PNAS
www.pnas.org

Supplementary Information for

A primarily serial, foveal accumulator underlies approximate numerical estimation

Samuel J. Cheyette & Steven T. Piantadosi

Samuel J. Cheyette
E-mail: sjcheyette@gmail.com

This PDF file includes:

Supplementary text
Fig. S1
Tables S1 to S6
References for SI reference citations

www.pnas.org/cgi/doi/10.1073/pnas.1819956116

Supporting Information Text

Experiment 1.

Effect of time on mean and variance. We ran a Bayesian hierarchical regression to estimate participant-level and group-level slopes and Weber fractions, as well as the effects of time on both. We assumed that each individual's mean estimate and standard deviation about that estimate varied linearly as a function of the quantity displayed and logarithmically as a function of time. We will call participants' baseline (independent of time) slopes and Weber fractions β_0 and w_0; we will denote time t; and we will call the effect of time on slopes β_t and Weber fractions w_t. In order to keep slopes and Weber fractions positive, we assume an exponential linking function between slope and the predictors. Specifically, Equations 1 and 2 show how the slope and Weber fractions for each participant are calculated:

$$\beta = e^{\beta_0 + \beta_t \cdot log(t)} \quad [1]$$

$$w = e^{w_0 + w_t \cdot log(t)} \quad [2]$$

Then, each participant's mean estimate is drawn from a Gaussian centered around $\beta \cdot N$ with standard deviation $w \cdot \beta \cdot N$.

Trials where the eye-tracker did not capture eye-movements (determined by Tobii software) were removed (17 trials); as well as estimates that were more or less than 3 standard deviations away from the mean (7 trials). Figure 2 shows the mean slope (2c) and Weber fraction (2d) in each time condition; the group-level means are shown in red and each participant's mean in black. If participants' estimates were un-biased on average then the group mean slopes would be 1; and if time did not have an effect, the group mean slopes and Weber fractions would remain constant across time. It is clear visually that participants' estimates are biased and that time had an effect on both the slope and Weber fraction. Table S1 provides the inferred group-level regression weights and the uncertainty of the estimate. The fact that the intercept is negative ($\beta_0 = -0.24; CI = [-0.28, -0.19]$) indicates a baseline tendency to underestimate. Most significantly, the effect of time on both mean slopes and Weber fractions is significantly different than 0: time increases the group mean slope ($\beta_t = 0.05$; $CI = [0.03, 0.09]$)*; and it decreases the group mean Weber fraction ($w_t = -0.11$; $CI = [-0.12, -0.09]$). Participants' average slope increases by about 17% (0.71 to 0.83) from the shortest to the longest time condition; and their average Weber fraction decreases by about 21% (0.28 to 0.22).

Var	Value	2.5%	97.5%
β_0	-0.24	-0.28	-0.19
β_t	0.05	0.03	0.09
w_0	-1.42	-1.59	-1.21
w_t	-0.11	-0.12	-0.09

Table S1. Group-level regression weights and their 95% credible intervals for each condition. For the mean slope, the inferred weights include the intercept (β_0), the effect of time (β_t). For Weber fractions, the inferred values are analogous, with the intercept (w_0), the effect of time (w_t).

Figure 4a shows the percent of dots within a participants' gaze for each time condition. As is visually apparent, there is a monotonic relationship such that an increase in time increases the number of dots foveated. The average proportion of dots within the foveal gaze for each time condition from shortest to longest are 0.18 ($SD = 0.03$), 0.22 ($SD = 0.04$), 0.39 ($SD = 0.08$), and 0.64 ($SD = 0.14$). Our prediction is that the increase in the proportion of dots foveated between time conditions may account for the observed effects of time on both accuracy and bias towards underestimation.

Visual fixations mediate effects of time. To quantify whether the effect of time was explained by eye movement patterns, we ran another Bayesian regression that was identical to the one reported above, with the addition of random variables for the effect of the *proportion of dots foveated* on the mean and variance of each participant's estimate. That is, we used the same terms as in equations 1 and 2, but added terms β_s and w_s to the way slope and Weber fraction are computed — each term is multiplied by s, representing the proportion of dots foveated. Equations 3 and 4 show the calculation of β and w in full.

$$\beta = e^{\beta_0 + \beta_t \cdot log(t) + \beta_s \cdot s} \quad [3]$$

$$w = e^{w_0 + w_t \cdot log(t) + w_s \cdot s} \quad [4]$$

Table S2 shows the results of this analysis in full. Two findings are worth highlighting. First, the proportion of dots foveated significantly affects the mean and variance of participants' estimates. Second, time no longer has a significant effect on either. The effect of foveation on the mean can be seen in Figure 4c: as the proportion of dots foveated increases, so do participants' mean estimates.† Congruently, the regression shows that the group-level mean β_s is significantly above 0 ($\beta_s = 0.43; CI = [0.26, 0.59]$). There is also an effect of foveation on the variance of estimates. Figure 4d shows that as foveation increases, Weber fractions tend to decrease. The group-level regression revealed that the effect of foveation is significant ($w_s = -0.67, CI = [-0.95, -0.27]$). Finally, consistent with our hypothesis, the effects of time were no longer significantly different than 0 when accounting for the visual samples. The lack of a time effect on the mean (when conditioning on percent of dots foveated) can be seen clearly in Figure 4b which shows the average deviation as a function of dots foveated and colored by time: if there were a significant effect of time over-and-above the differences driven by fixations, the regression lines for each time condition would be non-overlapping.

*CI here indicates the *credible interval*, not the *confidence interval*.

†Note that the lines are non-linear in Figures 4c and 4d because the y-axis measures are collapsed over other predictors.

Var	Value	2.5%	97.5%
β_0	-0.74	-0.87	-0.65
β_t	-0.00	-0.02	0.03
β_s	0.43	0.26	0.59
w_0	-0.98	-0.74	-1.21
w_t	0.05	-0.03	0.12
w_s	-0.67	-0.95	-0.27

Table S2. Group-level regression weights and their 95% confidence intervals for each condition. For the mean slope, the inferred weights include the intercept (β_0), the effect of time (β_t), and the effect of the percent of dots seen (β_s). For Weber fractions, the inferred values are analogous, with the intercept (w_0), the effect of time (w_t), and the effect of the percent of dots seen (w_s).

Analysis within time conditions. To verify the robustness of the the finding that mean estimates increase with greater foveation, we ran a secondary analysis to test the effect of foveation *within each time condition*. Specifically, we ran four (Bayesian) linear regressions with group-level and subject-level parameters representing the effect of foveation on slopes (β_s), in addition to independent slope (β_0) and Weber (w_0) parameters. We used the un-standardized proportion of dots foveated to allow us to compare the effect of foveation across the time conditions. For a given number of dots displayed, N, the model assumed that estimates were drawn from a Gaussian centered around $(\beta_0 + \beta_s) \cdot N$ with standard deviation $w_0 \cdot \beta \cdot N$. The results of these regressions are shown in Table S3. Corroborating the findings from previous analyses, the proportion of dots foveated has a significant positive effect on mean estimates within each time condition (though it is marginal at 0.1s). The inferred slopes do not significantly differ from one another.

Time (s)	β_s	2.5%	97.5%
0.1	0.27	-0.01	0.60
0.3	0.31	0.05	0.44
1	0.40	0.21	0.60
3	0.26	0.13	0.43

Table S3. Group-level regression weights and their 95% confidence intervals for the effect of foveation on slopes in each time condition.

Experiment 2. To analyze the discrimination experiment, we ran a regression analogous to the one used for the estimation task, but predicting the probability of guessing the second array given the participant's mean slope (β_0), mean Weber fraction (w_0), and the effect of the duration of the stimuli on the mean (β_t) and the Weber fraction (w_t). If n_1 is the number of dots in the first display and n_2 is the number of dots in the second display, then the ratio is defined as $\frac{(n_1-n_2)}{\sqrt{(n_1^2+n_2^2)}}$. If t_1 and t_2 are the display times of the first and second stimulus, then the probability a participant chooses the second screen is given below in Equation 5.

$$p(choose\ 2) = \frac{1}{2} + \frac{1}{2} erf\left(\frac{n_2 \cdot e^{\beta_0+\beta_t \cdot log(t_2)} - n_1 \cdot e^{\beta_0+\beta_t \cdot log(t_1)}}{\sqrt{2} e^{w_0+w_t \cdot log(t_1)+w_t \cdot log(t_2)} \sqrt{(n_1^2+n_2^2)}}\right) \quad [5]$$

The results of the regression are shown in Table S4. Corroborating the estimation task, the effects of time on both the group-level mean estimate (β_t) and Weber fraction (w_t) are significantly different than 0. Moreover, they were in the same direction as in the estimation task: mean estimates increase ($\beta_t = 0.05$; $CI = [0.04, 0.07]$) and Weber fractions decrease ($w_t = -0.07$; $CI = [-0.12, -0.02]$).

Var	Value	2.5%	97.5%
β_0	-1.22	-1.53	-0.99
β_t	0.05	0.04	0.07
w_0	-1.73	-1.14	-2.31
w_t	-0.07	-0.12	-0.02

Table S4. Group-level regression weights and their 95% credible intervals for each condition in the discrimination task using only time as a predictor for the mean and variance. For the mean slope, the inferred weights include the intercept (β_0), the effect of time (β_t). For Weber fractions, the inferred values are analogous, with the intercept (w_0), the effect of time (w_t).

Recency bias cannot explain effects of time. An alternative account of our finding that people respond that the second stimulus is more numerous than the first in the Short-Long condition than the Long-Short condition is that it is a mere effect of *recency* (1). We can test this possibility in our design by comparing responses across the conditions where the second stimulus' duration is held constant. That is, comparing responses in the Short-Long to the Long-Long condition; and comparing responses in the Long-Short condition to the Short-Short condition. In both cases, recency could not explain differences in perceived numerosity.

If longer duration does, in fact, increase perceived numerosity, then participants should rate the first stimulus in the Short-Long as less numerous than the second relative to the first stimulus in the Long-Long condition.

We ran two logistic regressions to determine the effect of increasing the first stimulus' duration. The first regression was run on only the conditions where the second stimulus was 0.1s (short) and the second regression was run on only the conditions where the second stimulus was 1s (long). The ratio of the numerosities presented was also entered as a predictor. Responding that the first stimulus was greater numerosity than the second was coded as *0* and responding that the second was greater than the first was coded as *1*.

The results of this analysis revealed effects of stimulus duration in the predicted direction. The first regression — looking at the conditions when the second stimulus was short — showed a significant effect of the first stimulus' duration ($\beta = -0.32;\ z = -4.04;\ p < 0.001$), such that increasing the presentation duration of the first stimulus increased participants' likelihood of reporting that the first stimulus was of greater numerosity. Likewise, the same effect was revealed in the second regression ($\beta = -0.19;\ z = -2.22\ p = 0.03$), looking at the conditions when the second stimulus was long. Taken together, these findings effectively rule out the possibility that a *recency bias* explains the patterns in our data; instead, longer duration of presentation itself seems to increase perceived numerosity.

Visual fixations mediate effects of time. To determine if this observed effect of time is mediated by visual fixations, we ran a regression identical to the one above with the effect of the percent of dots foveated added as a predictor of both the mean estimate and Weber fraction. The results of the regression are shown in Table S5. Most importantly, the effect of the percent of dots foveated on the mean is positive ($\beta_s = 2.93; CI = [0.24, 4.01]$); and the effect on the Weber fraction is negative ($w_s = -0.14; CI = [-0.21, -0.05]$). The effect of time on the slope is negligible ($\beta_t = -0.01; CI = [-0.07, 0.03]$); but there is still an effect of time on the Weber fraction ($w_t = -0.07; CI = [-0.14, -0.02]$), indicating that the percent of dots foveated probably does not entirely mediate the effect of time on accuracy.

Var	Value	2.5%	97.5%
β_0	-0.29	-0.94	0.41
β_t	-0.01	-0.07	0.03
β_s	2.93	0.24	4.01
w_0	-0.96	-0.61	-1.36
w_t	-0.07	-0.14	-0.02
w_s	-0.14	-0.21	-0.05

Table S5. Group-level regression weights and their 95% credible intervals for each condition in the discrimination task, including parameters for time an the percent of dots foveated as parameters affecting mean slopes and Weber fractions. For the mean slope, the inferred weights include the intercept (β_0), the effect of time (β_t), and the effect of the percent of dots seen (β_s). For Weber fractions, the inferred values are analogous, with the intercept (w_0), the effect of time (w_t), and the effect of the percent of dots seen (w_s).

Model.

Model implementation. The full model is given below. The subscripts for each variable denote whether that variable applies to *foveal* (F) or *peripheral* (P) dots; and whether that variable is at the *group*-level (G) or at the *subject*-level (S). All variables that have only one subscript apply to both foveal and peripheral dots, so the subscript denotes only whether it is a group- or subject-level variable. The mean of a participant's estimate, μ, is a function of five quantities: the number of foveated (N_F) and peripheral (N_P) dots; the proportion of screen area foveated (A_F) and peripheral (A_P); and the total number of times all the dots were re-fixated (N_D). Each of these parameters has a corresponding inferred weight in the regression.

$$\beta_{F,G},\ \beta_{P,G},\ \beta_{N,G},\ \beta_{D,G} \sim \mathcal{N}(0, 100)$$
$$\sigma_{F,G},\ \sigma_{P,G},\ \sigma_{N,G}, \sigma_{D,G} \sim |\mathcal{N}(0, 100)|$$
$$\gamma_{F,G},\ \gamma_{P,G} \sim Beta(1, 1)$$
$$\lambda_G \sim Exp(1)$$

$$\beta_{F,S} \sim \mathcal{N}(\beta_{F,G}, \sigma^2_{F,G})$$
$$\beta_{P,S} \sim \mathcal{N}(\beta_{P,G}, \sigma^2_{P,G})$$
$$\beta_{N,S} \sim \mathcal{N}(\beta_{N,G}, \sigma^2_{N,G})$$
$$\beta_{D,S} \sim \mathcal{N}(\beta_{D,G}, \sigma^2_{P,G})$$

$$\gamma_{F,S} \sim Beta(\lambda\gamma_{F,G},\ \lambda(1 - \gamma_{F,G}))$$
$$\gamma_{P,S} \sim Beta(\lambda\gamma_{P,G},\ \lambda(1 - \gamma_{P,G}))$$

$$\mu_S = (1 + \beta_{D,S} \cdot N_D)\ (\beta_{F,S} \cdot N_F \cdot (1/A_F)^{\gamma_{F,S}}) \quad + \quad \beta_{P,S} \cdot N_F \cdot (1/A_P)^{\gamma_{P,S}}$$
$$\sigma_S = \beta_{N,S} \cdot \mu_S$$
$$E \sim \mathcal{N}(\mu_S, \sigma^2_S)$$

Model inference. The values for parameters in the model were inferred using the gradient-based MCMC sampler NUTS (2) which was implemented in the Python package PyMC3 (as were all the hierarchical models) (3). Two chains were ran for 10,000 steps each.

Model results. The important group-level parameters are given in Table S6. Importantly, foveated dots have a significantly greater effect on the mean estimate ($\beta_{foveal} = 0.88$; $CI = [0.79, 0.96]$) than those on the periphery ($\beta_{peripheral} = 0.47$; $CI = [0.36, 0.55]$). Figure 6b shows the regression weights of β_{foveal} on and $\beta_{peripheral}$. Note also that the difference in β values here is actually underestimate of the difference in the relative contributions to the mean, since it does not take into account the amount of re-scaling by area, which happens to be slightly greater for peripheral regions. However, re-scaling is inferred to have only a miniscule influence for both the area foveated ($\gamma_{foveal} = 0.11; CI = [0.04, 0.18]$) and not foveated ($\gamma_{peripheral} = 0.11; CI = [0.01, 0.23]$). Figure 6c shows the values of γ_{foveal} and $\gamma_{peripheral}$. That no subject has a value above 0.5 for either γ_{foveal} or $\gamma_{peripheral}$ is strong evidence that re-normalization by area is not an important factor in estimating numerosity. Double-counting seems to also have minimal effect on the mean estimate ($\beta_{double} = 0.01; CI = [0.00, 0.05]$). The average contribution of each one of these factors at each time condition is shown in Figure 6d.

Var	Value	2.5%	97.5%
β_{foveal}	0.88	0.79	0.96
$\beta_{peripheral}$	0.47	0.36	0.55
γ_{foveal}	0.11	0.04	0.18
$\gamma_{peripheral}$	0.11	0.01	0.23
β_{double}	0.01	0.00	0.05

Table S6. Group-level regression weights and their 95% credible intervals for the effect of dots foveated (β_{foveal}) and not foveated ($\beta_{peripheral}$) on the mean estimate; for the effect of re-foveating on the mean (β_{double}) and for the foveal re-scaling factor (γ_{foveal}) and the peripheral re-scaling factor ($\gamma_{peripheral}$).

Foveation analysis. It will be important for future research to better determine how saccades are programmed, since this may explain some influence of the properties of visual displays on numerical estimation (4, 5), including biases introduced by object clustering (6). Research has shown that a pre-saccadic selection process takes place over competing regions in the visual periphery (7). Indeed, in our tasks, there is likely a non-random nature to participants' saccades — this is suggested by the fixation paths in Figure 3. After starting at center fixation, participants tended to fixate regions of the screen that had higher density; and their gaze remained in higher density regions for longer. We computed the mean x- and y- coordinates of the displayed dots for each trial in the 3-second time condition (where participants could saccade freely). We found that the mean x-coordinate of the dots significantly correlated with the participant's mean x-coordinate gaze ($r = 0.27, p < 0.001$); this is likewise true for the y-coordinates ($r = 0.34, p < 0.001$). This is consistent with previous results showing that people tend to look towards the greater of two quantities first and for longer in a quantity discrimination task (8).

Continuous variables.

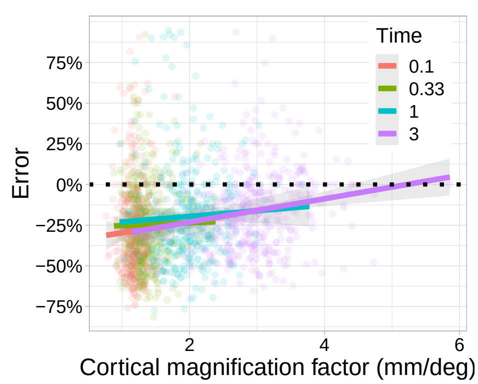

Fig. S1. The percent deviation of estimates from the true number of dots (y-axis) as a function of the percent of the average cortical magnification (x-axis) of each dot displayed in a trial. Each time condition is grouped by color.

Analysis using cortical magnification factor. In our primary analysis, we considered a dot "foveated" if it fell within a 5° window around someone's gaze for more than 50ms. While this measure has the benefit of simplicity, the exact values are somewhat arbitrary. A somewhat more complicated, but less ad hoc approach, would be to use the *cortical magnification factor*, which is known to predict visual acuity (e.g. 9). The cortical magnification factor (CMF) is inversely proportional to the eccentricity of an object from someone's gaze. For each trial, we calculated the CMF for each dot based on the minimum distance between each dot and the participants' gaze (in terms of visual degrees). We took the mean CMF over all dots and used that as a predictor of the mean and variance of estimates. The results revealed no substantial differences between this metric and the one used in the main text. Figure S1 shows, for example, the relationship between the CMF and participants' errors. There is clearly no substantial difference in the qualitative result, as can be seen by comparing this to Figure 4b.

Analysis using convex hull. There is an ongoing debate about the importance of continuous variables such as area, density, and convex hull on number estimation (e.g. 10–13). Our experimental design was not suited to testing whether people were using these types of heuristics (nor was it intended to). In particular, total area was perfectly correlated with number since the size of the dots was constant across trials. Convex hull was also strongly correlated with the total number of dots displayed, however there was enough random variation to allow us to test its influence on numerical estimation. Given the dependence of estimates on eye movements, one might expect greater underestimation from displays with greater convex hull. To evaluate this, we ran a regression to predict percent estimation error (signed, so not absolute error) from the number of dots shown, the proportion of dots foveated, and the convex hull of the dots. There was a significant effect of the number of dots on estimation error, such that participants' bias to under-estimate increased with the number of dots ($\beta = -0.24; t = -6.43; p < 0.001$), even controlling for convex hull. The effect of convex hull trended in the same direction, but 1/4th the size and was only marginally significant ($\beta = -0.06; t = -1.74; p = 0.08$). Consistent with previous analyses, foveating a greater proportion of the dots has the opposite effect, pushing estimates higher ($\beta = 0.14; t = 6.14; p < 0.001$). These results therefore indicate that there are strong effects of number over and above convex hull, and weak-to-nonexistent effects of convex hull controlling for number and foveation. Interestingly, the influence of convex hull (to the extent it is present) is in the opposite direction as has been found previously (12).

References

1. Van den Berg R, Lindskog M, Poom L, Winman A (2017) Recent is more: A negative time-order effect in nonsymbolic numerical judgment. *Journal of Experimental Psychology: Human Perception and Performance* 43(6):1084.
2. Hoffman MD, Gelman A (2014) The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research* 15(1):1593–1623.
3. Salvatier J, Wiecki TV, Fonnesbeck C (2016) Probabilistic programming in python using pymc3. *PeerJ Computer Science* 2:e55.
4. Burr D, Ross J (2008) A visual sense of number. *Current biology* 18(6):425–428.
5. Dakin SC, Tibber MS, Greenwood JA, Morgan MJ, , et al. (2011) A common visual metric for approximate number and density. *Proceedings of the National Academy of Sciences* 108(49):19552–19557.
6. Im HY, Zhong Sh, Halberda J (2016) Grouping by proximity and the visual impression of approximate number in random dot arrays. *Vision research* 126:291–307.
7. Fischer B, Weber H (1993) Express saccades and visual attention. *Behavioral and Brain Sciences* 16(3):553–567.
8. Odic D, Halberda J (2015) Eye movements reveal distinct encoding patterns for number and cumulative surface area in random dot arrays. *Journal of vision* 15(15):5–5.
9. Cowey A, Rolls E (1974) Human cortical magnification factor and its relation to visual acuity. *Experimental Brain Research* 21(5):447–454.
10. Gebuis T, Reynvoet B (2012) The interplay between nonsymbolic number and its continuous visual properties. *Journal of Experimental Psychology: General* 141(4):642.
11. Starr A, DeWind NK, Brannon EM (2017) The contributions of numerical acuity and non-numerical stimulus features to the development of the number sense and symbolic math achievement. *Cognition* 168:222 – 233.
12. Gebuis T, Reynvoet B (2012) The role of visual information in numerosity estimation. *PloS one* 7(5):e37426.
13. Ferrigno S, Jara-Ettinger J, Piantadosi ST, Cantlon JF (2017) Universal and uniquely human factors in spontaneous number perception. *Nature communications* 8:13968.