Title
Models of human visual clustering

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors
Marupudi, Vijay
Varma, Sashank

Publication Date
2023

Peer reviewed

Models of human visual clustering

Vijay Marupudi (vijaymarupudi@gatech.edu)
School of Interactive Computing
Atlanta, GA 30308 USA

Sashank Varma (varma@gatech.edu)
School of Interactive Computing
Atlanta, GA 30308 USA

Abstract

How humans visually cluster points is relevant for perception, information visualization, and many other domains. However, there are relatively few empirical studies and computational models of this ability. Here, we propose a new *competitive clustering* model that uses neurally plausible mechanisms such as hebbian learning and lateral inhibition. We evaluate its fit to the data from a behavioral study of visual clustering, as well as the fits of two categorization models (the Rational model and SUSTAIN) and one statistical learning algorithm (K-Means). We find that people are highly reliable in their clusterings of the same stimulus on two occasions, suggesting they are using a stable strategy. The models were generally successful at predicting human clusterings and were able to replicate the qualitative performance profile of human reliability over different numbers of points and levels of cluster structure. The performance of the competitive clustering model motivates further investigation of its computational properties and empirical validity.

Keywords: clustering, perception, numerosity estimation, K-Means, the Rational model, SUSTAIN

Introduction

Human visual clustering can be defined as dividing a set of points into "similar" groups. There have been relatively few studies and models of this ability. This is an important gap because visual clustering is highly relevant for perception, education, problem-solving, and information visualization. A scientific understanding of human visual clustering could be used to help assess students' pattern recognition in graphs (e.g., scatterplots) and ultimately to develop instructional materials to improve this understanding. It could also be used to inform the design of information visualizations and assessment of their visualization quality (Sips et al., 2009). It could also contribute to a better understanding of problem decomposition in human spatial decision-making, in contexts such as navigation or human solutions of the traveling salesperson problem (Marupudi et al., 2022).

A consideration of human visual clustering immediately suggests a number of interesting questions: Do people constrain themselves to a preferred number of clusters? How large are people's clusters relative to each other? How sensitive are people to the statistical structure of stimuli (i.e., relatively clustered vs dispersed)? How reliable are people in their clustering strategies? At the moment, these questions lack convincing answers in the literature. The current research begins to address this gap by proposing a new model of human clustering, the *competitive clustering* model, and comparing its ability to predict the clusters people draw to other unsupervised models of clustering from cognitive science and computer science.

Related theoretical proposals and empirical research

Early principles relevant for human visual clustering can be found in the Gestalt research tradition (Wertheimer, 1938). For example, the law of proximity states that objects that are close together are perceived as being part of a single group. To take another example, the law of continuity suggests that objects positioned in a continuous configuration such as a line or a curve are considered to be part of a single group. Many Gestalt principles have been demonstrated in carefully controlled experiments. However, they have not been formalized in models of human visual clustering that have been empirically evaluated against human clustering data.

Research relevant for clustering also comes from studies of numerosity estimation, where people report the number of points they see in point clouds similar to clustering stimuli. Research shows that people use many visual properties in parallel to make such judgments, with increasing redundancy of cues resulting in more accurate numerosity estimates (Bertamini et al., 2016; Gebuis & Gevers, 2011; Gebuis & Reynvoet, 2012).

In addition, studies have shown that people perceive regular patterns as more numerous (Ginsburg, 1980) and perceive stimuli with points arranged as clusters as less numerous than a single point cloud.

Studies and computational models of clustering

Designing an accurate model of visual clustering is a difficult task. One reason is that clustering is an ill-defined task with no right answer, making it difficult to evaluate the quality of solutions. One could legitimately perceive a stimulus as having any number of clusters of varying sizes, numerosities, shapes, and densities. Another barrier is the lack of studies of visual human clustering, with most work focusing on predicting the number of clusters (Aupetit et al., 2019; Im et al., 2016) or suggesting that people perceive Gaussian clusters (Sedlmair et al., 2012). Finally, it is also unclear if participants use a consistent strategy for their clustering, i.e., whether their clusterings are reliable.

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*.

There have been two main attempts to model human visual clustering. The first model, CODE, was proposed by van Oeffelen and Vos (1982). CODE involves centering a normal distribution for each point in a stimulus with a standard deviation that is half the distance to the point's nearest element. Clusters are then determined by using a threshold to partition regions of the cumulative distribution of all the points. An extensive analysis of the algorithm's various assumptions used exact clustering matches between participants and the model as a metric of algorithm performance (Compton & Logan, 1993). This study was an important first step in the study of human clustering. Here, we extend their analysis by using a more graded metric of cluster similarity and directly manipulating the cluster structure of the stimuli.

The current study expands the set of models of human visual clustering with a new model, competitive clustering, that is more neurally plausible than prior models. It also considers K-means, a standard statistical clustering algorithm, and two existing models that characterize human categorization as driven by clustering: SUSTAIN (Love et al., 2004) and the Rational model (Anderson, 1991). Both of these models contain an internal representation that clusters observations in order to categorize them, and both have completely unsupervised operating modes. They have been successful at explaining human categorization behavior. Here, we ask whether people use similar statistical learning abilities to cluster stimuli.

In this study, we address the following research questions:

- How reliable is human clustering? How do participants' clusterings of the same stimulus compare across different time points?
- How does pre-existing cluster structure impact clustering reliability?
- How well does the new competitive clustering model emulate human visual clustering?
- How well do two models of categorization and unsupervised learning, SUSTAIN and the Rational model, explain human visual clustering?
- How well do these models compare to K-Means, a standard clustering algorithm, at predicting human clustering?

Methods

We recruited 47 (M = 11, F = 38) undergraduate students at a large university in the Midwestern US. Of these, 21 completed the study in the lab. Following the onset of the COVID-19 pandemic, the remaining 26 participants completed the study online using the same interface. This study followed a $8 \times 2 \times 2$ within-subjects design. Number of Points had 8 levels (10, 15, 20, 25, 30, 35, 40) and statistical Cluster Structure had 2 levels (Clustered, Dispersed). Each condition contained four unique stimuli. Half of these stimuli were flipped when participants clustered them for the second time; this was the third factor. This design allowed us to evaluate whether increasing the complexity of clustering by increasing the number of points or reducing the cluster structure impacted participants' cluster quality and strategies.

Materials

We generated 4 stimuli for each level of Number of Points and Cluster Structure, resulting in a total of 56 unique stimuli. For each stimulus, we continuously generated candidate stimuli by sampling the appropriate number of points from a uniform random distribution onto an 800 by 500 pixels (width x height) space. We then filtered these candidates for appropriate values of the *vacuumed standardized Z score*, the cluster structure metric defined by Marupudi et al. (2022). This process continued until the desired number of stimuli were generated. Similarly to Marupudi et al. (2022), we defined Clustered and Dispersed stimuli as having vacuumed standardized Z scores of -2 ± 0.05 and 1 ± 0.05 respectively. These values were chosen so that stimuli of both cluster structures were qualitatively different while remaining difficult to notice during the experiment. Pilot testing confirmed that participants did not notice any patterns in cluster structure.

Procedure

After providing consent, participants first viewed instructions and completed three practice trials before moving on to the 56 experimental trials. On each trial, they were asked to "draw circles around clusters in the stimuli". Those who drew a single cluster around the entire stimulus were asked to cluster the stimulus again. To begin drawing clusters, participants clicked any spot on the stimulus, dragged their cursor around the outer edge of the cluster, and released the cursor to complete the shape they started. Points were initially displayed as black circles on a white background. However, points that were subsequently included in a cluster turned blue, allowing participants to easily identify which points remained unclustered.

After an unrelated distractor task, participants again clustered the same 56 experimental stimuli. These stimuli were randomly shuffled. Half were flipped vertically and horizontally from their initial presentation. Examples of participants' clusterings are shown in Figure 1.

Cluster similarity metric: Fowlkes-Mallows (FM) index

Participants clustered all 56 stimuli twice. We used these pairs of clusterings of identical stimuli to investigate the reliability of participants' clusterings. High clustering similarity between pairs of clusterings would suggest that participants were reliable and followed consistent strategies to determine visual clusters and/or that they are consistently sensitive to specific stimulus properties such as density, area, etc.

For our measure of clustering similarity, we chose the Fowlkes-Mallows (FM) index (Fowlkes & Mallows, 1983). It conveniently ranges between 0 and 1 and also accounts for chance. The FM trends towards 0 with increasing number of points in the stimulus for random clusterings. It is defined as:

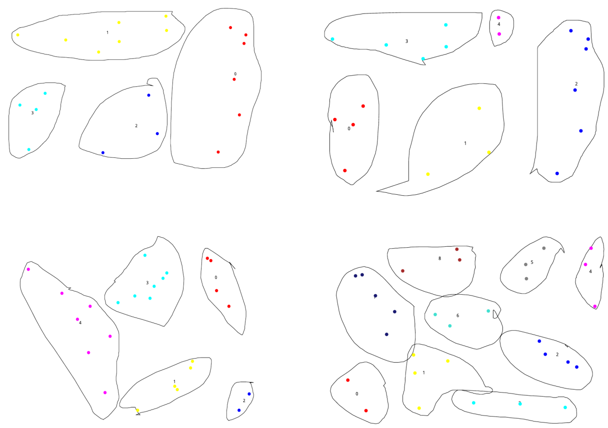

Figure 1: Example stimuli clustered by participants. Black lines represent cluster borders drawn by participants. Colors indicate cluster membership of the points and were not present in the stimuli clustered by participants. Numbers indicate the sequence that participants followed to cluster the stimuli.

$$\sqrt{\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \cdot \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}}$$

where TP and FP represent the number of true and false positives respectively, and FN represents the number of false negatives. Higher values of the FM index indicate higher clustering similarity and lower values indicate lower similarity.

Competitive clustering model

The competitive clustering algorithm is a competitive recurrent neural network (Grossberg, 1973; Rumelhart & Zipser, 1985) and can be seen as a modified version of a self-organizing map with additional lateral inhibition (Kohonen, 1982). It uses neurally plausible mechanisms such as hebbian learning, lateral inhibition, and finding the n most active neurons from a set of neurons. The algorithm relies on a k-winner-take-all computational principle that has successfully modeled data from the inferotemporal cortex during object recognition (Riesenhuber & Poggio, 1999). Approaches using similar operations and computational principles have also been shown to be implementable using spiking neural networks (Oster et al., 2009), supporting the claim of neural plausibility.

Like the K-Means algorithm discussed below, competitive clustering attempts to find the locations of cluster centroids for a set of points. Unlike K-Means, the number of clusters is dynamically determined by the activation and lateral inhibition mechanisms governed by model parameters.

The model continuously samples random observations from the stimulus to a group of neurons. It then allows the temporal dynamics of the neurons to come to convergence, allowing single neurons to survive that correspond to regions of an image. Applying a threshold operation ($w > 0.5$) then gives the centroids of the clusters.

In greater detail: For each stimulus, the model begins with an initial set of 100 neurons. Each neuron represents three numbers: x and y coordinates and also an activation a ranging from 0 to 1. The x and y values are initialized using a uniform random distribution while a is set to 0.5. The network is not particularly sensitive to the initial configuration as long as the neurons approximately cover the range of the observations.

The model then applies two processes simultaneously that change the locations (i.e., dimensions) and activations of the neurons.

1. Hebbian learning: Observations of the input stimulus are presented to all the neurons in the network. For each neuron, its sensitivity s to an observation is determined by calculating the inverse of the Euclidean distance between the observation and the neuron's dimension values, and then multiplying by the neurons' activation.

   The dimension values of the top $k = 10$ neurons with the highest s values are updated to move closer to the observation, with the amount determined by a learning rate $\eta = 0.1$. The activation of the most sensitive neuron is also increased by $\Delta a = 0.5a$.

   The 90 neurons that were not part of the top k set receive a flat 0.1 reduction in activation, with a floor of 0.

2. Lateral inhibition: The goal of this process is to reduce the activations of neurons close to other existing neurons. Every neuron is connected to every other neuron in the network and can pass an inhibitory message proportional to the strength of its activation. Each neuron reduces the activation of the k neurons that are sensitive (s) to its dimension value (using the same criteria as Hebbian learning) by $\Delta a = -sa$.

Intuitively, these processes move the dimension values of neurons closer to high density parts of the input stimulus, while reducing the activation of neurons that aren't particularly close to any of the observations. At the high density parts of the stimulus, the lateral inhibition mechanism results in a winner-take-all neuron.

After 14 iterations through the dataset, the neurons with activations greater than 0.5 are used as the centroids of the clusters, and the points of the stimulus are assigned to the cluster represented by the closest centroid.

For the current study, we used fixed parameter values and a single global variant to demonstrate the proof-of-concept clustering capabilities of the model. We did *not* compute separate parameter fits for each participant for the competitive clustering model. We defer such parameter fitting and further experiments to future work.

Rational model

The rational model was first presented in Anderson (1991). Unlike other models of categorization (for e.g., the prototype model and the exemplar model), the rational model learns to form clusters from observations using Bayes theorem. We chose the rational model as it represents an "optimal" sequential model of clustering where participants iteratively decide whether to expand their current clusters or create new clusters. A recent proof by Dasgupta and Griffiths (2022) suggests that the rational model can be seen as an optimal clustering algorithm that aims to maximize clustering fit while minimizing representational complexity, i.e., number of clusters.

The model begins with no clusters. When presented with a new point, it compares the probability that it belongs to an existing cluster with the probability that a new cluster should be generated. This latter probability is influenced by a coupling parameter c giving the prior probability that any new point comes from a new cluster.

In greater detail, our implementation follows the specification outlined in Anderson (1991) for continuous dimensions using a normal distribution centered on the mean prototype of the observations in a cluster. We fix the prior confidence parameters λ_0 and a_0 to 1 and set the prior variance of a new cluster to be 1/4 of the range of each dimension: for our stimuli, 800 / 4 for the width and 500 / 4 for the height. We used the Virtanen et al. (2020) `curve_fit` non-linear least square parameter optimization function to maximize the FM metric of clustering similarity with all trials of the human data, resulting in a c value of 0.7. For the current study, we ran one global variant of the rational model with the same parameter c for each participant.

SUSTAIN

Similarly to the rational model, SUSTAIN is a model of human categorization that organizes observations as clusters. It aims to prevent prediction failures by only recruiting a new cluster for an observation if activation of the existing clusters for an observation is below a threshold parameter. SUSTAIN has been shown to outperform many prominent models of categorization on a multitude of tasks (Love et al., 2004). We chose this model because its parameters have the potential to be useful for capturing individual differences at the clustering task.

Our implementation of the model used only the unsupervised clustering mechanism of the model outlined in Love et al. (2004). We considered two variants. The global variant used a set of parameters fit across all participants to maximize the median FM score. The per-participant variant was fit independently to maximize the median FM score for each participant. The parameters for the model were fit using a genetic algorithm optimization procedure.

K-Means

K-Means is a standard clustering algorithm in statistical and machine learning. It is parameterized by a number of clusters, k. It generates clusters with random centroids and assigns the closest points to the clusters. This makes the algorithm sensitive to the initial centroids. The algorithm then iteratively recalculates the value of the cluster centroids by calculating the mean of all the points in the clusters and then reassigning the closest points to the clusters (which can be different as the centroid of the clusters might have changed). This process terminates when the cluster centroids do not change. *K*-Means attempts to minimize the variance within each cluster, producing a locally optimal solution (the optimal algorithm is an NP-Hard problem).

We considered three ways to determine the value of k for *K*-Means, resulting in three variants of the model. The global model used the median number of clusters across all participants and all trials. The per-participant model used the median number of clusters across each participants' trial. Finally, the per-trial model used the number of clusters in the participant's clustering of each stimulus. We used the scikit-learn implementation of the Lloyd *K*-Means model (Pedregosa et al., 2011).

Because all models are either sensitive to the ordering of the stimuli in a stimulus or randomly vary in their initial clusterings, we shuffled and clustered each stimulus 10 times and used the median FM value as the final FM value of the model for a specific trial.

Results

Human performance

We first investigated whether participants' clusters were reliable, i.e., whether their second clustering of a stimulus was similar to their first clustering of the same stimulus. We calculated the FM index for the two clusterings, using their first clustering as the reference. Participants were remarkably consistent at clustering in general, with a mean FM index of 0.757. (Recall this index ranges between 0 and 1, with 1 indicating identical clusterings on the two occasions.) This high value lends support to the hypothesis that participants have reliable clustering strategies and justifies the effort to develop a model of human visual clustering.

We then compared the mean FM values between pairs of similarly oriented stimuli and pairs where the second stimulus was flipped. Flipped pairs were slightly less reliable than pairs with similar orientation (Flipped FM: 0.74, Same FM: 0.76). While this difference is significant, it explained very little variance in the data ($R^2 = 0.01$) and did not interact with either Number of Points or Cluster Structure. Therefore, we did not consider this factor in subsequent analyses.

To determine the effects of Number of Points and Cluster Structure on within-participant clustering reliability (i.e., when participants cluster the same stimulus a second time), we first aggregated the data set by computing the mean FM index for each participant for each value of Number of Points and each level of Cluster Structure, to reduce stimulus-specific variability; see the top-left panel of Figure 2. We then fit a linear regression model predicting the FM index using Number of Points, Cluster Structure, and their interaction. The model was significant ($F(3,654), p < 0.001$) and explained 15.7% of the variance. There was a main effect of Cluster Structure, with the FM index for dispersed stimuli -0.13 less than that for clustered stimuli ($t = -6.72, p <$

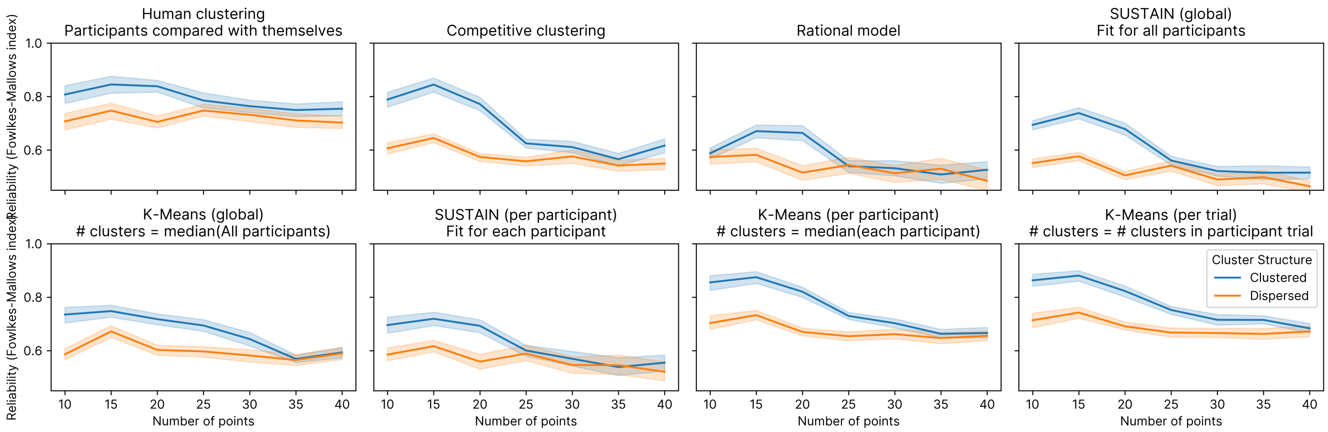

Figure 2: Clustering fits over Number of Points and the Cluster Structure of the stimuli. Top-left panel describes human compared to themselves, i.e., first clustering compared to their second clustering.

$0.001, \eta_p^2 = 0.11$). This suggests that statistical cluster structure plays a role in the variability of participants' clustering. Nevertheless, it is important to note that mean FM index for dispersed stimuli was still relatively high ($M = 0.72$), suggesting overall high reliability in clustering. There was also a main effect of Number of Points, where each additional point reduced the FM index by -0.003 ($t = -5.743, p < 0.001, \eta_p^2 = 0.028$). This implies that adding 10 points to a stimulus would reduce the FM index by 0.03. Finally, the interaction term in the model was significant ($\beta = 0.0026, t = 3.479, p = 0.001, \eta_p^2 = 0.016$). This suggests that Number of Points reduces the FM index less for dispersed stimuli compared to clustered stimuli. This is likely due to the high reliability for clustered stimuli at lower number of points; increasing number of points makes cluster structure less relevant for clustering reliability.

Model evaluation

There are multiple ways a model of clustering can be a good fit to human behavior. First, the clusterings of the model should be similar to the clusterings of the human participants. Second, the model should show the same reliability performance profile as human participants. Finally, the model should generate the same number of clusters as participants do for each stimulus. We tested the 7 models variants on these measures using the FM index as a metric of clustering similarity.

First, we compared the models on their ability to generate clusterings that are similar to humans. All models did relatively well., i.e., all had FM index values greater than 0.54; see Figure 3. The models are grouped left-to-right by increasing parameterization: competitive clustering, the Rational model, SUSTAIN (global), and K-Means (global) all use the same parameters for all participants; the remaining model variants use more specific parameters for each participant and/or each trial.

We then fit a linear model predicting the FM index with

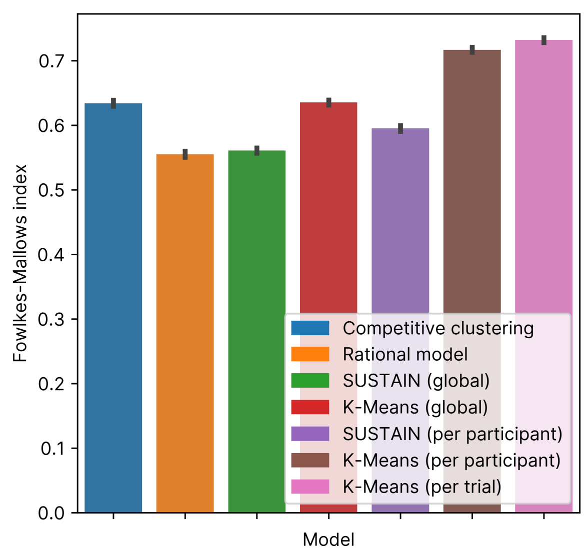

Figure 3: Clustering model predictions compared to human clustering data. Higher values of Fowlkes-Mallows indicate more human-like clusters. Error bars = SE.

the model used as the predictor variable; this explained 26% of the variance ($p < 0.001, F(6, 4599) = 273.7$). Post-hoc Tukey HSD tests with Benjamini-Hochberg p value correction revealed that all the models differed, i.e., $p < .05$ from one another with two exceptions: competitive clustering and the K-Means (global) model had comparable FM index values, as did the Rational model and SUSTAIN (global). The K-Means models with trial and participant parameters fit human data the best followed by the Competitive learning and the K-Means (global) model. Next, SUSTAIN (per participant) performs well, while SUSTAIN (global) and the Rational model are last in fitting human data.

We then asked whether the models' predictions of human clusters were correlated with human clustering reliability on the same stimuli. Qualitatively, it appears that every model tested replicates the difference in cluster reliability between clustered and dispersed stimuli (Figure 2). To test the performance profiles more precisely, we correlated the data for every model's panel in Figure 2 with human reliability data in the top-left panel. Higher correlation values indicate better modeling of the factors than determine variability in human clustering. In addition, this analysis tests whether models' sensitivity to the Number of Points and Cluster Structure in a stimuli correspond to human behavior. All correlations were significant ($p < 0.001$). The correlation results are displayed in Figure 4. Similarly to the overall model fit, K-Means (per trial) and K-Means (per participant) showed the highest correlation to the human reliability profile. Competitive clustering and the K-Means (global) models were next highest. Both SUSTAIN models correlated around 0.36, while the Rational model correlated poorly with the human performance profile, with $r = 0.18$.

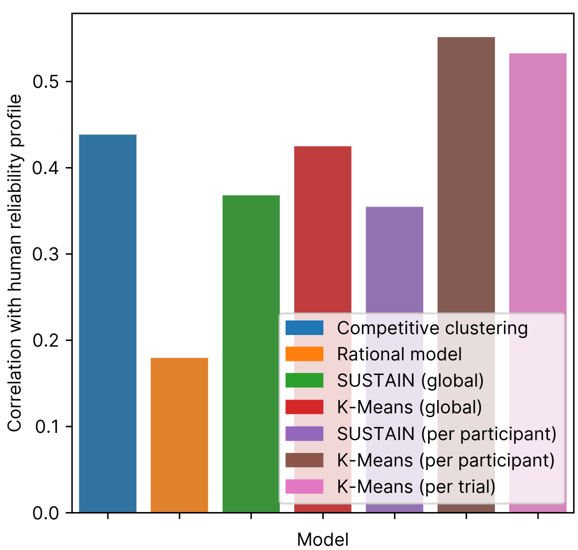

Figure 4: Correlations of the performance profiles of the models across Number of Points and Cluster Structure with human reliability data

Finally, we investigated whether the clusterings generated by the various models contained similar number of clusters as human clusterings. For each model, we correlated the number of clusters for each trial with the model's predictions. The K-Means models were excluded from this analysis because their number of clusters were explicitly determined from the empirical data. The competitive clustering and the SUSTAIN (global) models did not predict the number of clusters well ($p > 0.05$). Interestingly, the predictions of the SUSTAIN (per participant) model ($r = 0.42$) and the Rational model ($r = 0.22$) were significantly correlated with the human data.

Discussion

In this study, we presented experimental evidence that people are reliable when clustering points, and that they are more reliable when clustering points that are statistically clustered (i.e., less dispersed). This suggests that participants may be using a reliable strategy to cluster the points. To determine the strategy, we proposed a new competitive clustering model and compared its fit to the human data with two prominent models of unsupervised learning, the Rational model (Anderson, 1991) and SUSTAIN (Love et al., 2004), as well as K-Means. We ran variants of SUSTAIN and K-Means with parameters tuned for each participant and/or trial. The models that did not include these parameters were termed "global" models; they had less flexibility to fit the human data. We found that all models were generally successful at predicting human clusterings and succeeded in replicating the qualitative performance profile of human reliability over Number of Points and Cluster Structure (Figure 2). Nevertheless, some models performed better than others.

The competitive clustering model was the highest performing global model (i.e., models with the same parameters for all trials). This finding motivates further development of the model and investigation of its computational properties and empirical validity. It performs similarly to the K-Means (global) model in predicting participants' clusterings and following their reliability profile. However, it does not do very well at predicting the number of clusters drawn by participants. That said, the only model that performed well at this task was SUSTAIN (per participant) model, suggesting the presence of individual differences between participants and the need to fit parameters for each participant to give a good account of the data on this task.

Although SUSTAIN and the Rational model did not perform as well as the competitive clustering and K-Means models, both models did quite well at predicting the number of clusters drawn by participants. Future work could compare how well they perform compared to the dynamic K-Means model proposed by Im et al. (2016), which was purpose-built for that task. It is possible that their performance is impaired by the fact that they do not use Euclidean distance as the distance metric, and because they assume a specific distribution for the clusters. Modifications of these models that incorporate these features and/or allow for multiple distributions might improve their performance. An additional model to consider for future work would be the neural gas model, which shares many properties with the competitive clustering model (Martinetz & Schulten, 1991).

References

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, *98*, 409–429. https://doi.org/10.1037/0033-295X.98.3.409

Aupetit, M., Sedlmair, M., Abbas, M. M., Baggag, A., & Bensmail, H. (2019). Toward perception-based evaluation of clustering techniques for visual analytics. *2019 IEEE Visualization Conference (VIS)*, 141–145.

Bertamini, M., Zito, M., Scott-Samuel, N. E., & Hulleman, J. (2016). Spatial clustering and its effect on perceived clustering, numerosity, and dispersion. *Attention, Perception & Psychophysics*, *78*, 1460–1471. https://doi.org/10.3758/s13414-016-1100-0

Compton, B. J., & Logan, G. D. (1993). Evaluating a computational model of perceptual grouping by proximity. *Perception & Psychophysics*, *53*(4), 403–421. https://doi.org/10.3758/BF03206783

Dasgupta, I., & Griffiths, T. L. (2022). Clustering and the efficient use of cognitive resources. *Journal of Mathematical Psychology*, *109*, 102675. https://doi.org/10.1016/j.jmp.2022.102675

Fowlkes, E. B., & Mallows, C. L. (1983). A Method for Comparing Two Hierarchical Clusterings. *Journal of the American Statistical Association*, *78*(383), 553–569. https://doi.org/10.1080/01621459.1983.10478008

Gebuis, T., & Gevers, W. (2011). Numerosities and space; indeed a cognitive illusion! A reply to de Hevia and Spelke (2009). *Cognition*, *121*(2), 248–252. https://doi.org/10.1016/j.cognition.2010.09.008

Gebuis, T., & Reynvoet, B. (2012). The interplay between nonsymbolic number and its continuous visual properties. *Journal of Experimental Psychology: General*, *141*, 642–648. https://doi.org/10.1037/a0026218

Ginsburg, N. (1980). The Regular-Random Numerosity Illusion: Rectangular Patterns. *The Journal of General Psychology*, *103*(2), 211–216. https://doi.org/10.1080/00221309.1980.9921000

Grossberg, S. (1973). Contour Enhancement, Short Term Memory, and Constancies in Reverberating Neural Networks. *Studies in Applied Mathematics*, *52*(3), 213–257. https://doi.org/10.1002/sapm1973523213

Im, H. Y., Zhong, S.-h., & Halberda, J. (2016). Grouping by proximity and the visual impression of approximate number in random dot arrays. *Vision Research*, *126*, 291–307. https://doi.org/10.1016/j.visres.2015.08.013

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, *43*(1), 59–69. https://doi.org/10.1007/BF00337288

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A Network Model of Category Learning. *Psychological Review*, *111*, 309–332. https://doi.org/10.1037/0033-295X.111.2.309

Martinetz, T., & Schulten, K. (1991). A "neural-gas" network learns topologies. *Artificial Neural Networks*.

Marupudi, V., Harsch, R., Rao, V. N. V., Park, J., Bye, J. K., & Varma, S. (2022). Use of clustering in human solutions of the traveling salesperson problem. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *44*(44). Retrieved January 23, 2023, from https://escholarship.org/uc/item/3t16j57m

Oster, M., Douglas, R., & Liu, S.-C. (2009). Computation with Spikes in a Winner-Take-All Network. *Neural Computation*, *21*(9), 2437–2465. https://doi.org/10.1162/neco.2009.07-08-829

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*(11), 1019–1025. https://doi.org/10.1038/14819

Rumelhart, D. E., & Zipser, D. (1985). Feature discovery by competitive learning. *Cognitive Science*, *9*(1), 75–112. https://doi.org/10.1016/S0364-0213(85)80010-0

Sedlmair, M., Tatu, A., Munzner, T., & Tory, M. (2012). A Taxonomy of Visual Cluster Separation Factors. *Computer Graphics Forum*, *31*, 1335–1344. https://doi.org/10.1111/j.1467-8659.2012.03125.x

Sips, M., Neubert, B., Lewis, J. P., & Hanrahan, P. (2009). Selecting good views of high-dimensional data using class consistency. *Computer Graphics Forum*, *28*(3), 831–838. https://doi.org/10.1111/j.1467-8659.2009.01467.x

van Oeffelen, M. P., & Vos, P. G. (1982). Configurational effects on the enumeration of dots: Counting by groups. *Memory & Cognition*, *10*(4), 396–404. https://doi.org/10.3758/BF03202432

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., . . . SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods*, *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Wertheimer, M. (1938). Laws of organization in perceptual forms.