repeated the model fit using a hierarchical clustering with single linkage algorithm (Nielsen, 2016; Shetty & Singh, 2021) and found that the results were almost identical (Supplementary Figure 4).

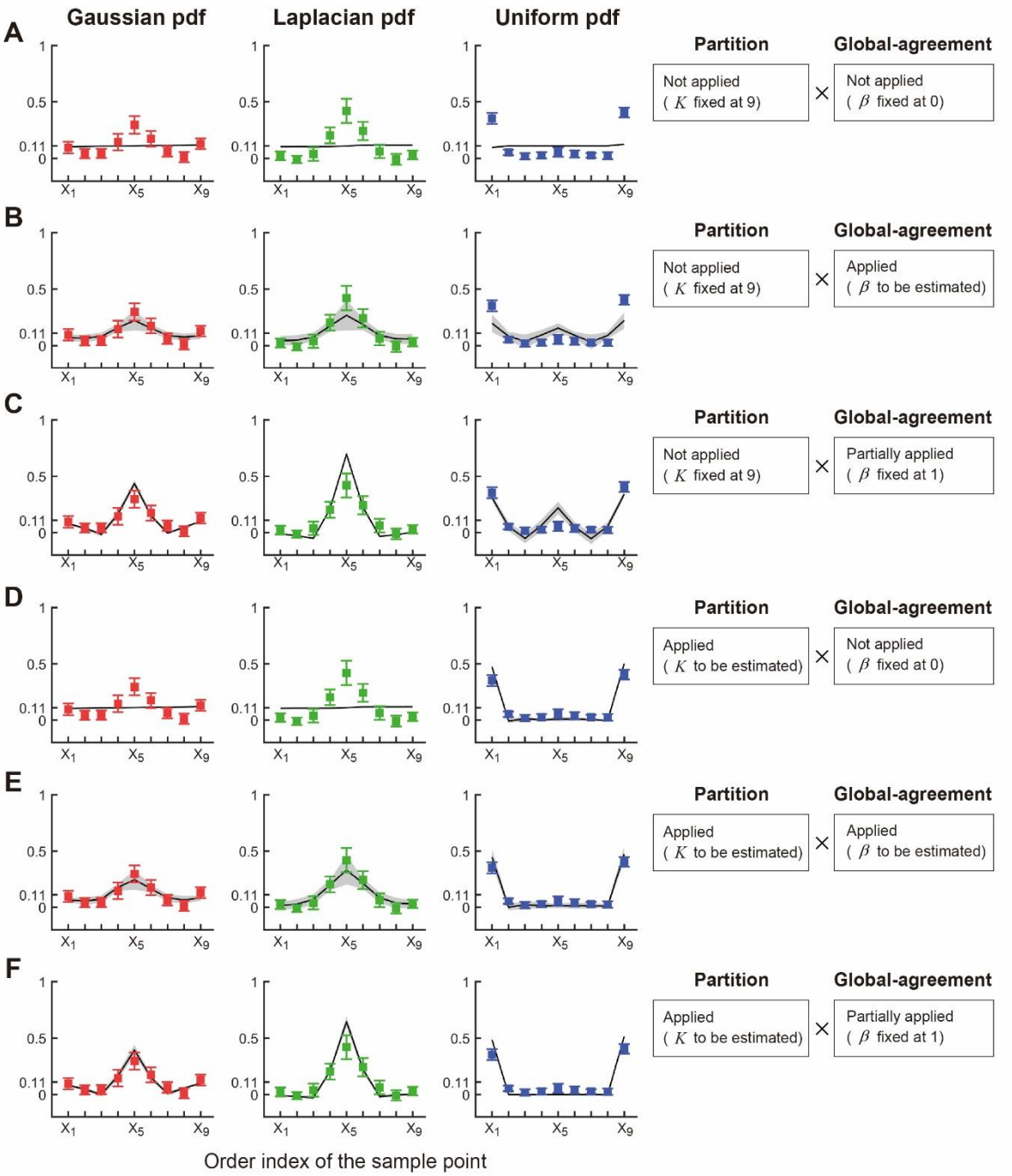


Figure 7. All variants of the visual cluster model.
The predictions of the estimated weights for each model in the factorial model comparison. Regression weights estimated from the data are taken from Figure 4 and replotted as colored squares. Mean model fits across observers are plotted as black solid lines while shaded gray areas denote ± 3 s.e.m.

Discussion

A central problem of visual perceptual organization is to model how the visual system combines parts together into an aggregated whole. When parts have spatial attributes such as location, how does the visual system combine the parts into an estimate of the location of the whole? We considered the case where a collection of points is sampled from a statistical distribution and we compare human performance to that of a Uniformly Minimum Variance Unbiased Estimator (UMVUE) for that distribution. The UMVUE for each distribution is unbiased and has the lowest possible variance among location estimators. The normative minimum variance estimator is different for the different distributions we consider, namely the Gaussian, the Laplacian, and the Uniform.

While Gaussian distributions are typically used in modeling many cue combination tasks (Ernst & Banks, 2002; Juni et al., 2010; Landy et al., 1995), it is less clear whether and how the human visual system compensates for non-Gaussian distributions (the Laplacian and the Uniform). We chose to use these three distributions because their UMVUEs are very different.

Using influence measures, we tested whether observers adopt the same estimator for the three different distributions. We found that observers used different location estimators for the three distributions. They did. We also tested whether the estimator for each distribution was the UMVUE for that distribution. The observers' rules were different from UMVUE for Gaussian and Uniform location tasks but we could not reject the hypothesis that observers used the UMVUE in the Laplacian location task.

Gestalt psychology (Koffka, 1935/2013) provided an early influential theory about perceptual organization. Visual objects are grouped together based on a similarity of spatial

attributes and these new *Gestalten* are assigned attributes including locations in space. Our findings suggest that *Gestalten* is not just a product that the visual system automatically operates on visual objects but *Gestalten* principle is used for a top-down system to guide visual judgments.

It is implausible that the human visual system has different, explicit rules of perceptual organization in our location task for any choice of location family. We wondered whether a common estimation algorithm based on Gestalt structure could account for how the visual system assigns locations to samples from arbitrary location families.

We developed a Two-Stage Visual Cluster Model (VCM). Our model comparison analysis validated the two stages on the VCM. The first partition stage edits individual visual points so that points can be grouped into a cluster which itself defines a location. Although we show participants just 9 sample points, models without the partition stage did not explain the patterns of observers' visual judgments well.

At the second stage, that we call global agreement, the VCM evaluates how likely each cluster represents a location of the distribution generating the points. A cluster with a high likelihood value is weighted more in the judgment than a cluster with a low likelihood value. The VCM then combines the locations of the clusters with weights that depend on the likelihood of the location of the underlying distribution. The distributional constraints provided information that was necessary to make the likelihood estimates of the location of the distribution. We demonstrated that the VCM accounts for the rules of the observer's visual judgments across all the three location tasks.

Our visual location tasks are different from the task used in Sun et al (2019) in the following points. First in their task, observers were asked to estimate the location of the empirical mean or mode of the sample of 70 visual points, and points were drawn from a

mixture of distributions that had three different peaks. While we generated just 9 points from a unimodal (Gaussian, Laplacian, or Uniform) distribution. Nevertheless, the visual cluster model that was used in that study also captured our data well.

The VCM offers two possible advantages. First, the visual system can reduce the amount of information that they need to process by summarizing the visual items as an aggregated whole. This becomes particularly advantageous with an increase in the number of visual items presented. Second, clustering anchors a specific spatial attribute for a collection of information that is useful for visual judgements. In the VCM, a cluster of points defined its own location and it anchored a position at which the likelihood of the location of the distribution generating the points was computed. The resulting computation depends on the number of clusters that the clustering algorithm creates, which helps reduce the complexity of computations, analogous to the resource rational principle in perception, cognition and decision-making (Lieder & Griffiths, 2019). For instance, in Figure 5C, the likelihood of the location of the distribution is evaluated only at five cluster locations. Nevertheless, it roughly approximates the maximum likelihood estimation as if we applied a low-pass filter at the limited sampling points. In this sense, Gestalt is not just a visual illusion but is functional for visual judgments.

Methods and Materials

Participants

Twenty participants (mean age 22.9, range 18–29, 13 females, 7 males, 0 other) participated in the experiment. This study was approved by the University Committee on Activities Involving Human Subjects of New York University (IRB-FY2019-3450) and carried out in accordance with its approved guidelines. Informed consent was given by each participant before the experiment. None of the participants was aware of the hypotheses under test in the experiment. Participants received $12 per hour plus a performance-related bonus (average bonus: $16.90, range: $11–$24). They were informed that there could be a bonus. The experiment took approximately 2 hours.

Apparatus

Stimuli were displayed on a horizontal monitor (VPIXX, VIEWPIXX, 514 mm × 288 mm). The monitor resolution was 1920 × 1080 pixels with a 60-Hz refresh rate and we refer to location on the monitor by horizontal coordinates ranging from -257 mm to 257 mm and vertical coordinates -144 mm to 144 mm. The participants were seated at a viewing distance of 43 cm and manipulated a computer mouse to carry out tasks. All stimuli were presented and controlled using the Matlab Psychophysics Toolbox (Brainard, 1997; Pelli, 1997).

Training Task

Participants were first familiarized with the three distributional families: Gaussian, Laplacian, and Uniform (see Supplementary Figure 1A and Supplementary Video 1). In this training stage, participants were not yet asked to make any estimates but were asked to observe the population pdfs, the population means, and the samples drawn from these population distributions. On each trial, we plotted a pdf drawn from one of the three distributional families and a sample from the same pdf. The origin (0,0) was set at the vertical and

horizontal center of the screen. The location μ of the pdf was drawn uniformly from the interval -100 mm and 100 mm. Following this, a sample of *N* points was randomly drawn. The sample points were displayed underneath the shape of the pdf. The points were discs with a radius of 1.5 mm and colored by the same color that was used to draw the underlying pdf (for instance, red Gaussian; green Laplacian; blue Uniform). The color codes were counter-balanced across participants.

The population variance σ^2 of all pdfs was fixed at 2500 mm^2. We added vertical jitter to the vertical positions of the sample points to reduce the chance that sample points would occlude one another. The variance of vertical jitter was fixed at 6.25 mm^2. To aid the observer in interpreting the pdf and sample during training, we marked the location μ of the pdf by a white solid line segment and dashed lines (width: 2 mm, height: 35 mm).

During training, the sample size *N* was initially 300 and was gradually reduced to 50, 20 and finally 9. With 300 points the displayed sample could be readily matched to the pdf from which it was drawn.

Presentation of the sample when N > 9 was extended across time: 18 additional points were drawn every 100 msec from the underlying pdf until the required number of points (300, 50, 20, or 9) points appeared on the screen. The sample points were displayed for 2s. After an inter-trial interval of 1s, a new pdf with randomly chosen location μ was displayed together with a sample drawn from the pdf.

A distribution was chosen from one of the three distributional families (Gaussian, Laplacian, and Uniform) and was repeated for five trials. In each trial, a new sample was generated from the distribution chosen. After the presentation of 300 sample points in groups of 15 = 5 x 3, the second block of 50 sample points started. The third block and the fourth block used 20 and 9 points, respectively, giving 60 trials in total. The final sample size

(N = 9) was the one used in the main experimental task, the location task. The training task provided substantial opportunity to observe typical Gaussian, Laplacian, and Uniform samples while learning the color codes.

In the instructions, we used the following analogy:

> On every trial, you will first see a white apple tree located somewhere on the screen. You will then see a curve which tells you where the apples are most likely to fall. The higher the curve is, the more likely an apple will fall below that point on the curve. Where the curve is high, there probably will be more apples. Where the curve is low, fewer. The density curve is always symmetric around the apple tree. The horizontal location of the apple tree is not fixed but it will vary from trial to trial along with the curve and the apples. The tree, the apples and the curve will all move together from left to right on each trial. There are three types of density curves from where apples will fall, each coded by color. One is bell-shaped (red), the second is "pointy" (green), and the last is like a box (blue).

Visual Location Task [Main Experiment]

The stimuli in the visual location task were the same as in the visual training task with the following differences: On each trial in the Visual Location Task, participants saw only a colored-coded sample of N = 9 points randomly drawn from the Gaussian, Laplacian or Uniform distribution. The pdf and the location of the pdf were not shown on the screen. To help participants remember the three density curves, small icons representing the shape of the three distributions were drawn near the top of the screen (see Supplementary Figure 1B and Supplementary Videos 2&3).

Participants were asked to estimate the unknown position of the center (i.e., location μ) of the underlying pdf using only the positions of sample points. They could move a white upward pointing arrow (10 mm) to any location on the screen using a mouse. Participants set the arrow at a position where they believed the horizontal center of the underlying pdf was located and clicked the mouse button to record their estimates. We ignored the vertical displacement of the arrow in analyzing the observer's settings.

Following a response, we presented a white vertical line (35 mm) marking the location of the underlying pdf as feedback on the error between the participant's estimates and the true position of the center of the pdf. Participants received 500 – D^2 points where D^2 is the squared magnitude of their error in mm. If the position of the estimates perfectly matched the true center ($D^2 = 0$), the participant received 500 points. Points received could be negative if the error in distance was above 22.4 mm or below -22.4 mm. At the end of each trial, the estimation error (mm), the score (points) and the cumulative score were presented as a graphical display for 1.5 s. Participants received a bonus payment proportional to the total score they earned across all trials ($1 per 10,000 points). The estimator that maximized the participants' expected bonus for each distributional family was the UMVUE for that distributional family.

We provided warm-up trials before the main experiment. During warm-up, the sample size was initially 300 and gradually reduced to 50, 20, and 9. The warm-up task became more difficult as the sample size was reduced. As in the training task, a distribution was chosen from one of the three distributional families (Gaussian, Laplacian, and Uniform) and was repeated for five trials. In each trial, a new sample was generated from the distribution chosen, giving 15 = 5 x 3 trials for the first block of 300 sample points. Then the second block of 50 sample points started. The third block and the fourth block used 20 and 9 points, respectively, giving 60 warm-up trials in total. Following a short break, the main experiment started. One block consisted of 10 trials. Samples from the three distributions were interleaved across blocks. Each participant ran 90 blocks (30 blocks for each distribution) with a short break after every 6 blocks.

Measuring influence

Order statistics. We performed a regression analysis to estimate the influence of each point in the sample on the participant's estimate of the position of the population mean of the distribution (Gaussian, Laplacian, or Uniform) that had generated the sample in the visual estimation task. We ordered 9 sample points in each sample in ascending order from the leftmost point to the rightmost point. For each of the three distributions, we regressed the vector of the participant's estimators $\mathbf{T^O}$ which contained 300 data points by the 300 by 9 matrix $\mathbf{X}$. Columns were vectors of order statistics for each trial $\mathbf{X}_{(1)},\ldots,\mathbf{X}_{(9)}$. Our linear regression model is thus $\mathbf{T^o} = \mathbf{X}\widehat{\mathbf{W}}$ where $\widehat{\mathbf{W}}$ is the estimated weight vector. We used an ordinary least squares regression with the constraint that weights add to 1 $\sum_{i=1}^{9} \widehat{w}_i = 1$. We estimated influence for each participant and each distributional family of the population distribution. In addition, we estimated the weight vector of the visual cluster model $\mathbf{W_{VCM}}$ by regressing the vector of the visual estimates predicted by the VCM $\mathbf{T^{VCM}}$ by matrix $\mathbf{X}$. For each variant of the VCM, we retrieved their visual estimates from the best fit parameters in each model (see *Model fitting and evaluation*). See Figure 7 for the estimated weights in each model. The influence measures are a sensitive way to uncover how participants behave in carrying out the task and discover discrepancies between human performance and optimal performance (Dal Martello et al., 2023; Ota et al., 2024; Ota et al., 2025b; Ota et al., 2025a; Ota & Maloney, 2024).

Model fitting and evaluation

We fitted the model's visual estimates [Eq 10] to the participant's estimates by minimizing the negative log-likelihood of the observed estimates using Bayesian adaptive direct search

(BADS) (Acerbi & Ma, 2017). The data were collapsed across the three distributional families. In the full VCM model, the grouping parameter K (i.e., the number of clusters) in K-means clustering was estimated individually from each participant and K was an integer between $K = 1$ and $K = 9$. The global agreement parameter β and the bias term b were estimated respectively with the following boundaries: $\beta \in [0,4]$, $b \in [-50,50]$. To verify that we had found the global minimum, we repeated the search process with different starting points. For model comparison, we applied AICc—Akaike information criterion with a correction for finite sample size—to each participant and each model as the information criterion for goodness-of-fit (Burnham & Anderson, 2002; Hurvich & Tsai, 1989). AICc has a penalty term which avoids potential overfitting and helps select the models that have fewer parameters as Bayesian information criterion (BIC) does. Based on AICc values across participants and across models, we computed the protected exceedance probabilities (Rigoux et al., 2014; Stephan et al., 2009) using a Matlab function 'spm_BMS' in SPM 12 (https://www.fil.ion.ucl.ac.uk/spm/). The sign of AICc values was reversed by multiplying with $-\frac{1}{2}$ such that an entry to the function 'spm_BMS' can be converted the log model evidence (https://mbb-team.github.io/VBA-toolbox/wiki/BMS-for-groupstudies/).

Competing Interests

The authors declare no competing interests.

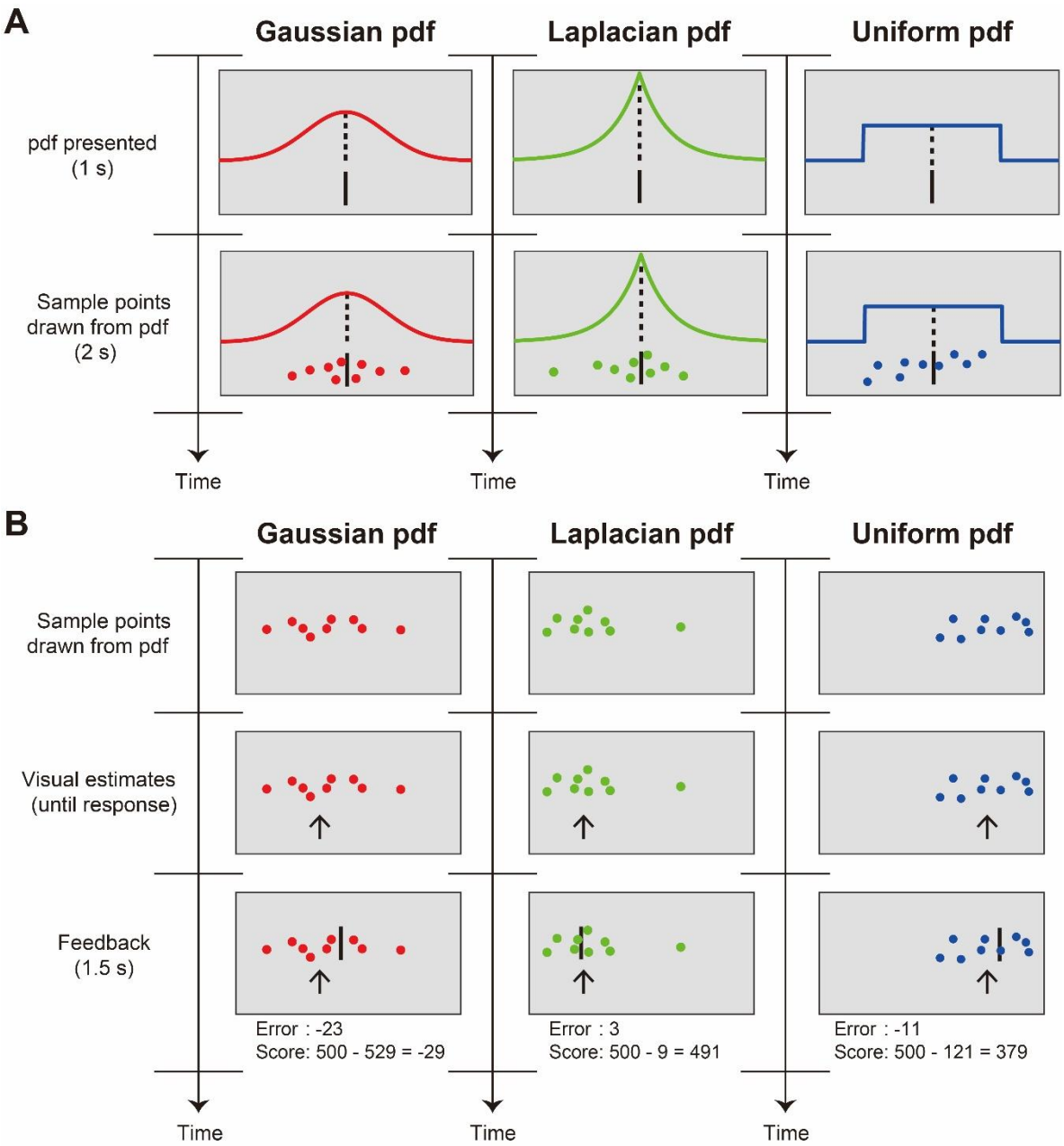
Author's Contributions

KO, QW, PM and LTM designed the experiment and analysis, QW and KO collected data. KO carried out the analyses. All authors contributed to developing and evaluating models and writing the article.

Acknowledgments

This research was supported by Grant-Aid for JSPS Fellows No. 17J07822 awarded to KO, by a Guggenheim Fellowship and a Fellowship from the Institute for Advanced Studies of Paris, awarded to LTM, and an ANR/NIH grant "Probabilistic models of perceptual grouping and segmentation in natural vision" (ANR-19-NEUC-0003-01) awarded to PM.

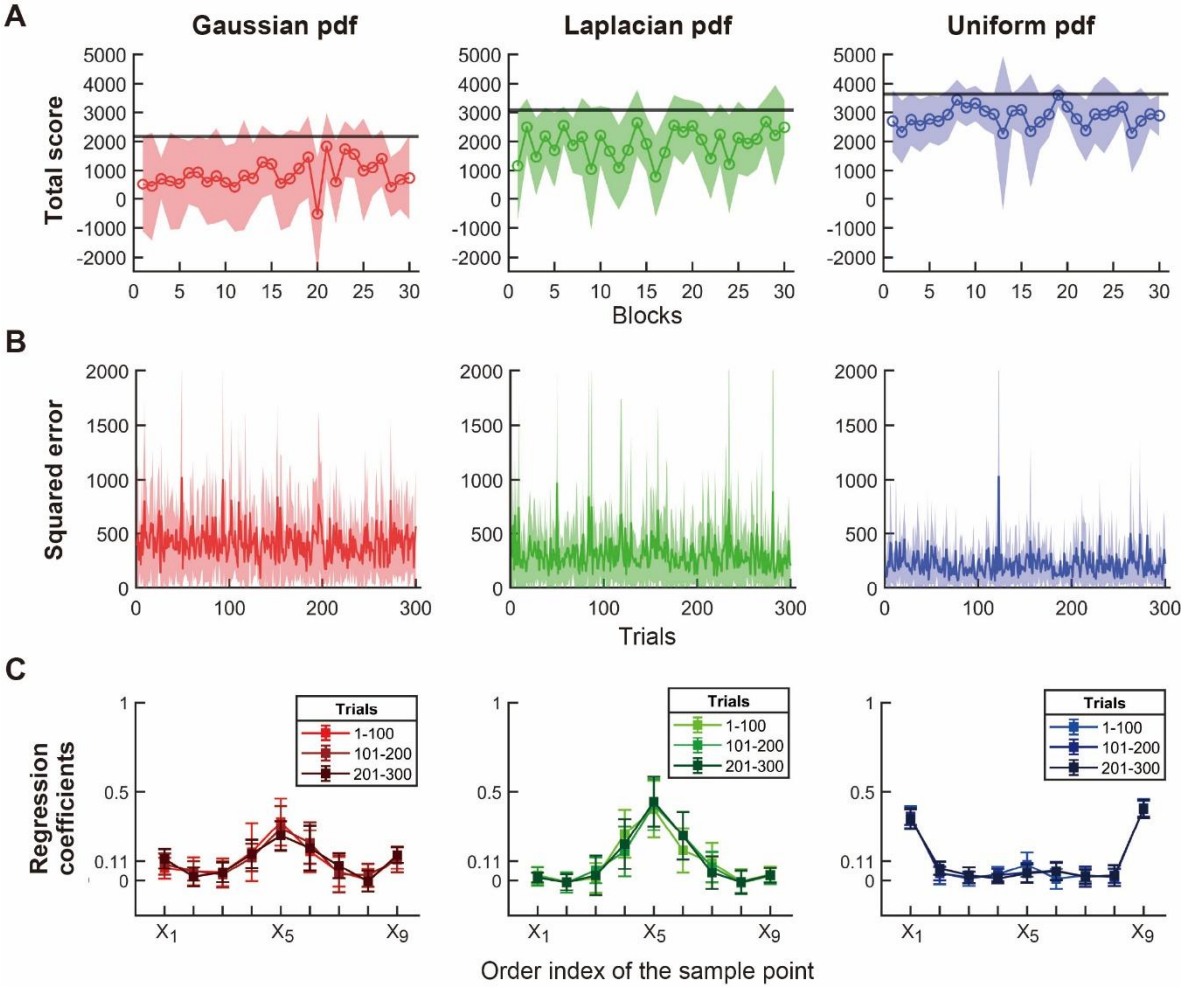
Supplementary Information



Supplementary Figure 1. Training task and visual location task.

A. Training task. Three distributional families were color-coded. On each trial, we presented one of the three pdfs and a sample drawn from the same pdf. The sample size reduced from $N = 300$ to $N = 9$ between blocks. Participants were asked to observe a color-coded sample drawn from the pdf with the same color code. **B. Visual location task**. On each trial, a sample of 9 points is drawn from one of the three distributional families with its center μ

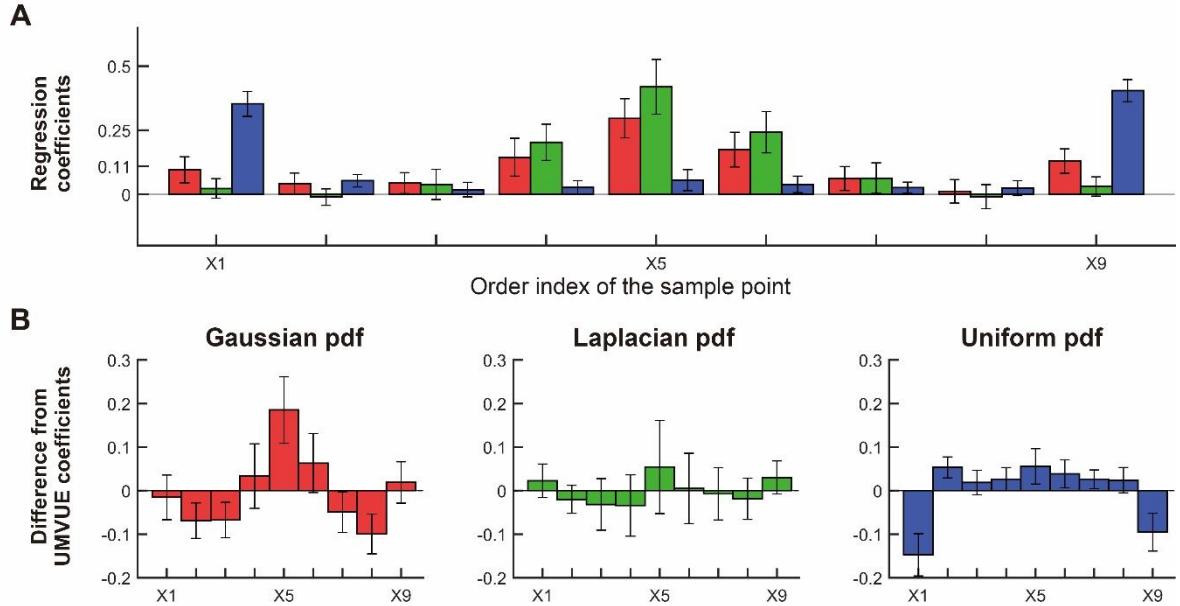
varied along the horizontal axis. Participants were asked to estimate the location of the invisible center of the pdf that had generated the sample by moving an arrow horizontally. Following a response, the true location of the center of the pdf was presented as a vertical line. A squared error and a score for a trial were also provided as feedback.



Supplementary Figure 2. Behavioral performance did not change systematically across trials.

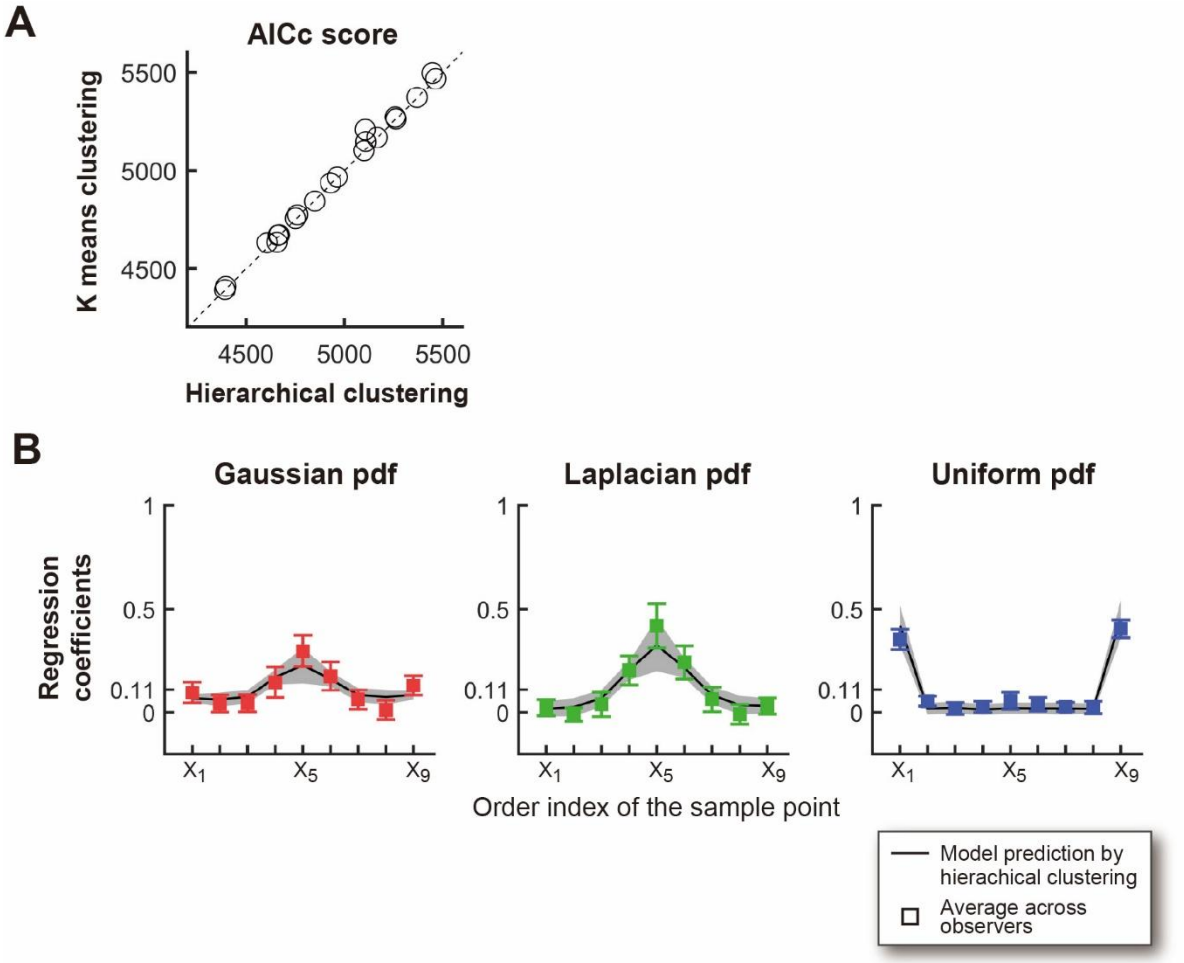
A. Total score. The participant's total score in each location task is plotted across blocks. Data are averaged across all participants. The shaded areas indicate ± 3 s.e.m. Black horizontal lines are the theoretically maximum total score obtained by UMVUE that corresponds to each distribution family. Two-way within participant ANOVA, using the number of blocks (30) and the distributional family (3) as independent variables, shows neither main effect of the block (F [299, 5681] = 1.11, p = 0.095, partial η^2 = 0.055) nor a significant interaction between the block and the distribution (F [598, 11362] = 1.08, p = 0.091, partial η^2 = 0.054). **B. Squared error.** The squared error between the participant's visual estimate and the true location of the center of the distribution is plotted versus trial. Two-way within-participant ANOVA, using the number of trials (300) and the distributional family (3) as independent variables, shows neither main effect of the block (F [29, 551] = 1.43, p = 0.069, partial η^2 = 0.070) nor a significant interaction between the block and the distribution (F [58, 1102] = 1.19, p = 0.155, partial η^2 = 0.059). **C. Estimated weights**

(Influence). To confirm a lack of learning effect over trials, we performed the linear regression separately for the first, the second, and the last 100 trials. The estimated weights are averaged across all participants. The error bars indicate ± 3 s.e.m. A two-way within participant ANOVA, using the order of points (9) and the bin of trials (3) as independent variables, shows no significant interaction between the order index of the sample point and the trial bin in the Gaussian location task (F [16, 304] = 0.70, p = 0.796, partial η^2 = 0.035), in the Laplacian location task (F [16, 304] = 0.74, p = 0.751, partial η^2 = 0.038), nor in the Uniform location task (F [16, 304] = 1.22, p = 0.250, partial η^2 = 0.060).



Supplementary Figure 3. Pairwise comparisons of the estimated weights.

A. The values of the estimated weights (influence measure) from the data are taken from Figure 4 and replotted as a bar graph as a function of the order statistics. **B.** Difference between the estimated weights and the weights employed by the *UMVUE* (*for the Laplacian the UMVUE*$_L$). In all panels, bars represent the average across participants. The error bars indicate ± 3 s.e.m.



Supplementary Figure 4. Comparison of the model fit across two clustering algorithms. A. A comparison of AICc scores. We adopted K-means clustering in the main text to fit the visual cluster model to the data. We repeated the clustering algorithm analysis using hierarchical clustering with single linkage. Here a grouping parameter $\boldsymbol{K}$ (i.e., the number of clusters) in K-means clustering was replaced by a grouping parameter $\boldsymbol{\theta} \in [\mathbf{0}, \mathbf{150}]$ in hierarchical clustering. This parameter is a cutoff point in millimetre to consider grouping two sample points. AICc scores are almost identical between the two clustering algorithms. **B. Prediction of the estimated weights (Influence).** Estimated regression coefficients (influence) estimated from the predictions of the VCM with hierarchical clustering. Compare the results with Figure 6. **C. Prediction of the cluster size using hierarchical clustering.** The number of clusters in a sample of 9 points was retrieved from the predictions of the VCM with hierarchical clustering for each trial and each observer and was plotted as a histogram for each location task.

References

Acerbi, L., & Ma, W. J. (2017). Practical Bayesian optimization for model fitting with Bayesian adaptive direct search. *Advances in neural information processing systems*, *30*, 1834-1844.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436. https://doi.org/https://doi.org/10.1163/156856897X00357

Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference* (2nd ed.). Springer.

Dal Martello, M. F., Ota, K., Pietralla, D. E., & Maloney, L. T. (2023). Detecting visual texture patterns in binary sequences through pattern features. *J Vis*, *23*(13), 1. https://doi.org/10.1167/jov.23.13.1

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429-433. https://doi.org/10.1038/415429a

Hampel, F. R., Ronchetti, E. M., & Rousseeuw, P. J. (1986). *Robust Statistics: The Approach Based on Influence Functions*. Wiley.

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. *Applied Statistics*, *28*(1). https://doi.org/10.2307/2346830

Huber, P. J. (1981). *Robust_statistics*. Wiley.

Hurvich, C. M., & Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika*, *76*(2), 297-307. https://doi.org/10.1093/biomet/76.2.297

Juni, M. Z., Singh, M., & Maloney, L. T. (2010). Robust visual estimation as source separation. *J Vis*, *10*(14), 2. https://doi.org/10.1167/10.14.2

Koffka, K. (1935/2013). *Principles of Gestalt psychology*. Routledge.

Landy, M. S., Goutcher, R., Trommershauser, J., & Mamassian, P. (2007). Visual estimation under risk. *J Vis*, *7*(6), 4. https://doi.org/10.1167/7.6.4

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Res*, *35*(3), 389-412. https://doi.org/10.1016/0042-6989(94)00176-m

Lehmann, E. L. (1998). *Theory of point estimation* (2nd ed.). Springer.

Lieder, F., & Griffiths, T. L. (2019). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behav Brain Sci*, *43*, e1. https://doi.org/10.1017/S0140525X1900061X

Maloney, L. T., & Landy, M. S. (1989). A statistical framework for robust fusion of depth information. *Proceedings of the SPIE: Visual Communications and Image Processing IV*(1199), 1154-1163.

Maloney, L. T., & Thomas, E. A. C. (1991). Distributional assumptions and observed conservatism in the theory of signal detectability. *Journal of Mathematical Psychology*, *35*(4), 443-470. https://doi.org/10.1016/0022-2496(91)90043-s

McGowan, J. W., Kowler, E., Sharma, A., & Chubb, C. (1998). Saccadic localization of random dot targets. *Vision Res*, *38*(6), 895-909. https://doi.org/10.1016/s0042-6989(97)00232-0

Melcher, D., & Kowler, E. (1999). Shapes, surfaces and saccades. *Vision Res*, *39*(17), 2929-2946. https://doi.org/10.1016/s0042-6989(99)00029-2

Mood, A. M., Graybill, F. A., & C., B. D. (1974). *Introduction to the theory of statistics* (3rd ed.). McGraw-Hill.

Nielsen, F. (2016). Hierarchical Clustering. In *Introduction to HPC with MPI for Data Science* (pp. 195-211). https://doi.org/10.1007/978-3-319-21903-5_8

Oruç, İ., Maloney, L. T., & Landy, M. S. (2003). Weighted linear cue combination with possibly correlated error. *Vision Research*, *43*(23), 2451-2468. https://doi.org/10.1016/s0042-6989(03)00435-8

Ota, K., Charles, L., & Haggard, P. (2024). Autonomous behaviour and the limits of human volition. *Cognition*, *244*, 105684. https://doi.org/10.1016/j.cognition.2023.105684

Ota, K., Christofilea, E., Charles, L., Daunizeau, J., & Haggard, P. (2025b). Freedom through understanding: instructed knowledge shapes voluntary action choices. *Preprint at PsyArXiv*. https://doi.org/https://doi.org/10.31234/osf.io/mgksv_v2

Ota, K., Ciston, A., Haggard, P., Gajdos, T., & Charles, L. (2025a). Integrating explicit reliability for optimal choices: effect of trustworthiness on decisions and metadecisions. *Preprint at bioRxiv*. https://doi.org/https://doi.org/10.1101/2025.01.12.632598

Ota, K., & Maloney, L. T. (2024). Dissecting Bayes: Using influence measures to test normative use of probability density information derived from a sample. *PLoS Comput Biol*, *20*(5), e1011999. https://doi.org/10.1371/journal.pcbi.1011999

Ota, K., Shinya, M., & Kudo, K. (2016). Sub-optimality in motor planning is retained throughout 9 days practice of 2250 trials. *Sci Rep*, *6*, 37181. https://doi.org/10.1038/srep37181

Ota, K., Shinya, M., Maloney, L. T., & Kudo, K. (2019). Sub-optimality in motor planning is not improved by explicit observation of motor uncertainty. *Sci Rep*, *9*(1), 14850. https://doi.org/10.1038/s41598-019-50901-x

Ota, K., Tanae, M., Ishii, K., & Takiyama, K. (2020). Optimizing motor decision-making through competition with opponents. *Sci Rep*, *10*(1), 950. https://doi.org/10.1038/s41598-019-56659-6

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*(4), 437-442. https://doi.org/10.1163/156856897x00366

Rensink, R. A. (2017). The nature of correlation perception in scatterplots. *Psychon Bull Rev*, *24*(3), 776-797. https://doi.org/10.3758/s13423-016-1174-7

Reverdy, P., & Leonard, N. E. (2016). Parameter Estimation in Softmax Decision-Making Models With Linear Objective Functions. *IEEE Transactions on Automation Science and Engineering*, *13*(1), 54-67. https://doi.org/10.1109/tase.2015.2499244

Rigoux, L., Stephan, K. E., Friston, K. J., & Daunizeau, J. (2014). Bayesian model selection for group studies - revisited. *Neuroimage*, *84*, 971-985. https://doi.org/10.1016/j.neuroimage.2013.08.065

Shetty, P., & Singh, S. (2021). Hierarchical Clustering: A Survey. *International Journal of Applied Research*, *7*(4), 178-181. https://doi.org/10.22271/allresearch.2021.v7.i4c.8484

Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., & Friston, K. J. (2009). Bayesian model selection for group studies. *Neuroimage*, *46*(4), 1004-1017. https://doi.org/10.1016/j.neuroimage.2009.03.025

Sun, J., Li, J., & Zhang, H. (2019). Human representation of multimodal distributions as clusters of samples. *PLoS Comput Biol*, *15*(5), e1007047. https://doi.org/10.1371/journal.pcbi.1007047

Tanae, M., Ota, K., & Takiyama, K. (2021). Competition Rather Than Observation and Cooperation Facilitates Optimal Motor Planning. *Frontiers in Sports and Active Living*, *3*. https://doi.org/10.3389/fspor.2021.637225

Trommershauser, J., Kording, K., & Landy, M. S. (2011). *Sensory cue integration*. Oxford University Press.

Trommershauser, J., Maloney, L. T., & Landy, M. S. (2008). Decision making, movement planning and statistical decision theory. *Trends Cogn Sci*, *12*(8), 291-297. https://doi.org/10.1016/j.tics.2008.04.010

Tufte, E. R. (2001). *The visual display of quantitative information* (2nd ed.). Graphics Press.

Tukey, J. W. (1977). *Exploratory data analysis*. Pearson.

Ullman, S. (1984). Visual routines. *Cognition*, *18*(1-3), 97-159. https://doi.org/10.1016/0010-0277(84)90023-4

Ullman, S., Vidal-Naquet, M., & Sali, E. (2002). Visual features of intermediate complexity and their use in classification. *Nat Neurosci*, *5*(7), 682-687. https://doi.org/10.1038/nn870

van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial comparison of working memory models. *Psychol Rev*, *121*(1), 124-149. https://doi.org/10.1037/a0035234