Perceived clusters may not explain people's judgments of approximate numerosity

Vijay Marupudi (vijaymarupudi@gatech.edu)
School of Interactive Computing, Georgia Tech, Atlanta, Georgia, United States
Department of Psychology, UC Berkeley, Berkeley, California, United States

Sashank Varma (varma@gatech.edu)
School of Interactive Computing & School of Psychology, Georgia Tech, Atlanta, Georgia, United States

V. N. Vimal Rao (raovnv@illinois.edu)
Department of Statistics, University of Illnois, Champaign, Illinois, United States

Abstract

The approximately number system (ANS) helps people quickly estimate the numerosity of objects in their environment. In this study, we explore one proposed mechanism for visually perceiving numerosity: visual clustering. Participants completed a magnitude comparison task, magnitude estimation task, and a clustering task using the same set of numerosity stimuli. The stimuli varied in the spatial configuration of the points (*cluster structure* – clustered or dispersed) and in the number of points present. Participants judged stimuli with dispersed cluster structure to be more numerous in the magnitude comparison task. However, there was a minimal effect of cluster structure in the magnitude estimation task and no effect of the number of clusters perceived in both tasks. We also found that the clusters people perceived in the third task did not explain the effects of cluster structure. These findings go against strong claims that people use visual clustering to judge numerosity and set the stage for further investigations into the mechanisms underlying the ANS.

Keywords: approximate number system; visual clustering; ensemble perception; mathematical cognition; numerical cognition; numerosity estimation

Introduction

People are remarkably quick and accurate at judging how many objects they see. This ability is supported by the approximate number system (ANS), which is correlated with activity in the intraparietal sulcus (IPS). However, it is unclear how people go from visual input to a judgment of numerosity.

This study investigates the role of visual clustering on numerosity estimation. Visual clustering is the process of perceiving groups in visual input. Prior work has suggested that people are sensitive to the *cluster structure* of a stimulus, i.e., the spatial configuration of points (Ginsburg, 1980). People find more evenly-spaced stimuli (dispersed) more numerous than stimuli with points that are closer together (clustered). In addition, Im et al. (2016) found evidence suggesting that the number of clusters people perceive is indicative of their numerosity estimates.

In light of recent work investigating the properties of visual clustering (Marupudi & Varma, 2024), this study examines whether people's clustering of visual input is associated with their perception of approximate numerosity. We manipulated the cluster structure and the number of points present in stimuli and asked people to compare their numeric magnitudes, to estimate their numerosities, and to cluster them. We find evidence of sensitivity to cluster structure, but find that the clusters people perceive do not affect their numerosity performance. These results go against strong claims that people use visual clustering to judge numerosity.

The approximate number system

The ability to quickly judge the numerosity of a set is a useful skill that is found across many species. Humans in particular have been shown to be able to estimate the numerosity of stimuli across multiple modalities and formats such as vision, audition, digits, and words (Buckley & Gillman, 1974; Moyer & Landauer, 1967; Varma et al., 2024). The cognitive system that underlies this ability is the ANS (Brannon, 2006; Cantlon, 2015; Feigenson et al., 2004).

The ANS has been found to convert numeric information into an analog magnitude representation, similar to those for loudness or brightness (Buckley & Gillman, 1974; Fechner, 1860). It is often investigated using a magnitude comparison task, where people are shown two numerosities (e.g., point clouds) and asked which is more numerous. As with sensory-perceptual quantities, people are more sensitive to magnitude differences the greater the ratio — rather than the absolute difference — in the two numerosities. This performance profile is consistent with a logarithmically scaled mental number line where the distance between smaller numbers is larger than the distance between larger numbers. Individual differences in people's ability to discriminate between numerosities — their so-called ANS acuity — have been found to be correlated with their academic achievement (Halberda et al., 2008). There is also evidence suggesting that more symbolic and abstract mathematical abilities in humans are built upon the ANS (Butterworth et al., 2011; Szkudlarek & Brannon, 2017). However, this remains a topic of ongoing research (Cochrane et al., 2019; Szkudlarek et al., 2021). People also rely on the ANS when they view information visualizations (e.g., bar charts) to make decisions (Park, 2023). Therefore, understanding the nature of the ANS can help improve mathematics education and support decision-making more generally.

The ANS is conceptualized to be specific to number and independent of modality. However, evidence has emerged of people's sensitivity to perceptual qualities when making judgments of number. Taking vision for example, people seem to be deceived when area, a continuous perceptual dimension, leads to a different judgment than numerosity (Brannon et

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

al., 2006; Hurewitz et al., 2006; Yousif & Keil, 2020). Other visual attributes that appear to interfere with numerosity include the convex hull of the set of points and the density of the entire display (Gebuis & Reynvoet, 2012; Sanford et al., 2024). Such evidence has led to magnitude sense and perceptual interdependence accounts of numerosity perception that emphasize the role of non-numeric perceptual dimensions playing an important role in determining numerosity (Aulet & Lourenco, 2021; Leibovich et al., 2017; Lourenco & Aulet, 2023). However, it has been shown that people are sensitive to numerosities even when many of these other attributes are incongruent between stimuli (Clarke & Beck, 2021; DeWind et al., 2015; Ferrigno et al., 2017).

Why do these attributes interfere with people's numerosity estimates? One hypothesis is that they are part of how number is perceived and encoded in the brain, contrary to the classical description of the ANS. This is supported by burgeoning evidence of sensitivity to number in the early visual cortex (EVC) (DeWind et al., 2019; Fornaciai et al., 2017). The EVC is upstream of the IPS, the primary neural correlate of the ANS, suggesting a role for both domain-specific and domain-independent processing of number.

Grouping effects on numerosity estimation

To further understand the domain-specific nature of the ANS, with the ultimate goal of better understanding its underlying mechanisms, we consider another visual attribute that may affect numerosity estimation, visual grouping. This involves the perception of (hierarchical) structure in visual input. The impact of grouping on numerosity perception has been understudied compared to visual attributes such as area, convex-hull, and density.

It is important to note that the primary interest for this study are the processes underlying approximate numerosity perception, not exact enumeration (counting). People have been found to perceive the exact numerosities of up to 4 objects in constant time (Jevons, 1871; Mandler & Shebo, 1982) and to take advantage of groups of points when counting points (Ciccione & Dehaene, 2020; Starkey & McCandliss, 2014). However, it is less clear whether people take advantage of such structure in more time-constrained situations involving approximation.

A prominent example of a grouping effect in numerosity estimation is the *connectedness illusion*. Franconeri et al. (2009) and He et al. (2009) found that connecting points with lines reduces people's judgments of numerosity (as measured by a magnitude comparison task) compared to a condition without any lines and to a condition with lines that do not connect any points. He et al. (2015) proposed that this might indicate a sensitivity to the topological properties of visual input. However, other studies have since shown similar effects for many non-topological grouping factors, such as the color of the points (Adriano & Ciccione, 2024; Chakravarthi et al., 2023; Li et al., 2025; Pan et al., 2021; Yu et al., 2019).

Of particular interest to the current study is the effect of the spatial arrangement of points. In a study investigating this phenomenon, Ginsburg (1980) presented participants with either a random set of points or an evenly spaced set of points and asked them to indicate how many points they saw, i.e., perform a magnitude estimation task. Participants provided larger estimates for the evenly spaced *dispersed* stimulus compared to the more *clustered* random one. This finding was termed the *regular-random illusion*. This effect is independent of the effect of overall density (Dakin et al., 2011), since it remains constant as the positioning of points is varied within a stimulus. Other studies have since replicated this illusion using magnitude comparison tasks (Allïk & Tuulmets, 1991; Anobile et al., 2015). Importantly, Chakravarthi and Bertamini (2020) demonstrated that the cause of the illusion is not just visual crowding, leaving a role for visual clustering processes.

There is some initial evidence that implicates visual clustering in approximate numerosity perception. Im et al. (2016) asked participants to report the number of clusters in a stimulus. They used this information to develop a modified version of the K-Means clustering algorithm and to fit the model to the data. They then asked a separate set of participants to estimate the numerosity of the same set of stimuli. A model measure, number of points divided by the number of predicted clusters, was correlated with participants' numerosity estimates. However, it is not clear whether other properties, for example, the clusters participants perceived (i.e., how they separated the points), were also predictive of their numerosity estimates. It is also not clear whether the clusters perceived by people explain the effect of cluster structure on numerosity perception. For example, it is possible that the number of clusters perceived by people can predict the underestimation of clustered stimuli.

Visual clustering has been found to be a reliable cognitive ability (Marupudi & Varma, 2024). Combined with evidence that people are able to perceive the numerosity of two sets in parallel (Halberda et al., 2006), it is possible that people use clustering to determine the numerosity of a stimulus. Clustering can be used to divide a large set of points into smaller subsets, each of which can be estimated more accurately due to the logarithmically-scaled nature of the ANS. Additionally, there is evidence of grouping effects such as the connectedness illusion (Clarke et al., 2022) and the regular-random illusion (Ginsburg & Deluco, 1979) becoming stronger with age, distinct from many other attributes. This opens up the possibility that clustering and other grouping strategies help drive people's increasing ANS acuity over development.

The current study

In this study, we directly test whether people use visual clustering to enumerate points. We focus particularly on whether the statistical cluster structure of the points of a stimulus and the properties of clusters people perceive (i.e., the number of clusters they draw and the numerosities of those clusters) impact their estimates. We test subtler manipulations of statistical cluster structure (clustered vs dispersed point clouds) by using the same stimuli as Marupudi and Varma (2024). We present these stimuli to participants in the format of a magnitude comparison task and also a magnitude estimation task, in part to reconcile some of the conflicting findings and different measurement paradigms in the literature. Finally, we ask participants to cluster the same stimuli, and use the properties of the clusters they draw, such as the number of clusters, to predict their performance on the magnitude comparison and magnitude estimation tasks.

Methods

Participants

420 undergraduate students from a large university in the Midwestern US participated in this study (μ = 19.19 years, σ^2 = 1.45). 300 (71%) participants identified as women, 108 (25%) participants identified as men, 7 (1%) as non-binary or gender non-conforming, and 2 preferred not to say. 241 (57%) participants identified as White, 86 (20%) participants identified as Asian, 47 (11%) participants identified as Black, and 9 (2%) participants identified as an American Indian, Alaskan Native or Native Hawaiian. 4 participants did not provide any demographic information. Participants were compensated with course extra credit for their time. Median duration of the study was 40.3 minutes. The study was approved by the university IRB.

Design and materials

This study followed a 7 x 2 within-subjects design, with participants completing tasks on stimuli with 7 levels of Number of Points (10, 15, 20, 25, 30, 35, 40) and 2 levels of Cluster Structure (Clustered vs Dispersed). Participants completed three tasks: a magnitude comparison task, a magnitude estimation task, and a clustering task.

For all tasks, we started with the set of stimuli that were used in the Marupudi and Varma (2024) study of clustering. These stimuli differed in 7 levels of Number of Points and 2 levels of Cluster Structure. Cluster structure was operationalized as specific ranges of the clustering coefficient metric, which takes into account the spatial positions of all the points and the whitespace surrounding the convex hull of the points. Clustered and dispersed stimuli were those with clustering coefficient values of -2 ± 0.05 and 1 ± 0.05, respectively. We used the same set of stimuli to enable comparisons between the results of that study, which concerns the reliability of clustering, and the results of the current study.

Magnitude comparison trials involve determining which of two stimuli is more numerous. We constructed pairs of point cloud stimuli for participants to compare. From the base set of stimuli, we first constructed a set with all possible pairs of stimuli and then filtered pairs that would be too easy for participants: pairs where the ratio of the larger numerosity to the smaller numerosity was greater than 1.43. From this reduced set, for every possible ratio, we included 4 pairs where one stimulus was clustered and the other dispersed, and also 2 clustered-clustered stimulus pairs and 2 dispersed-dispersed stimulus pairs.

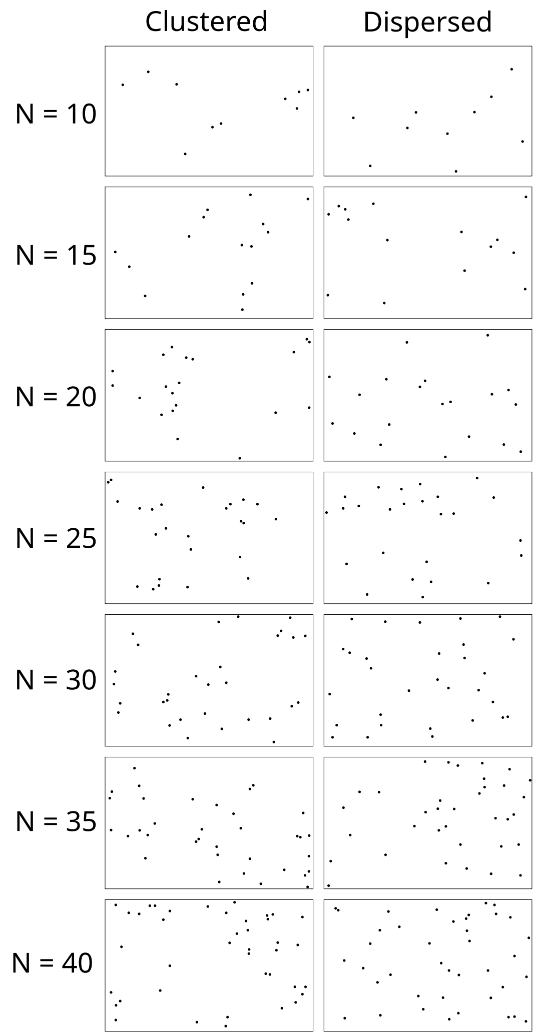

Figure 1: Examples of dispersed and clustered stimuli used for the magnitude comparison, magnitude estimation, and the clustering tasks

We added flipped versions of these pairs to prevent any potential of directional bias, resulting in participants judging each pair of stimuli twice. This process resulted in 112 magnitude comparison pairs. In addition, we included *faux comparison pairs* where both stimuli contained the same number of points. 8 faux comparison pairs (4 clustered-dispersed, 2 clustered-clustered, and 2 dispersed-dispersed) were included for each of the 30, 35, and 40 point stimuli. These pairs were included to provide additional evidence towards the perception of numerosity. If cluster structure is not an indicator of numerosity, we would expect participants to choose between the stimuli at chance. However, if people do use this indicator, then they should disproportionately choose the dispersed stimulus for the clustered-dispersed faux pairs. Adding the faux pairs resulted in a total of 136 stimulus pairs.

On magnitude estimation trials, participants viewed a single numerosity and were asked to enter the number of points they see. For these trials, we used the entire set of 56 stimuli from Marupudi and Varma (2024). On clustering trials, participants again viewed a single numerosity. This time, they were asked to draw clusters around groups of points in a stimulus. We again used the entire set of 56 from Marupudi and Varma (2024).

Procedure

After providing consent, participants first completed the magnitude comparison task. They were instructed to judge which of the two sets of points displayed to them was more numerous by pressing the [z] (left) or [m] (right) key. After a practice trial, participants completed the 136 magnitude comparison trials. Each magnitude comparison trial began with a fixation cross that was displayed for 1 second. Then, two sets of points were displayed on the screen, one to the right and the other to the left. Each set of points was outlined by a solid 1 pixel black border. The radius of each point was 2.5 pixels, displayed with antialiasing. While the points were displayed, participants responded with the [z] or [m] key to indicate their choice. Trials timed out if participants did not respond within 7 seconds.

Participants then proceeded to the magnitude estimation task. They were instructed to provide a numeric estimate of the number of points they perceived in each stimulus. Participants were unaware of the range of the numerosities being presented to them. First, they completed three practice trials with no feedback. Participants then moved on to complete the 56 experimental trials. Similar to the magnitude comparison trials, each magnitude estimation trial began with a fixation cross. Participants then saw a single set of points at the center of the screen. A text box was presented below the set for the participants to enter their estimate and press the enter key. Trials timed out if participants did not provide an estimate within 7 seconds.

Participants next completed the clustering task. The procedure was identical to Marupudi and Varma (2024). Participants viewed each stimulus and drew shapes around points to indicate the clusters they perceived. Participants were prohibited from drawing a single cluster consisting of all the points of a stimulus, but were otherwise free to cluster however they felt was most appropriate.

Finally, participants completed an arithmetic fluency task and then a series of questionnaires, the results of which are not reported in this paper.

Results

We first cleaned the data by excluding participants who provided poor quality responses. For the magnitude estimation task, we considered a trial invalid if participants provided a response outside the reasonable bounds (an estimate less than half or more than twice the correct numerosity). For the clustering task, we considered a trial invalid if the number of clusters drawn by participants was 80% of the number of points in a stimulus, a sign of misunderstanding the instructions. Data from participants who were responsible for more than 5 invalid trials were removed from the entire dataset. In addition, we removed the data from participants who performed at chance at the magnitude comparison task. 114 participants were excluded, resulting in the final dataset consisting of 306 participants.

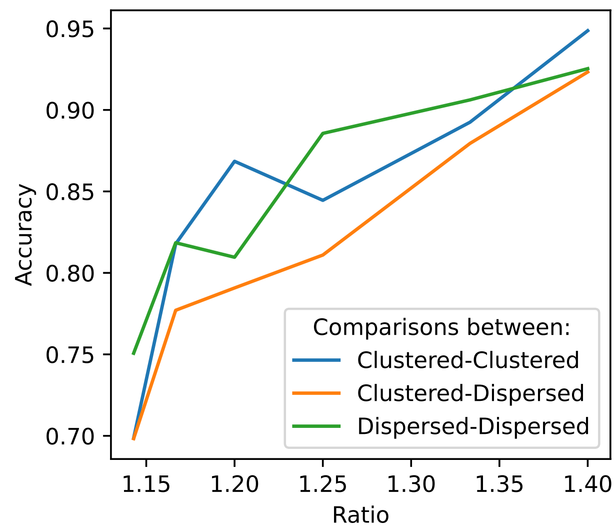

Figure 2: Participants' accuracy on the magnitude comparison task as a function of the ratio between the sets of points. Participants were less accurate on comparisons with stimuli of different cluster structures (Clustered vs Dispersed) than those with the same cluster structure (Clustered-Clustered & Dispersed-Dispersed).

Magnitude comparison task

To investigate the impact of cluster structure and the number of clusters perceived by participants on the magnitude comparison task, we determined the *cluster structure congruency* and *number of clusters congruency* for each trial. A trial was considered cluster structure congruent when the dispersed stimulus was more numerous. Likewise, a trial was considered number of clusters congruent when the stimulus for which the participants drew more clusters was more numerous.

We then fit a maximal generalized linear mixed effects model with a logistic link function predicting accuracy. We introduced an intercept term, the numerosity ratio between the pair, the number of clusters congruency, and the type of comparison (clustered-clustered, clustered-dispersed, dispersed-dispersed) as fixed effects and as random effects for each participant. The ratio term was z-score normalized. The data are depicted in Figure 2. We observed the expected main effect of the numerosity ratio between the pair of stimuli — participants were more accurate on comparison pairs with a higher ratio ($\beta = 0.60$, 95% CI [0.57, 0.63], $z = 34.63$, $p < 0.001$). Importantly, participants performed better on same cluster structure comparisons compared to trials with clustered-dispersed pairs ($\beta = 0.25$, 95% CI [0.17,

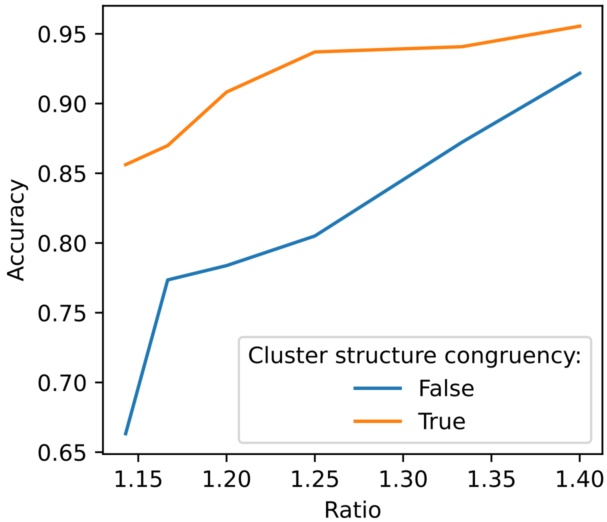

Figure 3: Accuracy on the magnitude comparison task among trials involving stimuli with different cluster structures (Clustered-Dispersed). Trials where the more numerous stimulus was more dispersed were considered congruent. Participants found congruent trials easier than incongruent trials.

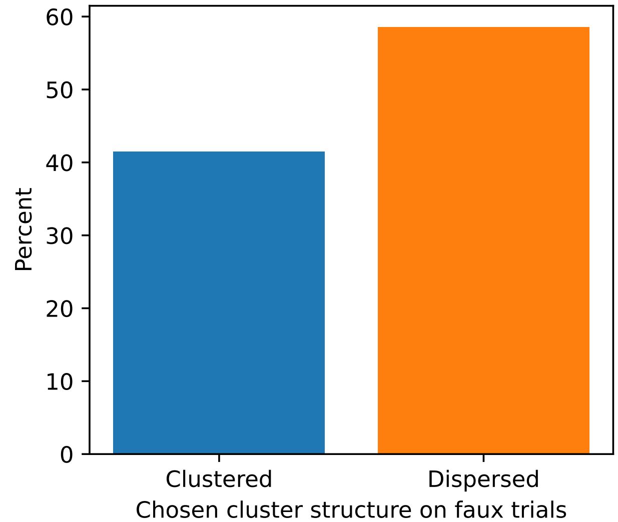

Figure 4: Percent of participants' responses on faux magnitude comparison trials where both stimuli contained the same number of points. Participants found dispersed stimuli as more numerous than clustered stimuli.

0.33], $z = 10.13$, $p < 0.001$). The odds of participants responding accurately on same cluster structure trials were 1.28 times the odds on different cluster structure trials. This suggests that cluster structure differences may be impacting people's magnitude judgments. On these trials, participants were strongly biased towards dispersed trials (cluster structure congruent trials) (Figure 3). However, we did not find an effect of the number of perceived clusters on magnitude comparison trials (number of clusters congruency; $\beta = 0.04$, 95% CI [-0.02, 0.10] $z = 1.3$, $p = 0.19$), contrary to prior work where the number of clusters was informative of participants' estimates (Im et al., 2016).

We then investigated the faux trials, for which there was no correct answer as both stimuli contained the same number of points. Specifically looking at the comparisons between clustered and dispersed stimuli, we see that participants picked the dispersed stimuli for a majority (63%) of the trials (binomial test, $p < 0.001$; Figure 4). This reinforces prior findings, both in this study and the regular-random illusion literature, that people find dispersed stimuli to be more numerous than clustered stimuli.

Magnitude estimation

Similarly to the magnitude comparison task, we investigated the role of cluster structure of the stimuli and the number of clusters perceived by participants in the magnitude estimation task. We fit a maximal linear mixed effects model predicting participants' estimates with an intercept term, the cluster structure of a stimulus, the number of clusters drawn by participants (z-score normalized), and the number of points in the stimulus (z-score normalized) as both fixed effects and random effects for each participant. Like in the magnitude comparison task, we observed the expected main effect of the number of points ($\beta = 8.45$, 95% CI [8.23, 8.67], $t = 71.616$, $p < 0.001$), where participants' estimates increase with increasing number of points in the stimulus; see Figure 5.

Additionally, we found evidence of a very small effect of cluster structure, as estimates of dispersed stimuli were 0.57 points higher than estimates of clustered stimuli ($\beta = 0.62$, 95% CI [0.21, 1.03], $t = 7.756$, $p < 0.001$, Figure 5). This strongly contrasts with the large effect of cluster structure on the magnitude comparison task. Similarly, we did not find evidence of a main effect of the number of clusters on participants magnitude estimates ($\beta = -0.26$, 95% CI [-0.53, 0.002], $t = -1.944$, $p = 0.052$). This null effect is consistent with participants' behavior on the magnitude comparison task and indicates that the number of clusters perceived may not affect participants' numerosity estimates.

Does numerosity distribution between clusters explain the effect of cluster structure?

Finally, we tested whether the clustering solutions produced by participants could explain the effect of cluster structure on numerosity estimation, i.e., why dispersed stimuli were overestimated. For example, it is possible that participants generate one large cluster with some smaller clusters for clustered stimuli and many evenly sized clusters for dispersed stimuli. Given the logarithmic scaling of approximate number perception, it is possible that the perceived value of one large set leads to greater underestimation than the sum of the perceived values of multiple smaller sets, e.g., $\log(20) <$

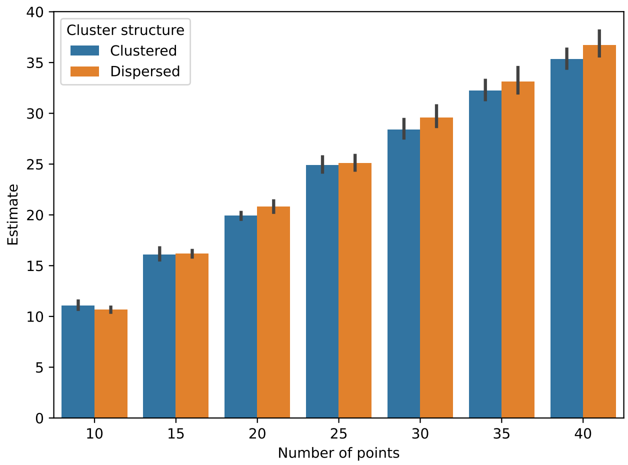

Figure 5: Estimates on the magnitude estimation task over the number of points in the stimuli. Participants estimated dispersed stimuli as slightly more numerous than clustered stimuli.

$\log(10) + \log(5) + \log(5)$.

To test this prediction, we calculated an f score for each participant and stimulus, which is the sum of the logarithm of the numerosity of each cluster they perceived in the stimulus:

$$f = \sum_c \log_b(kn_c)$$

where c is the index of a cluster, b is the base of the logarithm function, and k is a constant. We used grid search to find values of b and k that resulted in significantly different values of f for clustered and dispersed stimuli using a t-test. Despite this, we did not find any value of b and k (within reasonable ranges, $0.25 < b < 25$ and $0.25 < k < 100$) that resulted in meaningful differences in f between the cluster structure. This implies that the clustering perceived was independent of the effect of cluster structure on numerosity estimation.

Discussion

This study investigated the role human visual clustering plays in numerosity estimation. Building upon prior work on visual clustering (Marupudi & Varma, 2024) and the regular-random illusion (Ginsburg, 1980), we sought to systematically test the effects of more subtle differences in cluster structure, using stimuli for which cluster structure effects have been validated and which vary in the number of points. In addition, we directly tested the associations between the properties of people's perceived clusters, such as number of clusters drawn by participants, and numerosity performance. We did so while measuring the ANS using two different tasks: a magnitude comparison task where participants picked the more numerous stimulus, and a magnitude estimation task where they estimated of the number of points in a single stimulus.

In the magnitude comparison task, we found strong evidence of the regular-random illusion: people considered dispersed stimuli as more numerous. This was shown in two ways. First, in regular comparison trials, where the number of points differed between the stimuli, people were more accurate on trials where the dispersed stimulus was the correct answer compared to the trials where the clustered stimulus was the correct answer. Second, on trials with no correct answer, people chose the dispersed stimulus more often than the clustered stimulus. We also found that people were more likely to be accurate on trials on which they drew more clusters for the more numerous stimulus. However, the number of clusters perceived by participants was not associated with magnitude comparison performance, contrary to prior work.

We observed a different pattern in the magnitude estimation task. The effect of the cluster structure of the stimuli, despite being statistically significant, only slightly influenced participants' estimates. This contrasts with the rather large effect found by Ginsburg (1980), who also used a magnitude estimation task. In addition, the number of clusters participants perceived in a stimulus was not predictive of their estimates after controlling for the number of points in the stimulus. These results go against the hypothesis that Gestalt grouping processes may underlie the random-regular illusion (Chakravarthi et al., 2023) and suggest that they may not play a role in numerosity perception.

Looking to the future, one reason why the effect of visual clustering varied across the magnitude comparison and the magnitude estimation tasks might be due to the time spent by participants sampling the information on the screen. In both tasks, participants were provided 7 seconds to respond. However, participants completed the magnitude comparison task more quickly, with a median trial duration of 738 milliseconds compared to 2249 milliseconds in the magnitude estimation task. It is possible that cluster structure and number of cluster effects play a more important role in the early stages of visual numerosity perception, but are crowded out by other processes with more information and conscious deliberation. If this is the case, then if participants are allowed to view stimuli for the same amount of time in the two tasks, the effect of cluster structure would have been more prominent. Future research should test this prediction.

Conclusion

Our findings go against strong claims that people use visual clustering or spatial grouping processes to perceive approximate number. They also demonstrate that the clusters perceived by people do not explain the underestimation of clustered stimuli. Taken together, they shed light upon the mechanisms underlying the ANS and call for alternative explanations for the effect of cluster structure on numerosity perception.

References

Adriano, A., & Ciccione, L. (2024). The interplay between spatial and non-spatial grouping cues over approximate number perception. *Attention, Perception, & Psychophysics*. https://doi.org/10.3758/s13414-024-02908-4

Allïk, J., & Tuulmets, T. (1991). Occupancy model of perceived numerosity. *Perception & Psychophysics*, *49*(4), 303–314. https://doi.org/10.3758/BF03205986

Anobile, G., Turi, M., Cicchini, G. M., & Burr, D. C. (2015). Mechanisms for perception of numerosity or texture-density are governed by crowding-like effects. *Journal of Vision*, *15*(5), 4–4.

Aulet, L. S., & Lourenco, S. F. (2021). Numerosity and cumulative surface area are perceived holistically as integral dimensions. *Journal of Experimental Psychology: General*, *150*(1), 145–156. https://doi.org/10.1037/xge0000874

Brannon, E. M. (2006). The representation of numerical magnitude. *Current Opinion in Neurobiology*, *16*(2), 222–229. https://doi.org/10.1016/j.conb.2006.03.002

Brannon, E. M., Lutz, D., & Cordes, S. (2006). The development of area discrimination and its implications for number representation in infancy. *Developmental Science*, *9*(6), F59–F64. https://doi.org/10.1111/j.1467-7687.2006.00530.x

Buckley, P. B., & Gillman, C. B. (1974). Comparisons of digits and dot patterns. *Journal of Experimental Psychology*, *103*(6), 1131–1136. https://doi.org/10.1037/h0037361

Butterworth, B., Varma, S., & Laurillard, D. (2011). Dyscalculia: From brain to education. *science*, *332*(6033), 1049–1053. https://doi.org/10.1126/science.1201536

Cantlon, J. F. (2015). Analog origins of numerical concepts. *Mathematical Cognition and Learning*, *1*, 225–251. https://doi.org/10.1016/B978-0-12-420133-0.00009-0

Chakravarthi, R., & Bertamini, M. (2020). Clustering leads to underestimation of numerosity, but crowding is not the cause. *Cognition*, *198*, 104195. https://doi.org/10.1016/j.cognition.2020.104195

Chakravarthi, R., Nordqvist, A., Poncet, M., & Adamian, N. (2023). Fundamental units of numerosity estimation. *Cognition*, *239*, 105565. https://doi.org/10.1016/j.cognition.2023.105565

Ciccione, L., & Dehaene, S. (2020). Grouping Mechanisms in Numerosity Perception. *Open Mind*, *4*, 102–118. https://doi.org/10.1162/opmi_a_00037

Clarke, S., & Beck, J. (2021). The number sense represents (rational) numbers. *Behavioral and Brain Sciences*, *44*, e178.

Clarke, S., Qu, C., Luzzi, F., & Brannon, E. (2022). *The Connectedness Illusion Influences Numerical Perception Throughout Development*.

Cochrane, A., Cui, L., Hubbard, E. M., & Green, C. S. (2019). "Approximate number system" training: A perceptual learning approach. *Attention, Perception, & Psychophysics*, *81*(3), 621–636. https://doi.org/10.3758/s13414-018-01636-w

Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. A., & Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proceedings of the National Academy of Sciences*, *108*(49), 19552–19557. https://doi.org/10.1073/pnas.1113195108

DeWind, N. K., Adams, G. K., Platt, M. L., & Brannon, E. M. (2015). Modeling the approximate number system to quantify the contribution of visual stimulus features. *Cognition*, *142*, 247–265. https://doi.org/10.1016/j.cognition.2015.05.016

DeWind, N. K., Park, J., Woldorff, M. G., & Brannon, E. M. (2019). Numerical encoding in early visual cortex. *Cortex*, *114*, 76–89. https://doi.org/10.1016/j.cortex.2018.03.027

Fechner, G. T. (1860). Elements of psychophysics. *1*. https://doi.org/10.1037/11304-026

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002

Ferrigno, S., Jara-Ettinger, J., Piantadosi, S. T., & Cantlon, J. F. (2017). Universal and uniquely human factors in spontaneous number perception. *Nature Communications*, *8*(1), 13968. https://doi.org/10.1038/ncomms13968

Fornaciai, M., Brannon, E. M., Woldorff, M. G., & Park, J. (2017). Numerosity processing in early visual cortex. *NeuroImage*, *157*, 429–438. https://doi.org/10.1016/j.neuroimage.2017.05.069

Franconeri, S. L., Bemis, D. K., & Alvarez, G. A. (2009). Number estimation relies on a set of segmented objects. *Cognition*, *113*(1), 1–13. https://doi.org/10.1016/j.cognition.2009.07.002

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

Ginsburg, N., & Deluco, T. (1979). A Developmental Study of the Regular-Random Numerosity Illusion. *The Journal of Genetic Psychology*. Retrieved January 26, 2025, from https://www.tandfonline.com/doi/abs/10.1080/00221325.1979.10534071

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*(7213), 665–668. https://doi.org/10.1038/nature07246

Halberda, J., Sires, S. F., & Feigenson, L. (2006). Multiple Spatially Overlapping Sets Can Be Enumerated in Parallel. *Psychological Science*, *17*(7), 572–576. https://doi.org/10.1111/j.1467-9280.2006.01746.x

He, L., Zhang, J., Zhou, T., & Chen, L. (2009). Connectedness affects dot numerosity judgment: Implications for configural processing. *Psychonomic bulletin & review*, *16*(3), 509–517.

He, L., Zhou, K., Zhou, T., He, S., & Chen, L. (2015). Topology-defined units in numerosity perception. *Proceedings of the National Academy of Sciences*, *112*(41), E5647–E5655. https://doi.org/10.1073/pnas.1512408112

Hurewitz, F., Gelman, R., & Schnitzer, B. (2006). Sometimes area counts more than number. *Proceedings of the National Academy of Sciences*, *103*(51), 19599–19604.

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

Jevons, W. S. (1871). The Power of Numerical Discrimination. *Nature*, *3*(67), 281–282. https://doi.org/10.1038/003281a0

Leibovich, T., Katzin, N., Harel, M., & Henik, A. (2017). From "sense of number" to "sense of magnitude": The role of continuous magnitudes in numerical cognition. *Behavioral and Brain Sciences*, *40*, e164. https://doi.org/10.1017/S0140525X16000960

Li, Q., Ting, G., Kikuno, Y., & Kazuhiko, Y. (2025). The influence of increasing color variety on numerosity estimation and counting. *Psychonomic Bulletin & Review*, 1–11.

Lourenco, S. F., & Aulet, L. S. (2023). A theory of perceptual number encoding. *Psychological Review*, *130*(1), 155–182. https://doi.org/10.1037/rev0000380

Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, *111*, 1–22. https://doi.org/10.1037/0096-3445.111.1.1

Marupudi, V., & Varma, S. (2024). Human visual clustering of point arrays. *Psychological Review*. https://doi.org/10.1037/rev0000525

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, *215*(5109), 1519–1520.

Pan, Y., Yang, H., Li, M., Zhang, J., & Cui, L. (2021). Grouping strategies in numerosity perception between intrinsic and extrinsic grouping cues. *Scientific Reports*, *11*(1), 17605. https://doi.org/10.1038/s41598-021-96944-x

Park, J. (2023). *The role of numerical cognitive processes for bar graph comprehension* [Doctoral dissertation, University of Minnesota].

Sanford, E. M., Topaz, C. M., & Halberda, J. (2024). Modeling Magnitude Discrimination: Effects of Internal Precision and Attentional Weighting of Feature Dimensions. *Cognitive Science*, *48*(2), e13409. https://doi.org/10.1111/cogs.13409

Starkey, G. S., & McCandliss, B. D. (2014). The emergence of "groupitizing" in children's numerical cognition. *Journal of experimental child psychology*, *126*, 120–137.

Szkudlarek, E., & Brannon, E. M. (2017). Does the Approximate Number System Serve as a Foundation for Symbolic Mathematics? *Language Learning and Development*, *13*(2), 171–190. https://doi.org/10.1080/15475441.2016.1263573

Szkudlarek, E., Park, J., & Brannon, E. M. (2021). Failure to replicate the benefit of approximate arithmetic training for symbolic arithmetic fluency in adults. *Cognition*, *207*, 104521. https://doi.org/10.1016/j.cognition.2020.104521

Varma, S., Sanford, E. M., Marupudi, V., Shaffer, O., & Lea, R. B. (2024). Recruitment of magnitude representations to understand graded words. *Cognitive Psychology*, *153*, 101673.

Yousif, S. R., & Keil, F. C. (2020). Area, not number, dominates estimates of visual quantities. *Scientific Reports*, *10*(1), 13407. https://doi.org/10.1038/s41598-020-68593-z

Yu, D., Xiao, X., Bemis, D. K., & Franconeri, S. L. (2019). Similarity Grouping as Feature-Based Selection. *Psychological Science*, *30*(3), 376–385. https://doi.org/10.1177/0956797618822798