Endogenous Precision of the Number Sense

Arthur Prat-Carrabin[1,†,⋆] and Michael Woodford[1]

[1]Department of Economics, Columbia University, New York, USA
[†]*Present address:* Department of Psychology, Harvard University, Cambridge, USA
[⋆]*email: arthurpc@fas.harvard.edu*

April 10, 2024

Abstract

The behavioral variability in psychophysical experiments and the stochasticity of sensory neurons have revealed the inherent imprecision in the brain's representations of environmental variables. Numerosity studies yield similar results, pointing to an imprecise 'number sense' in the brain. If the imprecision in representations reflects an optimal allocation of limited cognitive resources, as suggested by efficient-coding models, then it should depend on the context in which representations are elicited. Through an estimation task and a discrimination task, both involving numerosities, we show that the scale of subjects' imprecision increases, but sublinearly, with the width of the prior distribution from which numbers are sampled. This sublinear relation is notably different in the two tasks. The double dependence of the imprecision — both on the prior and on the task — is consistent with the optimization of a tradeoff between the expected reward, different for each task, and a resource cost of the encoding neurons' activity. Comparing the two tasks allows us to clarify the form of the resource constraint. Our results suggest that perceptual noise is endogenously determined, and that the precision of percepts varies both with the context in which they are elicited, and with the observer's objective.

Significance statement

Results in neuroscience and psychology have suggested that the precision with which we represent the important variables of our environment, including numbers, proceeds from an optimized tradeoff between the objective and the cost of our representations. But the nature of this objective and of this cost remain unclear. By comparing the behavioral variability obtained in two experiments, and using several different ranges of numbers, we show that human observers optimize the objective of their current task (instead of a general-purpose objective, as often assumed), under a resource cost of the encoding neurons. This results in sublinear scaling laws, obtained in data, relating the degree of imprecision of internal representations to the range of stimuli expected in a given context.

Quartz wristwatches gain or lose about half a second every day. Still, they are useful for what one typically needs to know about the time, and they sell for as low as five dollars. The most recent atomic clocks carry an error of less than one second over the age of the Universe, and they are used to detect the effect of Einstein's theory of general relativity at a millimeter scale[1]; but they are much more expensive. Precision comes at a cost, and the kind of cost that one is willing to bear depends on one's objective. Here we argue that in order to make the many decisions that stipple our daily lives, the brain faces—and rationally solves—similar tradeoff problems, which we describe formally, between an objective that may vary with the context, and a cost on the precision of its internal representations about external information.

As a considerable fraction of our decisions hinges on our appreciation of environmental variables, it is a matter of central interest to understand the brain's internal representations of these variables—and the factors that determine their precision. An almost invariable behavioral pattern, in more than a century of studies in psychophysics, is that the responses of subjects exhibit variability across repeated trials. This variability has increasingly been thought to reflect the randomness in the brain's representations of the magnitudes of the experimental stimuli[2–4]. Substantiating this view, studies in neuroscience exhibit how many of these representations seem to materialize in the activity of populations of neurons, whose patterns of firing of action potentials (electric signals) are well described by Poisson processes: typically, average firing rates are functions ('tuning curves') of the stimulus magnitude, which is therefore 'encoded' in an ensemble of action potentials, i.e., in a stochastic, and thus imprecise, fashion[5–7]. Similar results have been obtained in studies on the perception of numerical magnitudes. People are imprecise, when asked to estimate the 'numerosity' of an array of items, or in tasks involving Arabic numerals[8,9]; and the tuning curves of number-selective neurons in the brains of humans and monkeys have been exhibited[10,11]. These findings point to the existence of a 'number sense' that endows humans (and some animals) with the ability to represent, imprecisely, numerical magnitudes[12].

The quality of neural representations depends on the number of neurons dedicated to the encoding, on the specifics of their tuning curves, and on the duration for which they are probed. Models of *efficient coding* propose, as a guiding principle, that the encoding optimizes some measure of the fidelity of the representation, under a constraint on the available encoding resources[13–25]. While they make several successful predictions (e.g., more frequent stimuli are encoded with higher precision[16,19,20,25,26]), including in the numerosity domain[27,28], several aspects of these models remain subject to debate[29,30], although they shape crucial features of the predicted representations. First, in many studies, the encoding is assumed to optimize the mutual information between the external stimulus and the internal representations[18–20,22], but it is seldom the case that this is actually the objective that an observer needs to optimize. An alternative possibility is that the encoding optimizes the observer's current objective, which may vary depending on the task at hand[24,31]. Second, the nature of the resource that constrains the encoding is also unclear, and several possible limiting quantities are suggested in the literature (e.g., the expected spike rate, the number of neurons[16,17,20], or a functional on the Fisher information, a statistical measure of the encoding precision[18,19,21,23,24]). Third, most studies posit that the resource in question is costless, up to a certain bound beyond which the resource becomes depleted. Another possibility is that there is a cost that increases with increasing utilization of the resource (e.g., action potentials come with a metabolic cost[32–34]). Together, these aspects determine how the optimal encoding, and thus the resulting behavior, depend on the task and on the 'prior' (the stimulus distribution).

Hence we shed light on all three questions by manipulating, in experiments, the task and the prior. In an estimation task, subjects estimate the numbers of dots in briefly presented arrays. In a discrimination task, subjects see two series of numbers and are asked to choose the one with the highest average. In both tasks, experimental conditions differ by the size of the range of numbers that are presented to subjects (i.e., by the width of the prior). In each case we examine closely the variability of the subjects' responses. We find that it depends on both the task and the prior. The scale of the subjects' imprecision increases *sublinearly* with the width of the prior, and this sublinear relation is different in the two tasks. We reject 'normalization' accounts of the behavioral variability, and in the estimation task we find no evidence of 'scalar variability', whereby the standard deviation of estimates for a number is proportional to the number, as sometimes reported in numerosity studies. The behavioral patterns we exhibit are predicted by a model in which the imprecision in representations is adapted to the observer's current task, whose expected reward it optimizes under a resource cost on the activity of the encoding neurons. The subjects' imprecision is thus endogenously determined, through the rational allocation of costly encoding resources.

Our experimental results suggest, at least in the numerosity domain, a behavioral regularity — a task-dependent quantitative law of the scaling of the responses' variability with the range of the prior — for which we provide a resource-rational account. Below, we present the results pertaining to the estimation task, followed by those of the discrimination task, before turning to our theoretical account of these experimental findings. The results we present here are obtained by pooling together the responses of the subjects; the analysis of individual data further substantiates our conclusions (see Methods).

Estimation task

In each trial of a numerosity estimation task, subjects are asked to provide their best estimate of the number of dots contained in an array of dots presented for 500ms on a computer screen (Fig. 1a). In all trials, the number of dots is randomly sampled from a uniform distribution, hereafter called 'the prior', but the width of the prior, w, is different in three experimental conditions. In the 'Narrow' condition, the range of the prior is $[50, 70]$ (thus the width w is 20); in the 'Medium' condition, the range is $[40, 80]$ (thus $w = 40$); and in the 'Wide' condition, the range is $[30, 90]$ (thus $w = 60$; Fig. 1b). In all three conditions the mean of the prior (which is the middle of the range) is 60. As an incentive, the subjects receive for each trial a financial reward which decreases linearly with the square of their estimation error. Each condition comprises 120 trials, and thus often the same number is presented multiple times, but in these cases the subjects do not always provide the same estimates. We now examine this variability in subjects' responses.

Studies on numerosity estimation with similar stimuli sometimes report that the standard deviation of estimates increases proportionally to the estimated number. This property, dubbed 'scalar variability', has been seen as a signature of numerical-estimation tasks, and more generally, of the 'number sense'[35]. However, looking at the standard deviation of estimates as a function of the presented number, we find that it is not well described by an increasing line. In the three conditions, the standard deviation seems to be maximal near the center of the range (60), and to slightly decrease for numbers closer to the boundaries of the prior (Fig. 1c). Dividing each prior range in five bins of similar sizes, we compute the variance of estimates in each bin (see Methods). In the three conditions, the variance in the middle (third) bin is greater than the variances in the fourth and fifth bins (which contain larger numbers). These differences are significative (p-values of Levene's tests of equality of variances: third vs. fifth bin, largest p-v. across the three conditions: 5e-6; third vs. fourth bin, Narrow condition: 0.009, Medium condition: 1.2e-5) except between the third and fourth bin in the Wide condition (p-v.: 0.12). This substantiates the conclusion that the standard deviation of estimates is not an increasing linear function of the number. Moreover, a hallmark of scalar variability is that the 'coefficient of variation', defined as the ratio of the standard deviation of estimates to the mean estimate, is constant[35]. We find that in our experiment, it is decreasing for most of the numbers, in the three conditions (Fig. 1e); this is consistent with the results of Ref.[36]. We conclude that the scalar-variability property is not verified in our data.

In fact, the most striking feature of the variability of estimates is not how it depends on the number, but how it strongly depends on the width of the prior, w (Fig. 1c,d). For instance, with the numerosity 60, the standard deviation of subjects' estimates is 4.2 in the

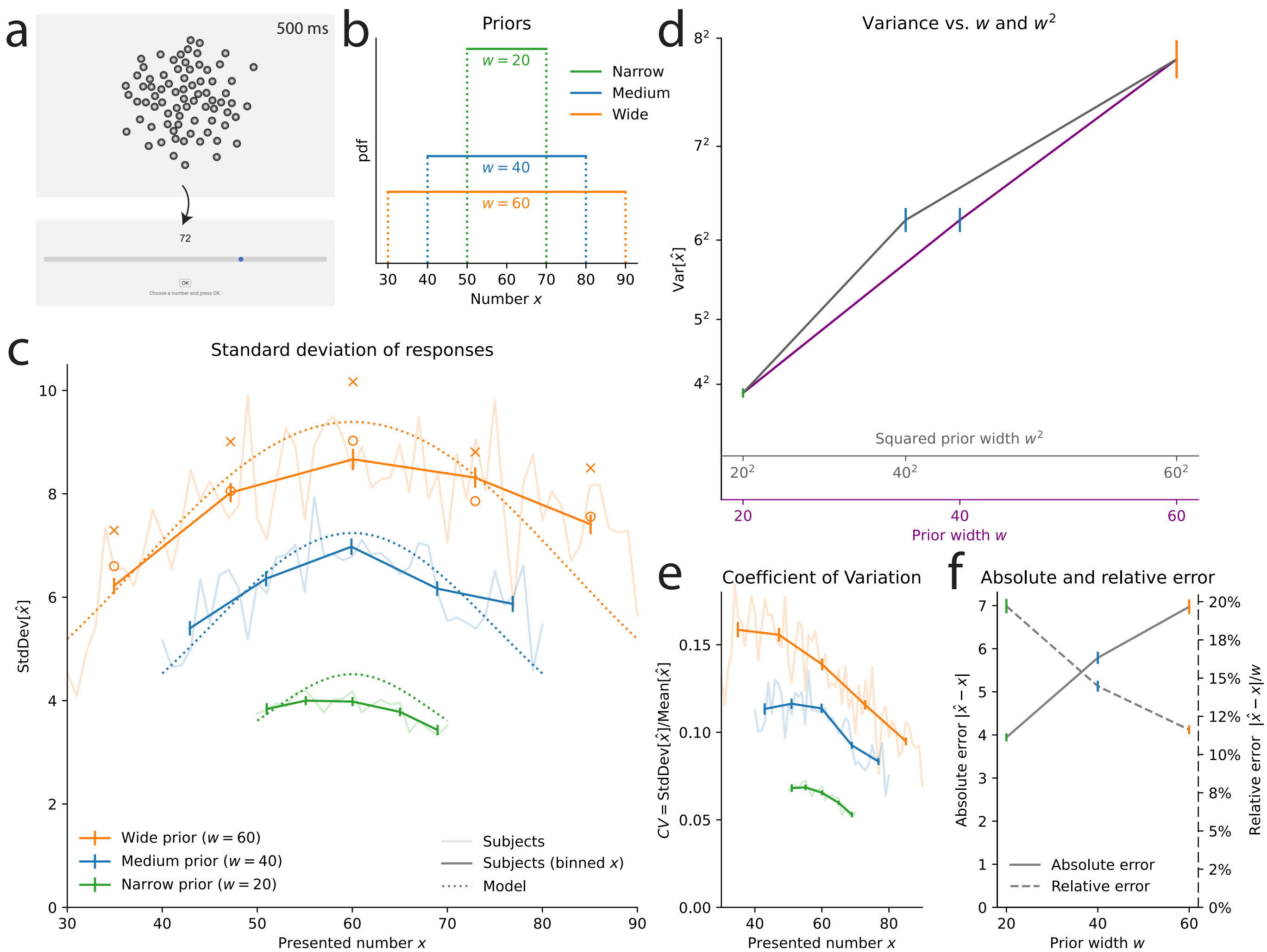

Fig. 1. Estimation task: the scale of subjects' imprecision increases sublinearly with the prior width. a. Illustration of the estimation task: in each trial, a cloud of dots is presented on screen for 500ms. Subjects are then asked to provide their best estimate of the number of dots shown. **b.** Uniform prior distributions (from which the numbers of dots are sampled) in the three conditions of the task. **c.** Standard deviation of the responses of the subjects (solid lines) and of the best-fitting model (dotted lines), as a function of the number of presented dots, in the three conditions. For each prior, five bins of approximately equal sizes are defined; subjects' responses to the numbers falling in each bin are pooled together (thick lines) or not (thin lines). **d.** Variance of subjects' responses, as a function of the width of the prior (purple line) and of the squared width (grey line). Both lines show the same data; only the x-axis scale has been changed. **e.** Subjects' coefficients of variations, defined as the ratio of the standard deviation of estimates over the mean estimate, as a function of the presented number, in the three conditions. **f.** Absolute error (solid line), defined as the absolute difference between a subject's estimate and the correct number, and relative error (dashed line), defined as the ratio of the absolute error to the prior width, as a function of the prior width. In panels **c-d**, the responses of all the subjects are pooled together; error bars show twice the standard errors.

Narrow condition, 6.8 in the Medium condition, and 8.4 in the Wide condition, although these estimates were all obtained after the presentations of the same number of dots (60). Testing for the equality of the variances of estimates across the three conditions, for each number contained in all three priors (i.e., all the numbers in the Narrow range,) we find that the three variances are significantly different, for all the numbers (largest Levene's test p-value, across the numbers: 1e-7, median: 2e-15).

The variability of estimates increases with the width of the prior. This suggests that the imprecision in the internal representation of a number is larger when a larger range of numbers needs to be represented. This would be the case if internal representations relied on a mapping of the range of numbers to a normalized, bounded internal scale, and the estimate of a number resulted from a noisy readout (or a noisy storage) on this scale, as in 'range-normalization' models[37–42]. Consider for instance the representation of a number x, obtained through its normalization onto the unit range $[0, 1]$, and then read with noise, as

$$r = \frac{x - x_{min}}{w} + \varepsilon, \tag{1}$$

where x_{min} is the lowest value of the prior, and ε a centered normal random variable with variance ν^2. Suppose that the estimate, $\hat{x}$, is obtained by rescaling the noisy representation back to the original range, i.e., $\hat{x} = x_{min} + rw$ (we make this assumption for the sake of simplicity, but the argument we develop here is equally relevant for the more elaborate, Bayesian model we present below). The scale of the noise, given by ν, is constant in the normalized scale; thus in the space of estimates the noise scales with the prior width, w. If we allow, in addition to the noise in estimates, for some amount of independent motor noise of variance σ_0^2 in the responses actually chosen by the subject, we obtain a model in which the variance of responses is $\sigma_0^2 + \nu^2 w^2$, i.e., an affine function of the *square* of the width of the prior.

With the numerosity 60, the variance of subjects' estimates is $4.2^2 = 17.64$ in the Narrow condition ($w = 20$), and $6.8^2 = 46.24$ in the Medium condition ($w = 40$): given these two values, the affine relation just mentioned predicts that in the Wide condition ($w = 60$) the variance should be $9.7^2 = 93.91$. We find instead that it is $8.4^2 = 70.56$, i.e., about 25% lower than predicted, suggesting a sublinear relation between the variance and the square of the prior width. Indeed the variance of estimates does not seem to be an affine function of the square of the prior width (Fig. 1d, grey line and grey abscissa). Our investigations reveal that instead, the variance is significantly better captured by an affine function of the width — and not of the squared width (Fig. 1d, purple line and purple abscissa).

As an additional illustration of this result, for each of the five bins mentioned above and defined for the three priors, we compute the predicted variance of estimates in the Wide condition on the basis of the variances in the Narrow and Medium conditions, and resulting either from the hypothesis of an affine function of the squared width, $\sigma_0^2 + \nu^2 w^2$, or from the hypothesis of an affine function of the width, $\sigma_0^2 + \nu^2 w$. The variances predicted with the former hypothesis all overestimate the variances of subjects' responses (Fig. 1c, orange crosses), but the predictions of the latter hypothesis appear consistent with the behavioral data (Fig. 1c, orange circles).

We further investigate how the imprecision in internal representations depends on the width of the prior through a behavioral model in which responses results from a stochastic encoding of the numerosity, followed by a Bayesian decoding step. Specifically, the presentation of a number x results in an internal representation, r, drawn from a Gaussian distribution with mean x and whose standard deviation, νw^α, is proportional to the prior width raised to the power α; i.e., $r|x \sim N(x, \nu^2 w^{2\alpha})$, where ν is a positive parameter that determines the baseline degree of imprecision in the representation, and α is a non-negative exponent that governs the dependence of the imprecision on the width of the prior. The observer derives, from the internal representation r, the mean of the Bayesian posterior over x, $x^*(r) \equiv \mathbb{E}[x|r]$. We note that this estimate minimizes the squared-error loss, and thus maximizes the expected reward in the task. The selection of a response includes an amount of motor noise: the response, $\hat{x}$, is drawn from a Gaussian distribution centered on the Bayesian estimate, $x^*(r)$, with variance σ_0^2, truncated to the prior range, and rounded to the nearest integer. This model has three parameters (σ_0, ν, and α).

The likelihood of the model is maximized for $\alpha = 0.48$, a value close to $1/2$ (and less close to 1), suggesting that the standard deviation is approximately a linear function of $\sqrt{w}$ (and the variance a linear function of w). The nested model obtained by fixing $\alpha = 1/2$ yields a slightly poorer fit (which is expected for a nested model), but the difference in log-likelihood is small (0.38), and the Bayesian Information Criterion (BIC), a measure of fit that penalizes larger numbers of parameters[43], is lower (i.e., better) by 8.70 for the constrained model with $\alpha = 1/2$. This indicates that setting $\alpha = 1/2$ provides a parsimonious fit to the data that is not significantly improved by allowing α to differ from $1/2$. A different specification, $\alpha = 1$, corresponds to a normalization model similar to the one described above, but here with a Bayesian decoding of the internal representation. The BIC of this model is higher by 244 than that with $\alpha = 1/2$, indicating a much worse fit to the data. (Throughout, we report the models' BICs even if they have the same number of parameters, so as to compare the values of a single metric). We emphasize that this large difference in BIC implies that the hypothesis $\alpha = 1$ can be confidently rejected, in favor of the hypothesis $\alpha = 1/2$ (in informal terms, it is not the case that the grey line in Fig. 1d, showing the variance vs. the squared width, only appears curved because of some sampling noise, in fact it is indeed *not* a straight line; while it is substantially more probable that the purple one, showing the variance vs. the width, corresponds indeed to a straight line).

The standard deviation of representations thus seems to increase linearly with the square root of the prior width, $\sqrt{w}$. The positive dependence results in larger errors when the prior is wider (Fig. 1f, solid line). But the sublinear relation implies that the subjects in fact make smaller *relative* errors (relatively to the width of the prior), when the prior is wider. In the Narrow condition, the ratio of the average absolute error to the width of the prior, $\frac{|\hat{x}-x|}{w}$, is 19.7%, i.e., the size of errors is about one fifth of the prior width. This ratio decreases substantially, to 14.5% and 11.6% in the Medium and Wide conditions, respectively, i.e., the size of errors is about one ninth of the prior width in the Wide condition (Fig. 1f, dashed line). In other words, while the size of the prior is multiplied by 3, the relative size of errors is multiplied by $\frac{5}{9} \simeq 0.56$, and thus the absolute size of errors is multiplied by $3 \cdot \frac{5}{9} \simeq 1.67$. If subjects had the same relative sizes of errors in both the Narrow and the Wide conditions, their absolute error would be multiplied by 3; conversely the absolute error would be the same in the two conditions if the relative error was divided by 3. The behavior of subjects falls in between these two scenarios: they adopt smaller relative errors in the Wide condition, although not so much so as to reach the same absolute error as in the Narrow condition. Below, we show how this behavior is accounted for by a tradeoff between the performance in the task and a resource cost on the activity of the mobilized neurons. But first, we ask whether subjects exhibit, in a discrimination task, the same sublinear relation between the imprecision of representations and the width of the prior.

Discrimination task

In many decision situations, instead of providing an estimate, one is required to select the better of two options. We thus investigate experimentally the behavior of subjects in a discrimination task. In each trial, subjects are presented with two interleaved series of numbers, five red and five blue numbers, after which they are asked to choose the series that had the higher average (Fig. 2a). Each number is shown for 500ms. Two experimental conditions differ by the width of the uniform prior from which the numbers (both blue and red) are sampled: in the Narrow condition the range of the prior is $[35, 65]$ (the width of the prior is thus $w = 30$) and in the Wide condition the range is $[10, 90]$ (the width is thus $w = 80$; Fig. 2b). After each decision, subjects receive a number of points equal to the average that they chose. At the end of the experiment, the total sum of their points is converted to a financial reward (through an increasing affine function).

Subjects in this experiment sometimes make incorrect choices (i.e., they choose the color whose numbers had the lower average), but they make less incorrect choices when the difference between the two averages is larger, and the proportion of trials in which they choose 'red' is a sigmoid function of the difference between the average of the red numbers, x_R, and the average of the blue numbers, x_B (Fig. 2c). In the Narrow condition, this proportion reaches 60% when the difference in the averages is 1, and 90% when the difference is 7. In the Wide condition, we find that the slope of this psychometric curve is less steep: subjects reach the same two proportions for differences of about 2.4 and 12.6, respectively.

In the Wide condition, it thus requires a larger difference between the red and blue averages for the subjects to reach the same discrimination threshold; put another way, the same difference in the averages results in more incorrect choices in the Wide condition than in the Narrow condition. As with the estimation task, this suggests that the degree of imprecision in representations is larger when the range of numbers that must be represented is larger. To estimate this quantitatively, we turn to the predictions of the model presented above, here considered in the context of the discrimination task: in this model, the average x_C, where C is 'blue' or 'red' (denoted by B and R, respectively), results in an internal representation, r_C, drawn from a Gaussian distribution with mean x_C and whose variance, $\nu^2 w^{2\alpha}$, is proportional to the prior width raised to the exponent 2α, i.e., $r_C|x_C \sim N\left(x_C, \nu^2 w^{2\alpha}\right)$. Given the (independent) representations r_B and r_R, the subject, optimally, compares the Bayesian estimates for each quantity, $x^*(r_B)$ and $x^*(r_R)$, and chooses the greater one. As the Bayesian estimate is an increasing function of the representation, the probability that the subject choose 'red', conditional on two averages x_B and x_R, is the probability that r_R be larger than r_B, i.e.,

$$P\left(\text{`red'}|x_B, x_R\right) = P\left(r_R > r_B|x_B, x_R\right) = \Phi\left(\frac{x_R - x_B}{\sqrt{2}\nu w^\alpha}\right), \tag{2}$$

where Φ is the cumulative distribution function of the standard normal distribution.

The choice probability is thus predicted to be a function of the ratio $\frac{x_R - x_B}{w^\alpha}$ of the dif-

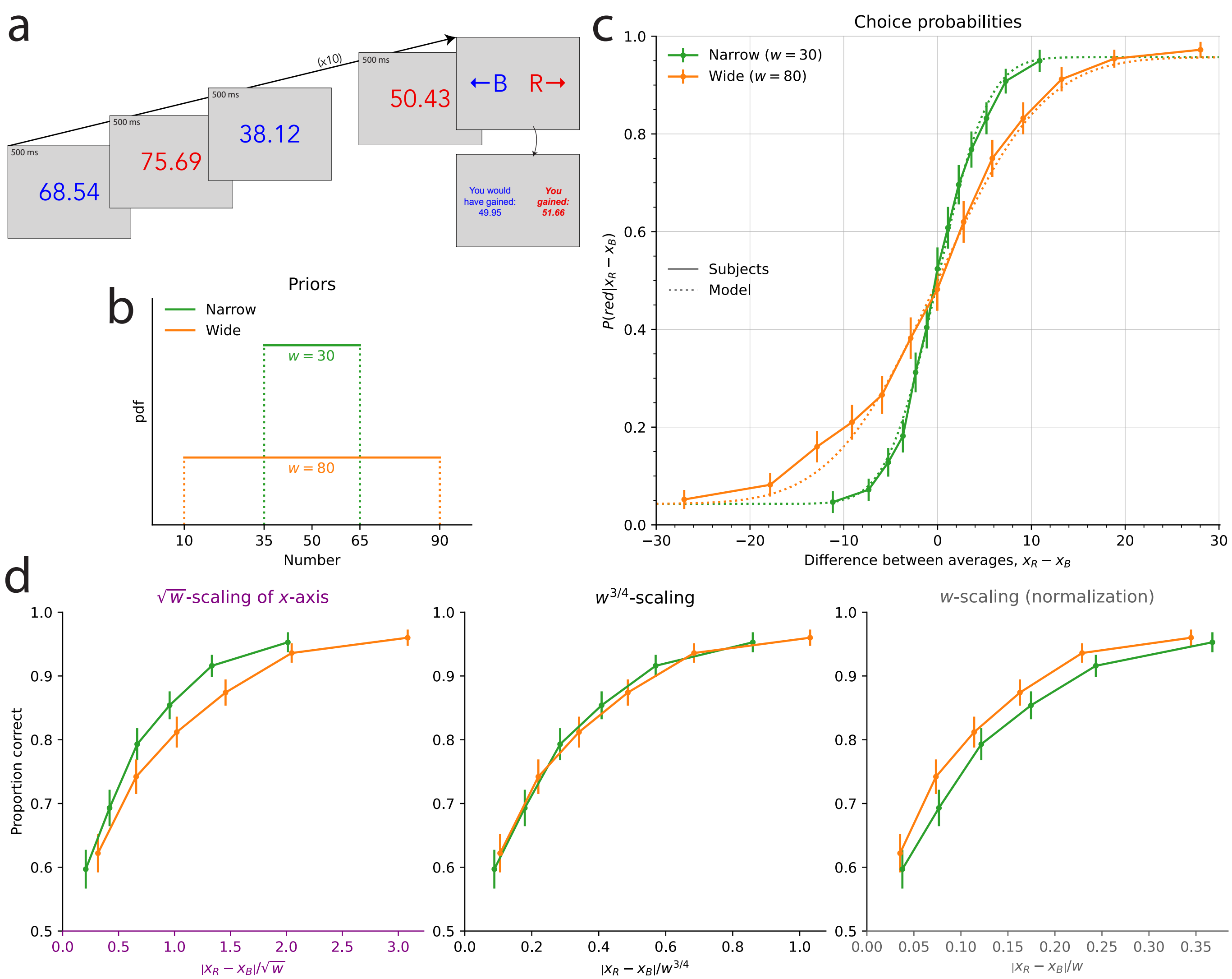

Fig. 2. Discrimination task: the scale of subjects' imprecision increases with the prior width; the relation is sublinear, but different than in the estimation task. **a.** Illustration of the discrimination task: in each trial, subjects are shown five blue numbers and five red numbers, alternating in color, each for 500ms, after which they are asked to choose the color whose numbers have the higher average. **b.** Uniform prior distributions (from which the numbers of dots are sampled) in the two conditions of the task. **c.** Proportion of choices 'red' in the responses of the subjects (solid lines) and of the best-fitting model (dotted lines), as a function of the difference between the two averages, in the two conditions. **d.** Proportion of correct choices in subjects' responses as a function of the absolute difference between the two averages divided by the square root of the prior width (left), by the prior width raised to the power 3/4 (middle), and by the prior width (right). The three subpanels are different representations of the same data. In panels **c** and **d**, the responses of all the subjects are pooled together; error bars show the 95% confidence intervals.

ference between the two averages over the width of the prior raised to the power α, and therefore the same choice probability should be obtained across conditions as long as this ratio is the same. In Figure 2d, we show for different values of α the subjects' proportions of correct responses as a function of the absolute value of this ratio, so as to be able to examine closely the difference between the resulting choice curves in the two conditions. The case $\alpha = 1$ corresponds, as above, to the hypothesis that the standard deviation of internal representations is a linear function of the width, w, i.e., a normalization of the numbers by the width of the prior. But we find that the proportion of correct choices as a function of the ratio $|x_R - x_B|/w$ is greater in the Wide condition than in the Narrow condition (Fig. 2d, last panel). In other words, in the Wide condition the subjects are more sensitive to the normalized difference than in the Narrow condition. This suggests that between the Narrow and the Wide conditions, the imprecision in representations does not change in the same proportions as does the prior width; specifically, it suggests a sublinear relation between the scale of the imprecision and the width of the prior.

As seen in the previous section, the behavioral data in the estimation task precisely suggest such a sublinear relation, and more precisely point to the exponent $\alpha = 1/2$, i.e., to a linear relation between the standard deviation and the square-root of the width, $\sqrt{w}$. But the proportion of correct choices as a function of the corresponding ratio, $|x_R - x_B|/\sqrt{w}$, is greater in the Narrow condition than in the Wide condition (Fig. 2d, first panel). The sublinear relation, thus, is not the same in the two tasks; and the data suggest in the case of the discrimination task an exponent α greater than 1/2, but lower than 1. Indeed, we find that the choice curves in the two conditions match very well with $\alpha = 3/4$ (Fig. 2d, middle panel).

Model fitting substantiates this result. We add to our model (in which the probability of choosing 'red' is given by Eq. 2) the possibility of 'lapse' events, in which either response is chosen with probability 50%; an additional parameter, η, governs the probability of lapses. (We reach the same conclusions with a model with no lapse, but this model with lapses yields a better fit; see Methods.) The BIC of this model with $\alpha = 3/4$ is lower (i.e., better) by 44.1 than that with $\alpha = 1/2$, and by 18.3 than that with $\alpha = 1$, indicating strong evidence rejecting the hypotheses $\alpha = 1/2$ and $\alpha = 1$, in favor instead of the hypothesis of an exponent α equal to 3/4. Notwithstanding the theoretical reasons, presented below, that motivate our focus on this specific value of the exponent in addition to the good fit to the data, we can let α be a free parameter, in which case its best-fitting value is 0.80 (and thus close to 3/4). This model's BIC is however higher (i.e., worse) by 7.9 than that of the model with α fixed at 3/4, which indicates strong evidence[44] in favor of the equality $\alpha = 3/4$. In sum, our best-fitting model is one in which the standard deviation of the internal representations is a linear function of the prior width raised to the power 3/4. As with the estimation task, this sublinear relation implies that subjects are relatively more precise when the prior is wider. This allows them to achieve a significantly better performance in the Wide condition than in the Narrow condition (with 80.2% and 77.4% of correct responses, respectively; p-value of Fisher's exact test of equality of the proportions: 9.5e-5).

Task-optimal endogenous precision

The subjects' behavioral patterns in the estimation task and in the discrimination task suggest that the scale of the imprecision in their internal representations increases sublinearly with the range of numerosities used in a given experimental condition. Specifically, the scale of the imprecision seems to be a linear function of the prior width raised to the power 1/2, in the estimation task, and raised to the power 3/4, in the discrimination task. We now show that these two exponents, 1/2 and 3/4, arise naturally if one assumes that the observer optimizes the expected reward in each task, while incurring a cost on the activity of the neurons that encode the numerosities.

Inspired by models of perception in neuroscience[16–18,20–25,45–47], we consider a two-stage, encoding-decoding model of an observer's numerosity representation. In the encoding stage, a numerosity x elicits in the brain of the observer an imprecise, stochastic representation,

r, while the decoding stage yields the mean of the Bayesian posterior, which is the optimal decoder in both tasks. The model of Gaussian representations that we use throughout the text is one example of such an encoding-decoding model.

The encoding mechanism is characterized by its Fisher information, $I(x)$, which reflects the sensitivity of the representation's probability distribution to changes in the stimulus x. The inverse of the square-root of the Fisher information, $1/\sqrt{I(x)}$, can be understood as the scale of the imprecision of the representation about a numerosity x. More precisely, it is approximately — when $I(x)$ is large — the standard deviation of the Bayesian-mean estimate of x derived from the encoded representation. (For smaller $I(x)$, the standard deviation of the Bayesian-mean estimate increasingly depends on the shape of the prior; with a uniform prior, it decreases near the boundaries.) The variability in subjects' responses in the estimation task, and their choice probabilities in the discrimination task, reported above, are thus indirect measures of the Fisher information of their encoding process.

Moreover, the expected squared error of the Bayesian-mean estimate of x is approximately the inverse of the Fisher information, $1/I(x)$. We thus consider the generalized loss function

$$L_a[I] = \int \frac{\pi(x)^a}{I(x)} \mathrm{d}x, \tag{3}$$

where $\pi(x)$ is the prior distribution from which x is sampled. With $a = 1$, this quantity approximates the expected quadratic loss that subjects in the estimation task should minimize in order to maximize their reward. And with $a = 2$, minimizing this loss is approximately equivalent to maximizing the reward in the discrimination task[24]. (The squared prior, in the expression of $L_2[I]$, corresponds to the probability of the co-occurrence of two presented numerosities that are close to each other, which is the kind of event most likely to result in errors in discrimination.)

In both cases, a more precise encoding, i.e., a greater Fisher information, results in a smaller loss. This precision, however, comes with a cost. We assume that the encoding results from an accumulation of signals, each entailing an identical cost (e.g., the energy resources consumed by action potentials[32–34].) The more signals the observer collects, the greater the precision; but also the greater the cost, which is proportional to the number of signals. Formally, we consider a continuum-limit model, in which a representation proceeds from a Wiener process (Brownian motion) with infinitesimal variance s^2, observed for a duration T (the continuum equivalent of the number of collected signals). The drift of the process, $m(x)$, encodes the number: it can be, for instance, some normalized value of x; but here we only assume that the function $m(x)$ is increasing and bounded. The resulting representation, r, is normally distributed, as $r|x \sim N(m(x)T, s^2T)$, and its Fisher information is $T(m'(x))^2/s^2$ and thus it is proportional to T. The bound on $m(x)$ puts a constraint on the Fisher information: specifically, it implies that the quantity

$$C[I] = \left(\int \sqrt{I(x)} \mathrm{d}x \right)^2 \tag{4}$$

is bounded by a quantity proportional to the duration, i.e., $C[I] \leq KT$, where $K > 0$. Other studies[18,21,24] have posited a bound on the quantity $C[I]$, but here we emphasize that the bound is a linear function of the duration of observation, and we assume, crucially, that the observer can choose this duration, T, but at the expense of a cost that is proportional to T.

Specifically, we assume that the observer chooses the function $I(.)$ and the duration T that solve the minimization problem

$$\min_{I(.),T} L_a[I] + \lambda T \text{ subject to } C[I] \leq KT, \tag{5}$$

where $\lambda > 0$. In this problem, any increase of the Fisher information, within the bound, improves the objective function; and thus the solution saturates the bound, i.e., $C[I] = KT$. Hence the problem reduces to that of choosing the function $I(.)$ that solves the minimization problem

$$\min_{I(.)} L_a[I] + \theta C[I], \tag{6}$$

where $\theta = \lambda/K$. The solution is

$$I(x) = \frac{\pi(x)^{2a/3}}{\sqrt{\theta \int \pi(\tilde{x})^{a/3} \mathrm{d}\tilde{x}}}. \tag{7}$$

This implies that the optimal Fisher information vanishes outside of the support of the prior; and in the case of a uniform prior of width w, $I(x)$ is constant, as

$$\begin{aligned} I(x) &= \frac{1}{\sqrt{\theta} w} \qquad \text{for the estimation task,} \\ \text{and } I(x) &= \frac{1}{\sqrt{\theta} w^{3/2}} \quad \text{for the discrimination task,} \end{aligned} \tag{8}$$

for any x such that $\pi(x) \neq 0$.

The scale of the imprecision of internal representations, $1/\sqrt{I(x)}$, is thus predicted to be proportional to the prior width raised to the power 1/2, in the estimation task, and raised to the power 3/4, in the discrimination task. As shown above, we find indeed that in these tasks, the imprecision of representations not only increases with the prior width, but it does so in a way that is quantitatively consistent with these two exponents. As for the model of Gaussian representations that we have considered throughout the text, it is in fact equivalent to the model just presented, up to a linear transformation of the representation that does not impact its Fisher information (nor the resulting estimates). Its Fisher information is the inverse of the variance, i.e., $1/\left(\nu^2 w^{2\alpha}\right)$, and thus Eq. 8 implies $\alpha = 1/2$ for the estimation task, and $\alpha = 3/4$ for the discrimination task, i.e., the two values that indeed best fit the data.

Many efficient-coding models in the literature feature a different objective, the maximization of the mutual information[18–20]; but a single objective cannot explain our different findings in the two tasks (namely, the different dependence on the prior width). Many models also feature a different kind of constraint: a *fixed* bound on the quantity in Eq. 4, or on a generalization of this quantity[18,19,21,23]. But here also, as this bound is usually saturated, the optimal Fisher information, which is constant, here, due to the uniform prior, is entirely determined by the constraint—irrespective of the objective of the task. This hypothesis thus cannot account either for the difference that we find between the two tasks. By contrast, we assume that it is the task's expected reward that is maximized, and that the amount of utilized encoding resources can be endogenously determined: our model is thus able to predict not only that the behavior should depend on the prior, but also that this dependence should change with the task; and it makes quantitative predictions that coincide with our experimental findings.

We compare the responses of the subjects and of the Gaussian-representation model, with $\alpha = 1/2$ in the estimation task and $\alpha = 3/4$ in the discrimination task. In both cases, the parameter ν governs the imprecision in the internal representation, and a second parameter corresponds to additional response noise: the motor noise, parameterized by σ_0^2, in the estimation task, and the lapse probability, η, in the discrimination task. The behavior of the model, across the two tasks and the different priors, reproduces that of the subjects (Figs. 1c and 2c, dotted lines). In the estimation task, the standard deviation of estimates increases as a function of the prior width, as it does in subjects' responses. The Fisher information in this model is constant with respect to x, and thus the variance of the internal representation, r, is also constant; but the Bayesian estimate, $x^*(r)$, depends on the prior, and its variability decreases for numerosities closer to the edges of the uniform prior. Hence the standard deviation of the model's estimates adopts an inverted U-shape similar to that of the subjects (Fig. 1c). In the discrimination task, the model's choice-probability curve is steeper in the Narrow condition than in the Wide condition, and the two predicted curves are close to the subjects' choice probabilities (Fig. 2c). We emphasize that how the internal imprecision scales with the prior width is entirely determined by our theoretical predictions (Eq. 8); these quantitative predictions allow our model to capture the subjects' imprecise responses simultaneously across different priors.

Discussion

In this study, we examine the variability in subjects' responses in two different tasks and with different priors. We find that the precision of their responses depends both on the task and on the prior. The scale of their imprecision about the presented numbers increases sublinearly with the width of the prior, and this sublinear relation is different in each task. The two sublinear relations are predicted by a resource-rational account, whereby the allocation of encoding resources optimizes a tradeoff, maximizing each task's expected reward while incurring a cost on the activity of the encoding neurons. Different formalizations of this tradeoff suggested in several other studies cannot reproduce our experimental findings.

The model and the data suggest a scaling law relating the size of the representations' imprecision to the width of the prior, with an exponent that depends on the task at hand. An important implication is that the relative precision with which people represent external information can be modulated by their objective and by the manner and the context in which the representations are elicited. In the model, the scaling law results from the solution to the encoding allocation problem (Eq. 6) in the special case of a uniform prior, and in the contexts of estimation and discrimination tasks. We surmise that with non-uniform priors and with other tasks (that imply different expected-reward functions), the behavior of subjects should be consistent with the optimal solution to the corresponding resource-allocation problem, provided that subjects are able to learn these other priors and objectives. Further investigations of this conjecture will be crucial in order to understand the extent to which the formalism of optimal resource-allocation that we present here might form a fundamental component in a comprehensive theory of the brain's internal representations of magnitudes.

References

[1] Tobias Bothwell, Colin J. Kennedy, Alexander Aeppli, Dhruv Kedar, John M. Robinson, Eric Oelker, Alexander Staron, and Jun Ye. Resolving the gravitational redshift across a millimetre-scale atomic sample. *Nature*, 602(7897):420–424, 2022.

[2] L. L. Thurstone. Psychophysical Analysis. *The American Journal of Psychology*, 38(3):368, jul 1927.

[3] Wilson P. Tanner and John A. Swets. A decision-making theory of visual detection. *Psychological Review*, 61(6):401–409, 1954.

[4] George A. Gescheider. *Psychophysics*. Psychology Press, jun 1997.

[5] D. H. Hubel and T. N. Wiesel. Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, 148(3):574–591, oct 1959.

[6] G. H. Henry, B. Dreher, and P. O. Bishop. Orientation specificity of cells in cat striate cortex. *Journal of Neurophysiology*, 37(6):1394–1409, 1974.

[7] KH Britten, MN Shadlen, WT Newsome, and JA Movshon. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *The Journal of Neuroscience*, 12(12):4745–4765, dec 1992.

[8] E. L. Kaufman, M. W. Lord, T. W. Reese, and J. Volkmann. The Discrimination of Visual Number. *The American Journal of Psychology*, 62(4):498, oct 1949.

[9] Robert S. Moyer and Thomas K. Landauer. Time required for Judgements of Numerical Inequality. *Nature*, 215(5109):1519–1520, sep 1967.

[10] Andreas Nieder and Earl K. Miller. Coding of cognitive magnitude: Compressed scaling of numerical information in the primate prefrontal cortex. *Neuron*, 37(1):149–157, 2003.

[11] Esther F. Kutter, Jan Bostroem, Christian E. Elger, Florian Mormann, and Andreas Nieder. Single Neurons in the Human Brain Encode Numbers. *Neuron*, 100(3):753–761.e4, 2018.

[12] Stanislas Dehaene. *The Number Sense: How the Mind Creates Mathematics.* Oxford University Press, New York, dec 2011.

[13] H. B. Barlow. Possible Principles Underlying the Transformations of Sensory Messages. In Walter A. Rosenblith, editor, *Sensory Communication*, chapter 13, pages 217–234. The MIT Press, Cambridge, MA, sep 1961.

[14] Nicolas Brunel and Jean-Pierre Nadal. Optimal tuning curves for neurons spiking as a Poisson process. *Proceedings of the ESANN Conference*, 1997.

[15] Mark D. McDonnell and Nigel G. Stocks. Maximally Informative Stimuli and Tuning Curves for Sigmoidal Rate-Coding Neurons and Populations. *Physical Review Letters*, 101(5):058103, aug 2008.

[16] Deep Ganguli and Eero P Simoncelli. Implicit encoding of prior probabilities in optimal neural populations. In J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, and A. Culotta, editors, *Advances in Neural Information Processing Systems 23*, pages 658—-666. Curran Associates, Inc., 2010.

[17] Deep Ganguli and Eero P. Simoncelli. Efficient Sensory Encoding and Bayesian Inference with Heterogeneous Neural Populations. *Neural Computation*, 26(10):2103–2134, oct 2014.

[18] Xue-Xin Wei and Alan A. Stocker. A Bayesian observer model constrained by efficient coding can explain 'anti-Bayesian' percepts. *Nature Neuroscience*, 18(10):1509–1517, oct 2015.

[19] Xue-Xin Wei and Alan A Stocker. Mutual Information, Fisher Information, and Efficient Coding. *Neural Computation*, 326:305–326, 2016.

[20] Deep Ganguli and Eero P. Simoncelli. Neural and perceptual signatures of efficient sensory coding. *ArXiv e-prints*, pages 1–24, feb 2016.

[21] Zhuo Wang, Alan A. Stocker, and Daniel D. Lee. Efficient Neural Codes That Minimize Lp Reconstruction Error. *Neural Computation*, 28(12):2656–2686, dec 2016.

[22] Il Memming Park and Jonathan W. Pillow. Bayesian Efficient Coding. *bioRxiv*, 2017.

[23] Michael Morais and Jonathan W Pillow. Power-law efficient neural codes provide general link between perceptual bias and discriminability. *Advances in Neural Information Processing Systems 31*, 2(1):5076–5085, 2018.

[24] Arthur Prat-Carrabin and Michael Woodford. Bias and variance of the Bayesian-mean decoder. In M Ranzato, A Beygelzimer, Y Dauphin, P S Liang, and J Wortman Vaughan, editors, *Advances in Neural Information Processing Systems*, volume 34, pages 23793–23805. Curran Associates, Inc., 2021.

[25] Ling Qi Zhang and Alan A. Stocker. Prior Expectations in Visual Speed Perception Predict Encoding Characteristics of Neurons in Area MT. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 42(14):2951–2962, 2022.

[26] Ahna R Girshick, Michael S Landy, and Eero P Simoncelli. Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience*, 14(7):926–932, jul 2011.

[27] Samuel J. Cheyette and Steven T. Piantadosi. A unified account of numerosity perception. *Nature Human Behaviour*, 2020.

[28] Arthur Prat-Carrabin and Michael Woodford. Efficient coding of numbers explains decision bias and noise. *Nature Human Behaviour*, pages 845–848, may 2022.

[29] Falk Lieder and Thomas L. Griffiths. Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, 43:e1, feb 2020.

[30] Wei Ji Ma and Michael Woodford. Multiple conceptions of resource rationality. *Behavioral and Brain Sciences*, 43:e15, mar 2020.

[31] Jonathan Schaffner, Sherry Dongqi Bao, Philippe N. Tobler, Todd A. Hare, and Rafael Polania. Sensory perception relies on fitness-maximizing codes. *Nature Human Behaviour*, 7(7):1135–1151, 2023.

[32] S B Laughlin, R R de Ruyter Van Steveninck, and J C Anderson. The metabolic cost of neural information. *Nature neuroscience*, 1(1):36–41, 1998.

[33] Andrea Hasenstaub, Stephani Otte, Edward Callaway, and Terrence J. Sejnowski. Metabolic cost as a unifying principle governing neuronal biophysics. *Proceedings of the National Academy of Sciences of the United States of America*, 107(27):12329–12334, 2010.

[34] Biswa Sengupta, Martin Stemmler, Simon B. Laughlin, and Jeremy E. Niven. Action potential energy efficiency varies among neuron types in vertebrates and invertebrates. *PLoS Computational Biology*, 6(7):35, 2010.

[35] Véronique Izard and Stanislas Dehaene. Calibrating the mental number line. *Cognition*, 106(3):1221–1247, 2008.

[36] Alberto Testolin and James L. McClelland. Do estimates of numerosity really adhere to Weber's law? A reexamination of two case studies. *Psychonomic Bulletin and Review*, 28(1):158–168, 2021.

[37] Camillo Padoa-Schioppa. Range-adapting representation of economic value in the orbitofrontal cortex. *Journal of Neuroscience*, 29(44):14004–14014, 2009.

[38] Shunsuke Kobayashi, Ofelia Pinto de Carvalho, and Wolfram Schultz. Adaptation of Reward Sensitivity in Orbitofrontal Neurons. *The Journal of Neuroscience*, 30(2):534–544, jan 2010.

[39] Xinying Cai and Camillo Padoa-Schioppa. Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex. *Journal of Neuroscience*, 32(11):3791–3808, 2012.

[40] Alireza Soltani, Benedetto De Martino, and Colin Camerer. A Range-Normalization Model of Context-Dependent Choice: A New Model and Evidence. *PLoS Computational Biology*, 8(7):e1002607, jul 2012.

[41] Antonio Rangel and John A. Clithero. Value normalization in decision making: Theory and evidence. *Current Opinion in Neurobiology*, 22(6):970–981, 2012.

[42] Kenway Louie and Paul W. Glimcher. Efficient coding and the neural representation of value. *Annals of the New York Academy of Sciences*, 1251(1):13–32, mar 2012.

[43] Gideon Schwarz. Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2):461–464, mar 1978.

[44] Robert E. Kass and Adrian E. Raftery. Bayes Factors. *Journal of the American Statistical Association*, 90(430):773–795, jun 1995.

[45] Norberto M. Grzywacz and Rosario M. Balboa. A Bayesian framework for sensory adaptation. *Neural Computation*, 14(3):543–559, 2002.

[46] Alan A. Stocker and Eero P. Simoncelli. Sensory adaptation within a Bayesian framework for perception. *Advances in Neural Information Processing Systems*, 18:1291–1298, 2006.

[47] Alan A. Stocker and Eero P. Simoncelli. Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience*, 9(4):578–585, apr 2006.

Methods

Estimation task

Task and subjects 36 subjects (20 female, 15 male, 1 non-binary) participated in the estimation-task experiment (average age: 21.4, standard deviation: 2.8). The experiment took place at Columbia University, and complied with the relevant ethical regulations; it was approved by the university's Institutional Review Board (protocol number: IRB-AAAS8409). All subjects experienced the three conditions.

In the experiment, subjects provide their responses using a slider (Fig. 1a), whose size on screen is proportional to the width of the prior. Each condition comprises three different phases. In all the trials of all three phases the numerosities are randomly sampled from the prior corresponding to the current condition. This prior is explicitly told to the subject when the condition starts. In each of the 15 trials of the first, 'learning' phase, the subject is shown a cloud of dots together with the number of dots it contains (i.e., its numerosity represented with Arabic numerals). These elements stay on screen until the subject chooses to move on to the next trial. No response is required from the subject in this phase. Then follow the 30 trials of the 'feedback' phase, in which clouds of dots are shown for 500ms without any other information on their numerosities. The subject is then asked to provide an estimate of the numerosity. Once the estimate is submitted, the correct number is shown on screen. The third and last phase is the 'no-feedback' phase, which is identical to the 'feedback' phase, except that no feedback is provided. In both the 'feedback' phase and the 'no-feedback' phase, subjects respond at their own pace. All the analyses presented here use the data of the 'no-feedback' phase, which comprises 120 trials.

At the end of the experiment, subjects receive a financial reward equal to the sum of a \$5 show-up fee (USD) and of a performance bonus. After each submission of an estimate, an amount equal to $0.10 - (\hat{x} - x)^2/600$, where x is the correct number and $\hat{x}$ the estimate, is added to the performance bonus. If at the end of the experiment the performance bonus is negative, it is set to zero. The average reward was \$11.80 (standard deviation: 6.98).

Bins defined over the priors, and calculation of the variance The ranges of the three priors (50-70, 40-80 and 30-90), contain 21, 41, and 61 integers, respectively, and thus none of them can be split in five bins containing the same number of integers. Hence the ranges defining each of the five bins were chosen such that the third bin contains an odd number of integers, with at its middle the middle number of the prior (60 in each case), and such that the second and fourth bins contain the same number of integers as the third one; the first and last bins then contain the remaining integers. In the Narrow condition, the ranges of the five bins are: 50-52, 53-57, 58-62, 63-67, and 68-70. In the Medium condition, the ranges of the five bins are: 40-46, 47-55, 56-64, 65-73, and 74-80. In the Wide condition, the ranges of the five bins are: 30-40, 41-53, 54-66, 67-79, and 80-90.

In our calculation of the variance of estimates, when pooling responses by bins of presented numbers, we do not wish to include the variability stemming from the diversity of numbers in each bin. Thus we subtract from each estimate $\hat{x}$ of a number the average of all the estimates obtained with the same number, $\langle \hat{x} \rangle$. The calculation of the variance for a bin then makes use of these 'excursions' from the mean estimates, $\hat{x} - \langle \hat{x} \rangle$.

Model fitting and individual subjects analysis The Gaussian-representation model used throughout the text has three parameters: α, ν, and σ_0. We fit these parameters to the subjects' data by maximizing the model's likelihood. For each parameter, we can either

α	ν	σ_0	Num. param.	BIC
Fixed $\alpha = 1$	Shared	Shared	2	81762.79
Fixed $\alpha = 1/2$	Shared	Shared	2	*81519.07
Shared ($\alpha = .48$)	Shared	Shared	3	81527.78
Fixed $\alpha = 1$	Indiv.	Shared	37	81729.64
Fixed $\alpha = 1$	Shared	Indiv.	37	81746.34
Fixed $\alpha = 1$	Indiv.	Indiv.	72	81657.67
Fixed $\alpha = 1/2$	Indiv.	Shared	37	81427.11
Fixed $\alpha = 1/2$	Shared	Indiv.	37	81467.71
Fixed $\alpha = 1/2$	Indiv.	Indiv.	72	*81346.37
Shared ($\alpha = .43$)	Indiv.	Shared	38	81437.93
Shared ($\alpha = .45$)	Shared	Indiv.	38	81472.85
Shared ($\alpha = .44$)	Indiv.	Indiv.	73	81350.90
Indiv.	Shared	Shared	38	81444.60
Indiv.	Indiv.	Shared	73	81571.48
Indiv.	Shared	Indiv.	73	81366.40
Indiv.	Indiv.	Indiv.	108	81453.52

Table 1. Estimation task: model fitting supports the hypothesis $\alpha = 1/2$, both with pooled and individual responses. Number of parameters (second-to-last column) and BIC (last column) of the Gaussian-representation model under different specifications regarding whether all subjects share the same values of the three parameters α, ν, and σ_0 (first three columns). 'Shared' indicates that the responses of all the subjects are modeled with the same value of the parameter. 'Indiv.' indicates that different values of the parameter are allowed for different subjects. For the parameter α, 'Fixed' indicates that the value of α is fixed (thus it is not a free parameter); when the parameter α is 'Shared', it is a free parameter, and we indicate its best-fitting value in parentheses. In the first three lines of the table, all three parameters are shared across the subjects (the three lines differ only by the specification of α); while in the remaining lines at least one parameter is individually fit. In both cases the lowest BIC (indicated by a star) is obtained for a model with a fixed parameter $\alpha = 1/2$.

allow for 'individual' values of the parameter that may be different for different subjects, or we can fit the responses of all the subjects with the same, 'shared' value of the parameter. In the main text we discuss the model with 'shared' parameters; the corresponding BICs are shown in the first three lines of Table 1. The other lines of the Table correspond to specifications of the model in which at least one parameter is allowed to take 'individual' values. In both cases the lowest BIC is obtained for models with a fixed exponent $\alpha = 1/2$, common to all the subjects, consistently with our prediction (Eq. 8). Overall, the best-fitting model allows for 'individual' values of the parameters ν and σ_0, and a fixed, shared value for α. This suggests that the parameters ν and σ_0, which govern, respectively, the degrees of "internal" and "external" (motor) imprecision, capture individual traits characteristic of each subject, while the exponent α reflects the solution to the optimization problem posed by the task, which is the same for all the subjects.

Discrimination task

Task and subjects 111 subjects (61 male, 50 female) participated in the discrimination-task experiment (average age: 31.4, standard deviation: 10.2). Due to the COVID crisis, the experiment was run online, and each subject experienced only one condition. 31 subjects participated in the Narrow condition, and 32 subjects participated in the Wide condition. This experiment was approved by Columbia University's Internal Review Board (protocol number: IRB-AAAR9375).

In this experiment, each condition starts with 20 practice trials. In each of these trials, five red numbers and five blue numbers are shown to the subject, each for 500ms. In the first 10 practice trials, no response is asked from the subject. In the following 10 practice trials, the subject is asked to choose a color; choices in these trials do not impact the reward. Then follow 200 'real' trials in which the averages chosen by the subject are added to a score. At the end of the experiment, the subject receives a financial reward that is the sum of a \$1.50 fixed fee (USD) and of a non-negative variable bonus. The variable bonus is equal to $\max(0, 1.6(\text{AverageScore} - 50))$, where AverageScore is the score divided by 200. The average reward was \$6.80 (standard deviation: 2.15).

Individual subjects analysis In the Gaussian-representation model, a numerosity x yields a representation that is normally-distributed, as $r|x \sim N(x, \nu^2 w^{2\alpha})$. Fitting the model to the pooled data collected in the two conditions has enabled us to identify separately the two parameters ν and α. But fitting to the responses of individual subjects, who experienced

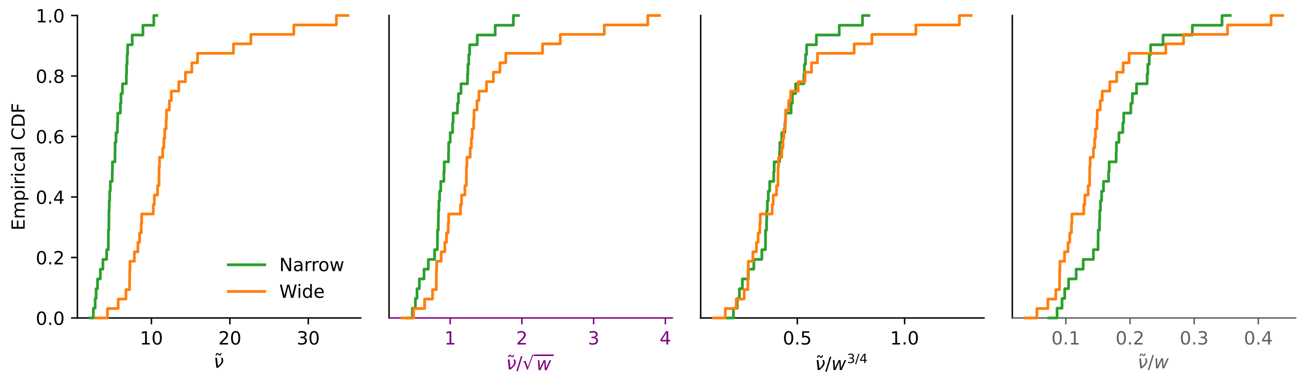

Fig. 3. Discrimination task: empirical across-subjects distribution of scaled best-fitting standard-deviation parameter. The first panel shows the empirical cumulative distribution function (CDF) of the fitted parameter $\tilde{\nu}$, unscaled. The second, third, and fourth panels show the empirical CDF of $\tilde{\nu}$ divided by w^α, with $\alpha = 1/2$, $3/4$, and 1, respectively.

only one of the two conditions, only allows to identify the variance $\tilde{\nu}^2 \equiv \nu^2 w^{2\alpha}$, and not ν and α separately. However, an important difference between these two parameters is that the baseline variance ν^2 is idiosyncratic to each subject (and thus we expect inter-subject variability for this parameter), while the exponent α, in our theory, is determined by the specifics of the task, and thus it should be the same for all the subjects; in particular, we predict $\alpha = 3/4$. Therefore, as subjects were randomly assigned to one of the two conditions, we expect the distribution of $\nu = \tilde{\nu}/w^\alpha$ to be identical across the two conditions. We thus look at the empirical distributions of this quantity, with different values of α, in the two conditions. We find that the distributions of $\tilde{\nu}$, $\tilde{\nu}/\sqrt{w}$, and $\tilde{\nu}/w$, in the two conditions, do not match well; but the distributions of $\tilde{\nu}/w^{3/4}$ in the two conditions are close to each other (Fig. 3). In each of these four cases, we run a Kolmogorov-Smirnov test of the equality of the underlying distributions. With $\tilde{\nu}$, $\tilde{\nu}/\sqrt{w}$, and $\tilde{\nu}/w$, the null hypothesis is rejected (p-values: 1e-10, 0.008, and 0.001, respectively), while with $\tilde{\nu}/w^{3/4}$ the hypothesis (of equality of the distributions in the two conditions) is not rejected (p-value: 0.79). Thus this analysis, based on the individual model-fitting of the subjects, substantiates our conclusions.

Models' BICs We fit the Gaussian-representation model, with or without lapses, to the subjects' responses in the discrimination task. In the main text we discuss the model-fitting results of the model with lapses. The corresponding BICs are reported in the last four lines of Table 2, while the first four lines report the BICs of the model with no lapses. Table 2 shows that including lapses in the model yields lower BICs, but also that in both cases (with or without lapses), the lowest BIC is obtained with the model with a fixed parameter $\alpha = 3/4$, consistently with our theoretical prediction (Eq. 8).

α	Lapses	Num. param.	BIC
Fixed $\alpha = 1$	No	1	11737.03
Fixed $\alpha = 3/4$	No	1	*11721.22
Fixed $\alpha = 1/2$	No	1	11815.86
Free ($\alpha = .84$)	No	2	11723.22
Fixed $\alpha = 1$	Yes	2	11635.59
Fixed $\alpha = 3/4$	Yes	2	*11617.24
Fixed $\alpha = 1/2$	Yes	2	11661.35
Free ($\alpha = .80$)	Yes	3	11625.14

Table 2. Discrimination task: model fitting supports the hypothesis $\alpha = 3/4$. Number of parameters (second-to-last column) and BIC (last column) of the Gaussian-representation model under different specifications regarding the parameter α (first column) and the absence or presence of lapses (second column). In the bottom four lines the model features lapses, while it does not in the top four lines; in both cases the lowest BIC (indicated with a star) is obtained with the specification $\alpha = 3/4$.

Data availability statement

Requests for the data can be sent via email to the corresponding author.

Code availability statement

Requests for the code used for all analyses can be sent via email to the corresponding author.

Acknowledgments

We thank Jessica Li and Maggie Lynn for their help as research assistants, Hassan Afrouzi for helpful comments, and the National Science Foundation for research support (grant SES DRMS 1949418).

Competing interest declaration

The authors declare no conflict of interest.