nature human behaviour

Article

https://doi.org/10.1038/s41562-023-01709-3

Distinct neuronal representation of small and large numbers in the human medial temporal lobe

Received: 12 February 2023

Accepted: 31 August 2023

Published online: 02 October 2023

Check for updates

Esther F. Kutter[1,2], Gert Dehnen[1], Valeri Borger[3], Rainer Surges[1], Florian Mormann[1,4] ✉ & Andreas Nieder[2,4] ✉

Whether small numerical quantities are represented by a special subitizing system that is distinct from a large-number estimation system has been debated for over a century. Here we show that two separate neural mechanisms underlie the representation of small and large numbers. We performed single neuron recordings in the medial temporal lobe of neurosurgical patients judging numbers. We found a boundary in neuronal coding around number 4 that correlates with the behavioural transition from subitizing to estimation. In the subitizing range, neurons showed superior tuning selectivity accompanied by suppression effects suggestive of surround inhibition as a selectivity-increasing mechanism. In contrast, tuning selectivity decreased with increasing numbers beyond 4, characterizing a ratio-dependent number estimation system. The two systems with the coding boundary separating them were also indicated using decoding and clustering analyses. The identified small-number subitizing system could be linked to attention and working memory that show comparable capacity limitations.

When asked to judge the number of briefly presented items in a set, humans show a behavioural dichotomy[1]. Participants respond fast and accurately for small numbers up to about 4 in a process termed 'subitizing'[2]. However, for larger numbers beyond 4, participants show increasingly slower and more imprecise number 'estimation' that is dependent on the ratio between the numbers to be compared[2–5].

On the basis of behavioural measures, it has been argued that the observed judgement differences arise from one and the same estimation system whose negligible ratio-dependent imprecision for small numbers gives rise to a seeming dichotomy in underlying mechanisms[6,7]. Others, in contrast, maintain that subitizing and estimation reflect two distinct mechanisms for assessing small versus large numbers[2–5]. Explorations into underlying brain mechanisms using blood flow imaging or electroencephalography remained similarly inconclusive; while some studies argue for a single underlying mechanism[8–12], others propose two separable number systems[13–15].

In this Article, to address this century-old debate about a single or two distinct mechanisms for number representations, we recorded single-neuron activity in the medial temporal lobe (MTL) of neurosurgical patients who judged numerical quantity[16,17]. If small and large numbers are represented by the same neuronal mechanism, a continuous code across small and large numbers is anticipated. However, if small and large numerosities engage distinct mechanisms, two different coding schemes with a discontinuity reflecting the change from one mechanism to the other is expected.

Results

We asked 17 human participants to quickly judge the parity (even versus odd) of numbers ranging from 0 to 9 shown as dot arrays on a computer screen. The simple parity task is suited to test a broad range of explicit number representations devoid of other cognitive factors (such as working memory), and in short time for the participants. In each trial, a numerosity

[1]Department of Epileptology, University of Bonn Medical Center, Bonn, Germany. [2]Animal Physiology, Institute of Neurobiology, University of Tübingen, Tübingen, Germany. [3]Department of Neurosurgery, University of Bonn Medical Center, Bonn, Germany. [4]These authors contributed equally: Florian Mormann, Andreas Nieder. ✉e-mail: florian.mormann@ukbonn.de; andreas.nieder@uni-tuebingen.de

https://doi.org/10.1038/s41562-023-01709-3

was flashed for 500 ms after a short fixation period, followed by a brief delay during which the number stimulus was removed (Fig. 1a). Afterwards, participants decided whether the number of dots had been even or odd by pressing the left or right arrow key, respectively, on the keyboard as indicated on the response screen. The keys associated with the respective response were switched between blocks to control for potential motor bias. Different stimulus protocols were used to control for non-numerical visual parameters: dots were shown in a standard (variable dot size and arrangement) and two control displays (constant total dot area and dot density across numerosities, and linear arrangement) (Fig. 1b). The numerosity and protocol of the stimuli varied randomly from trial to trial.

Behaviour

The participants' performance showed well-known behavioural effects indicative of two different representational systems. Small countable numerosities from 1 to 4 were equally effortlessly judged with only few errors (Fig. 1c) and short reaction times (RTs) (Fig. 1d), as expected for subitizing[2–4]. In contrast, numbers 5 and higher were judged with noticeably increasing error rates and RTs indicative of number estimation. This observation was bolstered by calculating the discontinuity point that signals a change in the slope[18], which could be determined for 14 of the 17 participants. We found average discontinuity points of 3.7 and 3.6 for error rates and RTs, respectively, as the upper boundary of the subitizing range (Fig. 1c,d). The errors seen for numbers larger than 5 argue that participants were not symbolically counting items as serial counting would be error-free and moreover has been shown to be impossible in afterimages[19,20]. Asymmetric switch cost effects for the transition from subitizing to estimation versus the transition from estimation to subitizing were not observed[21,22] ($P_{\text{switch condition}} = 0.88$; two-factor analysis of variance (ANOVA) with factors 'numerical value' (0–9) × 'switch condition' (switch versus non-switch)). Consistent with previous reports[23,24], the empty set (number zero) elicited distinct behavioural effects due to its special status as a latecomer in number concepts[25].

Neuronal responses

To test the long-standing hypothesis of different enumeration systems for small versus large numbers, we recorded action potentials of 801 single neurons in the MTL of the 17 participants while they performed the number task. Many neurons were activated in a tuned fashion to the numerical value of the sample stimulus. They responded strongest to their respective preferred numerosities and decreased their activity progressively with increasing numerical distance (Fig. 2a–d). We statistically identified number-selective neurons by applying a sliding-window analysis to all cells[16]. We combined a two-factor ANOVA with factors 'numerical value' (0–9) × 'protocol' (standard versus control) to detect tuning to numerical values, and a separate Mann–Whitney *U* test with factor 'parity' (even versus odd) to exclude neurons responsive to parity judgements (both evaluated at $\alpha = 0.01$). Across all four areas individually, a substantial proportion of neurons showed a significant main effect for the factor 'number' ($P < 0.001$; binomial test with $P_{\text{chance}} = 0.01$), but no effect for the factors 'protocol' or 'parity' (Fig. 3a). Across all four areas combined, 15.1% of MTL neurons (121/801) showed an exclusive significant main effect for the factor 'number' (Fig. 3a). Each of the tested numerosities (0–9) constituted the preferred numerosity of individual selective neurons (Fig. 3b); differences in these proportions were not due to response preferences for specific numerosities, but consistent with random variation ($P = 0.15$; Mantel–Haenszel test). Similarly, response latencies across the four MTL regions (parahippocampal cortex (PHC), entorhinal cortex (EC), hippocampus (HIPP) and amygdala (AMY)) did not reveal significant differences ($P = 0.87$; Kruskal–Wallis test).

Neuronal tuning characteristics

To explore hypothesized different physiological mechanisms for the representation of small and large numerosities, we first analysed the tuning curves of number-selective neurons. We calculated the numerosity tuning functions of all numerosity-selective neurons using standardized activity (*z*-score relative to baseline activity) (Fig. 3c). Apart from the well-known number-tuning characteristics, that is, maximum activity to the preferred number and progressively reduced firing rates as distance from the preferred number increased, several distinguishing features in the tuning to small versus large numbers emerged.

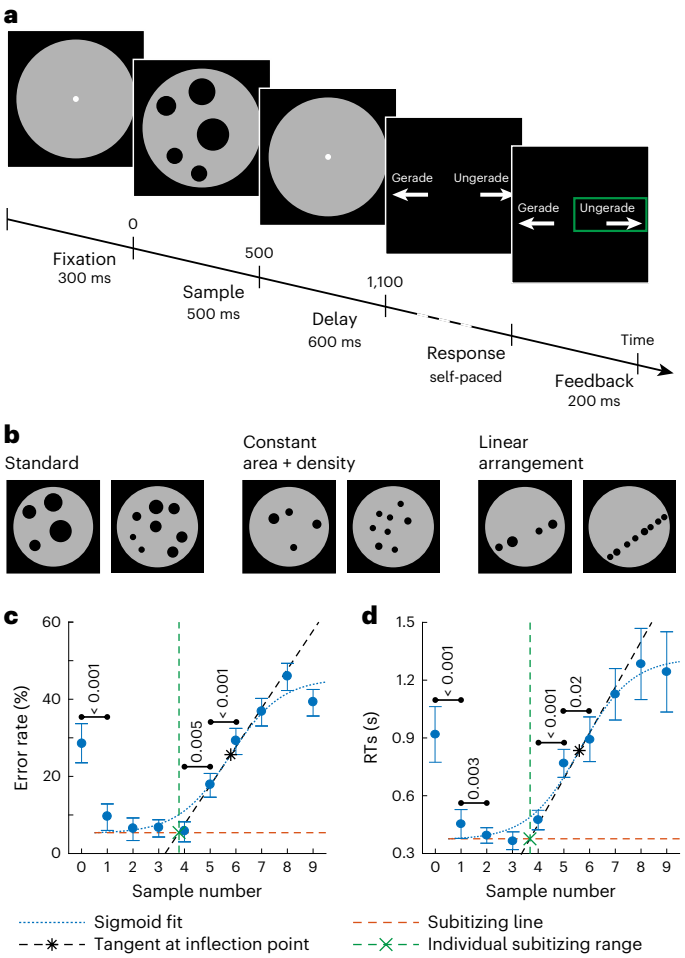

Fig. 1 | Behavioural task, stimuli and behavioural performance. a, Parity judgement task. Participants were required to indicate whether the number of dots was even ('gerade') or odd ('ungerade') by pressing the left or right arrow key, respectively (or vice versa). **b**, Sample number protocols. Dot arrays represented numerosity. They were shown in a standard layout with variable dot size and position (left), in a control layout with equalized total area and density of the dots (middle), and additionally as linearly arranged dots (right). Numerical values covered the range 0–9; exemplary dot displays for numbers 4 and 8 for each protocol are shown. **c**, Behavioural performance. Mean error rates and error bars denoting standard error of the mean (s.e.m.) are shown ($n = 17$). Values above small horizontal bars indicate *P* values for pair-wise comparisons (two-sided, paired *t*-test, Bonferroni-corrected for multiple comparisons of numbers ($n = 9$)); all other pair-wise comparisons were not different ($P > 0.05$). The subitizing boundary (green dashed line) is defined as the intersection point of the tangent (black dashed line) at the inflection point (black star) of a sigmoid fit (blue dotted line) to the error rates (excluding zero), and the subitizing line (red dashed line) at which the sigmoid curve intersects the *y* axis. **d**, RTs. Median and error bars denoting s.e.m. are shown ($n = 17$). Values above small horizontal bars indicate *P* values for pair-wise comparisons (two-sided, Wilcoxon signed rank test, Bonferroni-corrected for multiple comparisons of numbers ($n = 9$)); all other pair-wise comparisons were not different ($P > 0.05$). Conventions for the subitizing boundary as in **c**.

https://doi.org/10.1038/s41562-023-01709-3

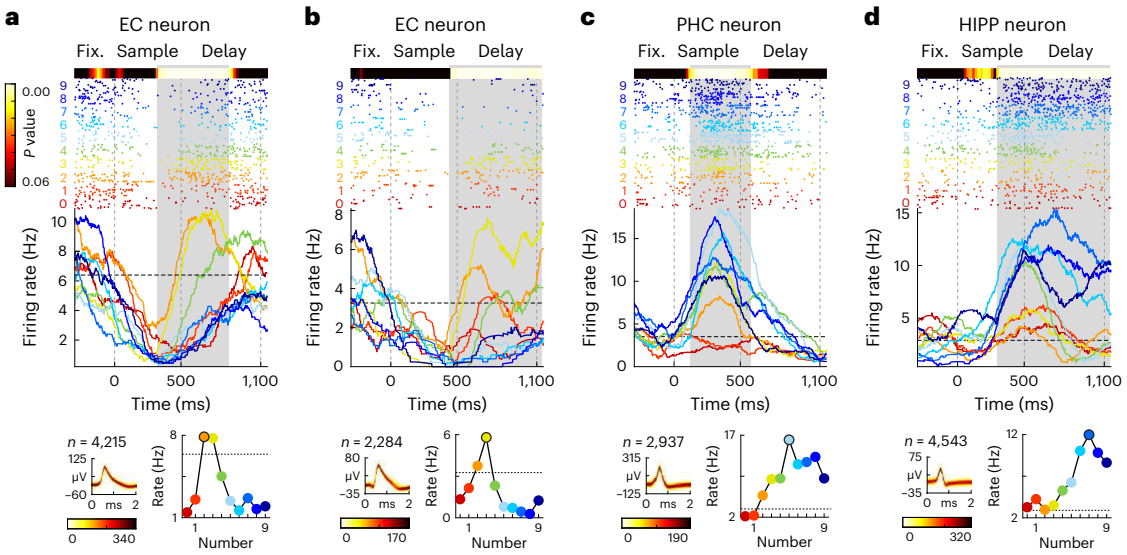

Fig. 2 | Responses of number-selective neurons. a, Example neuron from EC tuned to small number 2. Top: dot-raster histogram. Each row indicates one trial (colours correspond to presented numbers); each dot represents one action potential. Middle: corresponding mean instantaneous firing rates across trial time obtained by averaging responses to each number (smoothed using a 150 ms Gaussian kernel). Colours correspond to sample number. The horizontal dotted line depicts spontaneous activity (average across fixation periods). The grey shaded area represents the significant number-selective interval according to the sliding-window ANOVA (colour-coded *P* values above each panel). Bottom (left): density plot of the recorded action potentials, colour darkness indicating the number of overlapping wave forms according to the colour scale at the bottom. Bottom (right): number tuning function (average firing rate in the selective trial interval plotted against sample number). The horizontal dotted line indicates spontaneous firing rate. **b**, Example neuron from EC tuned to small number 3. Same layout as in **b**. **c**, Example neuron from PHC tuned to large number 5. Same layout as in **b**. **d**, Example neuron from HIPP tuned to large number 7. Same layout as in **b**.

First, tuning functions to small preferred numerosities 0–3 showed systematic surround suppression below spontaneous activity to non-preferred numbers, whereas tuning functions to large numerosities 4–9 returned to spontaneous activity for non-preferred numbers (Fig. 3c; see also example neuron tuning functions in Fig. 2a–d). We compared the firing rates elicited by non-preferred numbers (that is, at the flanks of the tuning curves) to the neurons' baseline activity. We found that firing rates at the flanks were significantly smaller than baseline activity in neurons tuned to each of the preferred numbers 0 to 3 (*P* values numerosity 0: 1.4×10^{-15}; numerosity 1: 0.0024; numerosity 2: 8.7×10^{-14}; numerosity 3: 0.0086; one-sided Wilcoxon signed rank tests), but not different from baseline in neurons tuned to each of the preferred numbers 4 to 9 (all *P* values >0.98; one-sided Wilcoxon signed rank tests) (Fig. 3d). The sharp cut in surround suppression between 3 and 4 was not due to tuning functions for preferred numbers larger than 3 becoming too wide to detect suppression. This argues for a physiological effect rather than a tuning-function resolution issue.

Second, and correlating with this tuning-flank suppression, we observed systematic differences in the amplitudes of the tuning curves. We fitted Gauss functions to the tuning curves and derived the amplitude value as a quantitative measure for the amplitude of the tuning functions[25]. Tuning curve amplitudes of neurons tuned to small numbers (0–3) were significantly smaller compared to large number (4–9) tuning curves ($P < 0.001$; one-sided Mann–Whitney *U* test), whereas tuning amplitudes were indifferent within the groups of neurons tuned to small ($P = 0.38$; Kruskal–Wallis test) and large numbers ($P = 0.69$; Kruskal–Wallis test) (Fig. 3e).

Third, tuning selectivity showed a dichotomy between small and large numbers. Since small numbers in the subitizing range can be discriminated more accurately (Fig. 1c), and more accurate discrimination is linked to more selective (that is, narrower) tuning functions[26–28], systematic differences in number tuning selectivity between the subitizing versus estimation range are expected. Thus, we derived the sigma value from the Gauss fits to quantify tuning width[29]. The tuning widths for neurons tuned to numbers 0–3 were small and not different in value ($P = 0.9$; Kruskal–Wallis test). Note that sigma as a measure of tuning width can be much smaller than 1, which is why the stable tuning widths in the subitizing range are not due to a floor effect. Around preferred number 4 or 5, a turning point emerged with tuning widths systematically increasing in a linear fashion towards larger numbers, as expected for ratio-dependent estimation (Fig. 3f). The selectivity dichotomy of neurons across the range of numbers is in agreement with behavioural predictions and suggests separate mechanisms for the coding of small versus large numerosities.

To explore the categorically distinct representation of small versus large numerosities further, we performed a representational similarity analysis (RSA) by calculating the correlation coefficients of the *z*-scored firing rates between all pairs of numbers for number-selective neurons ($n = 121$). We hypothesized that neurons tuned to small numbers would show more similar firing rates to other small numbers and thus higher correlation coefficients within pairs of small numbers, whereas neurons tuned to large numbers would show higher correlation coefficients within pairs of other large numbers. The resulting matrix of correlation coefficient values suggests radically different coding for numerosity 0 (which was therefore excluded from this analysis), but also categorical differences between small and large countable numbers (Fig. 3g). We then quantified for which of the eight number boundaries (that is, 1|2, 2|3, …, 8|9) the difference between within- and across-category correlation values was most significant and thus best segregated these data into small versus large number representations. The highest and most significant correlation value difference between within- and across-categories ($r = 0.27$; $P = 1.12 \times 10^{-5}$) was found for the boundary 3 versus 4 (Mann–Whitney *U* test, Bonferroni-corrected for multiple comparisons of boundaries, $n = 8$) (Fig. 3h). This correlation analysis suggests categorically different encoding of small versus large numbers based on the selective neurons' firing rates. When applied to the entire set of single units regardless of numerosity selectivity (801 units), this analysis yielded qualitatively similar results (Supplementary Fig. 1).

https://doi.org/10.1038/s41562-023-01709-3

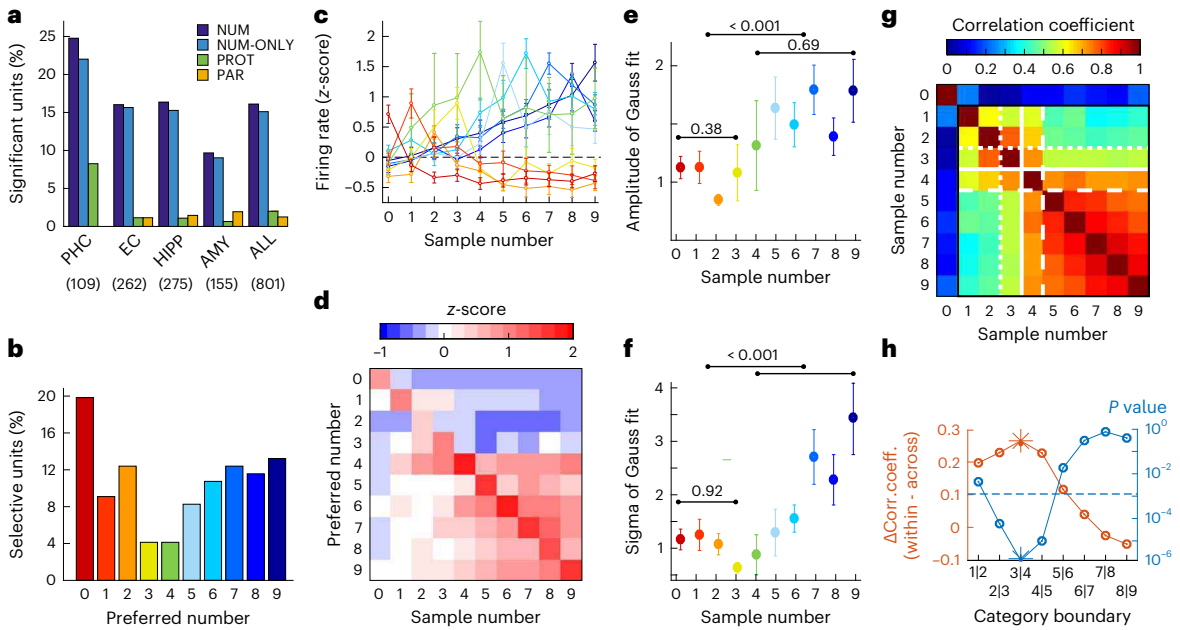

Fig. 3 | Tuning characteristics of number-selective neurons. a, Proportions of neurons with a significant main effect for 'number' (NUM), 'exclusively number' (NUM-ONLY), 'number protocol' (PROT) in a two-way ANOVA, or 'parity' (PAR) in a Mann–Whitney *U* test, evaluated at $\alpha = 0.01$, separately for each MTL region (number of recorded units in brackets). **b**, Proportion of neurons tuned to different preferred numbers. **c**, Average *z*-scored tuning curves of number-selective neurons tuned to the ten numbers (colour-coded as depicted in **b**). Error bars denote standard error of the mean (s.e.m.). **d**, Average (*z*-scored) firing rate per preferred number (rows) colour-coded relative to baseline activity. Blueish colours indicate suppression; reddish colours indicate enhancement of firing rates relative to baseline activity. **e**, Average tuning amplitude per preferred number derived from Gauss fits to tuning curves. Standard errors denote s.e.m. Amplitudes did not differ for units preferring small numerosities 0–3 ($n = 55$) and for units preferring large numerosities 4–9 ($n = 73$) (Kruskal–Wallis tests; $P > 0.05$, n.s.) but were significantly different between both groups, as indicated by the *P* value above the small horizontal bar (one-sided Mann–Whitney *U* test). **f**, Average tuning selectivity per preferred number as measured by sigma from Gauss fits to tuning curves. Error bars denote s.e.m. Sigma was small and constant for small numbers but increased in proportion with the value of large numbers. Sigmas did not differ for units preferring small numerosities 0–3 ($n = 55$), but were significantly different between both groups, as indicated by the *P* value above the small horizontal bar (one-sided Mann–Whitney *U* test). **g**, Correlation coefficients of the *z*-scored firing rates between pairs of numbers for all number-selective neurons ($n = 121$). Firing rates were more similar (higher correlations, corresponding to reddish colours) for numbers from the same number (small or large) category (upper-left and lower-right square), compared with responses for numbers from a different category (lower-left and upper-right square). White lines depict significant number category boundaries (solid line is most significant), dividing correlations into small versus large number categories. **h**, Evaluation of the goodness-of-fit of different number boundaries. Orange values depict the differences of correlation coefficients when segregating small versus large number categories (excluding zero) at different boundaries. The corresponding *P* values (two-sided Mann–Whitney *U* test) for these coefficient differences are shown in blue. Boundary 3 versus 4 (asterisks) divides the data most significantly into two number categories. The blue dotted line indicates $\alpha = 0.01$, Bonferroni-corrected for multiple comparisons ($n = 8$).

Population decoding using SVM classifiers

In addition to single neurons, neural populations carry information about neuronal computations[30]. Next, we therefore explored potential decoding discontinuities for numbers at the level of the population of selective neurons ($n = 121$) (Fig. 4a). Using a support vector machine (SVM) classifier, we first identified the time window of significant above-chance classification during a sliding-window classification analysis (60 ms to 1,200 ms after sample number onset; $\alpha = 0.01$) (Fig. 4b). Next, another classifier was trained with 50% of the data to discriminate the ten numbers, and then tested on the remaining 50% novel data from the same neuronal population and in the same time window to evaluate how well the model could decode each number based on information extracted from trials used during training. The classification probability of predicted numbers per truly presented number was then used to assemble a confusion matrix, with the main diagonal indicating correct labelling (Fig. 4c).

The classifier predominantly confused numbers from within the small-number category (upper-left square) and from within the large-number category (lower-right square), but not numbers across the small and large number category (in the lower-left and upper-right squares) (Fig. 4c). Zero was again excluded from the categorization analysis due to its distinctive difference relative to countable numbers. We quantified which of the eight number boundaries (that is, 1|2, 2|3, …, 8|9) resulted in the largest statistically significant differences in classification probabilities. A boundary between numbers 4 and 5 resulted in the largest and most significant difference in classification probability (difference 15%) between within- and across-categories ($P = 2.22 \times 10^{-6}$; Mann–Whitney *U* test, Bonferroni-corrected for multiple comparisons of boundaries, $n = 8$) (Fig. 4d). This population decoding analysis again indicates categorically different encoding of small versus large numbers, with a boundary between numbers 4 and 5. Again, this analysis yielded similar results when performed for the entire population of single units ($n = 801$; Supplementary Fig. 2).

Multi-dimensional state-space analysis and cluster analysis

Finally, to explore the dynamics of neuronal coding differences potentially pointing to two different number systems, we performed a multi-dimensional state-space analysis for the population of numerosity-selective neurons. At each point in trial time, the activity of *n* recorded neurons is defined by a point in *n*-dimensional space, with each dimension representing the activity of a single neuron ($n = 121$). The multi-dimensional space (used for quantitative analyses) is reduced to the three most informative dimensions for graphical depiction in three-dimensional (3D) state space. This results in 3D trajectories that are traversed for different neuronal states, that is, for the ten different numerical values (Fig. 5a). These trajectories reflect

https://doi.org/10.1038/s41562-023-01709-3

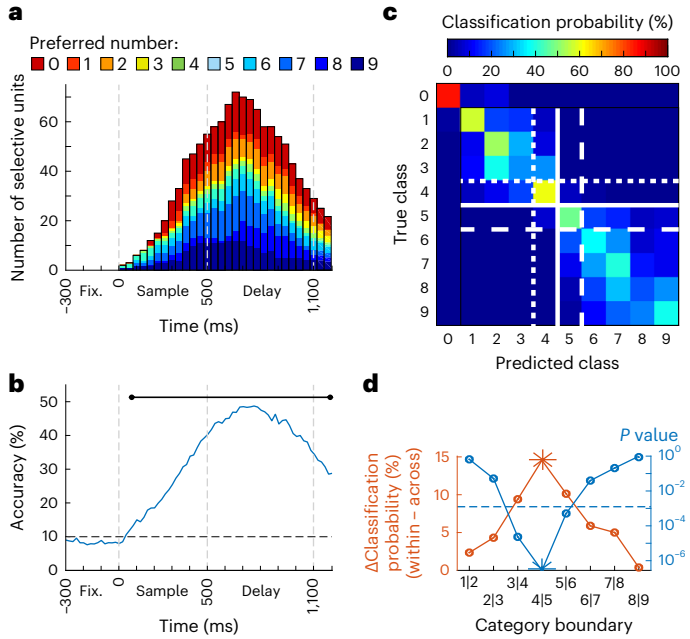

Fig. 4 | SVM classification analysis. a, Number of neurons that become selective to their respective preferred number during the course of the trial. **b**, Classification accuracy for decoding number information after training an SVM classifier on the instantaneous firing rates across the trial period for all number-selective neurons ($n = 121$). The dashed line represents chance level (10% for ten classes). The black bar above the data indicates significance ($P < 0.01$, one-sided permutation test compared with SVM trained on shuffled data). **c**, Confusion matrix derived from training an SVM classifier on firing rates averaged across the significant time window in the sliding-window analysis in **b** (60–1,200 ms). White lines depict the significant boundaries (highest significance for the solid, thick line) that divide the number range into small and large number categories. **d**, Evaluation of the goodness of different number boundaries. Orange values depict the difference in classification probabilities when segregating small versus large number categories (excluding zero) at different boundaries. The corresponding *P* values (two-sided Mann–Whitney *U* test) for these probability differences are shown in blue. Boundary 4 versus 5 (asterisks) divides the data most significantly into two number categories. The blue dotted line indicates $\alpha = 0.01$, Bonferroni-corrected for multiple comparisons ($n = 8$).

the instantaneous firing rates of the population of selective neurons as they evolve over time. Spatial closeness (that is, small distances) of the trajectories represents similarity in coding for the respective numerosities, whereas spatial disparity (that is, large distances) reflects coding dissimilarity.

The spatial layout of the trajectories evolving after sample onset until the end of the delay period again suggests two categorically different state spaces for numerosities 0–4 versus numerosities 5–9 (Fig. 5a). The trajectories representing numerosities 0–4 run in close vicinity to each other but (as expected) with increasing spatial gaps according to ordinal numerical distance. The same spatial pattern emerges within the group of trajectories representing numerosities 5–9. However, both trajectory categories are spatially segregated from each other by a large gap.

To statistically quantify this graphical grouping effect, we performed a cluster analysis on the neuronal population state space with averaged firing rates across the previously defined time window (60–1,200 ms after sample onset). The neural state space was then orthonormalized using principal component analysis. For visualization, only the first two dimensions (that is, PC1 and PC2) are shown (Fig. 5b). We first determined the optimal number of clusters for the data set by applying two measures: the Caliński–Harabasz index (also termed 'variance ratio criterion, VRC')[31], and the 'gap criterion' that determines the most dramatic decrease in error measurement (the 'elbow' or 'gap') of different cluster numbers[32]. Both measures indicated two clusters as the optimal cluster number for the dataset (Fig. 5c). We then applied unsupervised *k*-means clustering to partition all trials ($n = 160$) into two clusters[33]. The clustering algorithm detected one cluster that comprised the state spaces for numbers 0–4, and a second cluster consisting of state spaces for numbers 5–9 (Fig. 5d). Thus, the number state space is optimally described by two clusters that border between number representations 4 and 5. Again, performing this analysis for the entire population of single units ($n = 801$) yielded similar results (Supplementary Fig. 3).

Discussion

Our results provide evidence for two mechanisms encoding the continuous range of number. The number space from 0 to 9 was uninterruptedly covered by single neurons' overlapping tuning functions inherently ordered by number, and the activity of neuron populations was systematically arranged by numerical distances[16,34–36]. However, a coding dichotomy mirroring behavioural findings emerged within this representational continuum: neuronal tuning to small numbers in the subitizing range was more selective and ratio independent, whereas tuning widths increased in a ratio-dependent manner after a turning point around number 4. We also observed strong evidence for this coding dichotomy at the neuronal population level. This argues for a separate enumeration system for subitizing in addition to an estimation system[2–5]. Whether the current findings in the MTL transfer to other brain regions is currently not known and requires further exploration.

A defining feature of neuronal tuning in the subitizing range was surround suppression below baseline activity. Surround inhibition is a basic neuronal circuit operation[37,38] known to increase contrast sensitivity. Here, excitatory neurons firing in response to preferred stimuli recruit broadly tuned inhibitory interneurons that in turn suppress firing of neurons tuned to different preferred stimuli. Inhibition via interneurons is supposed to shape and sharpen the tuning to numerosities in the animal brain[39,40] and could mechanistically explain the more accurate number discrimination in the subitizing range. The time scale of surround inhibition to enable selective encoding in the subitizing range could be very fast. Moreover, the time delay of surround suppression with respect to classical receptive field excitation in the primate visual system has been reported to range from 15 to 60 ms (ref. 41), but (with a delay of 9 ms) can also act almost as suddenly as the direct-driving classical receptive field excitation signals[42]. Such short delays in surround inhibition are thought to emerge from a combination of feedforward, lateral and feedback connections to the target area[39,43]. While these mechanisms of surround suppression are a realistic assumption to explain the enhanced neuronal tuning in the subitizing range, they need direct testing in future experiments. With excitatory and inhibitory neurons identified in the human MTL[44–46], the necessary circuit components would readily be available to implement almost instantaneous surround inhibition for selective coding in the subitizing range.

Subitizing has been suggested to tap a different system in addition to that that for number estimation[5]. In contrast to number estimation, which is unaffected by attentional manipulations, subitizing is heavily dependent on attentive resources[47–49]. Attention-based processes that determine how many elements of information can be kept active in working memory have a very limited capacity of up to around four items[50,51], precisely the same set-size limit found for subitizing[52]. The mechanisms we discovered for subitizing may therefore well play a role for other capacity-limited processes, such as attention and working memory[49]. Similar to the observed surround suppression in small-number tuning curves, tuning flank suppression is a known mechanism to contrast task-relevant and task-irrelevant stimulus features in attention- and working-memory-related operations[53–55]. Here as well, a suppressive zone below baseline is seen in the surround of the

https://doi.org/10.1038/s41562-023-01709-3

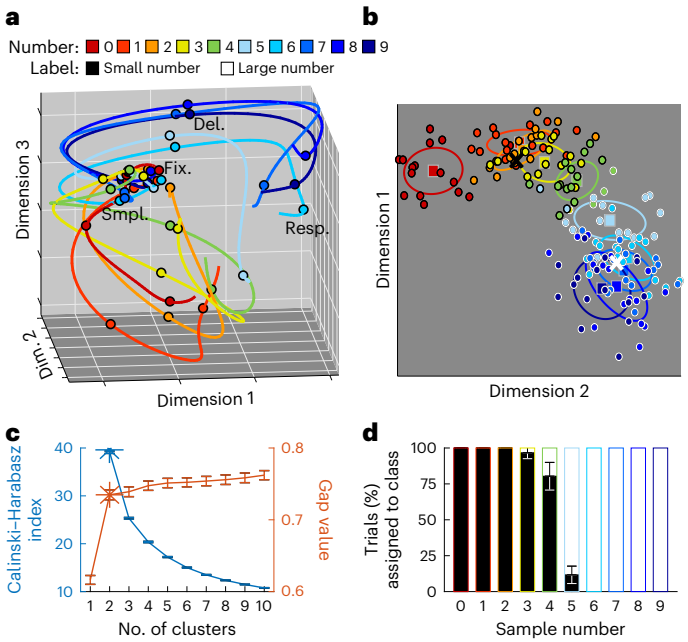

Fig. 5 | Population state-space analysis and *k*-means clustering. a, Averaged state-space trajectories of number-selective neurons for all number conditions, reduced to the three principal dimensions for visualization. Each trajectory depicts the temporal evolution in the time window −300 to 1,200 ms (stimulus onset to 100 ms after delay offset). The state space shows a gap between trajectories for numbers 0–4 versus 5–9. Circles indicate boundaries between task phases. Fix, Fixation; Smpl., Sample; Del., Delay; Resp., Response. **b**, Neural states, reduced to the two principal dimensions, after averaging firing rates per trial across the significant time window in the SVM classification analysis (60–1,200 ms). Different colours correspond to different number conditions. Each dot represents one trial; squares and ellipses indicate condition mean and covariance ellipse per condition. The colours of the dot outlines (black for 0–4 or white for 5–9) indicate the class label assigned by the *k*-means classifier. The black and grey crosses show the centroids of each class. **c**, Evaluation of different numbers of clusters using the Caliński–Harabasz criterion (blue) and the gap criterion (orange). Data are presented as mean values, error bars denote standard deviation (s.d.) of cross-validations ($n = 50$). Asterisks indicate the optimal number of clusters. Note that, unlike the Caliński–Harabasz criterion, the gap criterion would also be defined for clustering solutions containing only one cluster. **d**, Proportion of trials per number condition that were labelled as belonging to class 'small numbers' (black) or class 'large numbers' (white). Data are presented as mean values; error bars denote s.d. of cross-validations.

preferred stimulus parameter[56–59]. In intermediate-level visual area V4, surround suppression caused by spatial attention can be very quick and as early as 75 ms post stimulus onset[58]. The hypothesis is that with high attentional demand, the subitizing system overrides the estimation system. Thus, the subitizing system would enhance rather than replace estimation for small numbers[5]. This would also explain why estimation processes can in principle work also for small numbers, as seen many times in both human and animal brain studies[7,8,12,60–63]. This hypothesis is consistent with our current findings but needs to be tested empirically by contrasting the responses of neurons with and without attentional demands assigned to number representations. Complementing our parity judgement task with richer and more explicit number tasks could also help to support the generality of findings.

Methods

Experimental model and participant details

All studies conformed to the guidelines of the Medical Institutional Review Board at the University of Bonn, Germany, and were approved by this Board (licence no. 146/19). Seventeen human participants (five male, mean age 37.6 years) with medically refractory focal epilepsy undergoing invasive pre-surgical assessment participated in the study. Informed written consent was obtained from each patient; participants received no financial compensation for participating in the study.

Neurophysiological recording

Participants were implanted bilaterally with chronic intracerebral depth electrodes in the MTL to localize the seizure-onset zone for possible neurosurgical resection. The implantation site of the electrodes was determined exclusively by clinical criteria and varied across participants. To record neuronal signals, we used 9–10 clinical Behnke–Fried depth electrodes (AD-Tech Medical Instrument Corp.). Each depth electrode contained a bundle of nine platinum–iridium micro-wires protruding ~4 mm from the tip of each electrode: eight high-impedance active recording channels, and one low-impedance reference wire. Using a 256-channel ATLAS neurophysiology system (Neuralynx), differential neuronal signals (recording range ±3,200 μV) were filtered (bandwidth 0.1–9,000 Hz), amplified and digitized (sampling rate 32,768 Hz). Recorded spikes and behavioural data were synchronized via 64-bit timestamps using the Cheetah software (Neuralynx).

After bandpass-filtering (bandwidth 300–3,000 Hz) the local field potentials, spikes were automatically detected and pre-sorted using the Combinato package[64]. Classification as artefact, multi-unit or single unit was verified manually on the basis of spike shape and its variance, inter-spike-interval distribution per cluster, and the presence of a plausible refractory period. Only units that responded with an average firing rate of >1 Hz during stimulus presentation were included in the analyses. Across 28 recording sessions from all 17 participants, a total of 801 single units were identified in the PHC (109 units), EC (262 units), HIPP (275 units) and AMY (155 units).

Stimuli

All stimuli were presented within a filled grey circle (diameter approximately 6° of visual angle) on a black background. During fixation and delay phase, a white fixation spot was presented in the centre of the grey area. It was removed during stimulus presentation to avoid confusion with non-symbolic stimuli.

Numerical values of the stimuli ranged from 0 to 9 and consisted of black sets of dots with the number of dots corresponding to the respective numerical value ('numerosities'). Given that we needed zero for a balanced count of even and odd numbers, and acknowledging that zero is a set (even if empty) and a whole number like the natural numbers, we included zero in the stimulus presentation. We used different 'protocols' to control for low-level visual features. For the standard protocol, diameter and location of each dot varied randomly within a given range (diameters of 0.3° to 0.8° of visual angle). In the control displays, the total dot area and dot density (mean distances between centres of the dots) across numerosities was equated. Additionally, in half of the control trials, the dots were linearly arranged. Standard and control protocols for the non-symbolic stimuli were shown with equal probability of 50%.

Experimental task

Participants performed a parity judgement task sitting in bed and facing a laptop (display diagonal 11.7 inches, resolution 1,366 × 768 px) on which stimuli were presented at a distance of approximately 50 cm. Participants were not informed about hypotheses or purposes of the experiment to exclude any bias.

Before the experiment, the task instruction was displayed on the screen in addition to verbal explanation by the experimenter, specifying which numbers were 'even' and which ones 'odd'. Furthermore, to reduce confusion about the 'zero' stimulus, we added some familiarization trials preceding the recordings, during which the experimenter pointed out, once more, that an empty grey circle represented the 'even number zero'.

Each trial started with a fixation period of 300 ms. Afterwards, a number stimulus was presented for 500 ms, followed by a 600 ms delay display. After delay offset, participants had to decide whether the number had been even or odd by pressing the left or right arrow key on the keyboard, respectively, as indicated on the response screen ('gerade' (even) or 'ungerade' (odd)). To control for potential motor bias, the keys associated with the respective response were balanced and switched across blocks. Participants responded in a self-paced manner, but were asked to respond as fast and accurately as possible. After a 200 ms feedback display, the next trial started automatically. Each number stimulus was presented 16 times, resulting in 160 trials. A session was divided into four blocks, comprising all conditions in pseudo-random order. Stimuli and experimental task were programmed in MATLAB R2017a (The MathWorks), using the Psychtoolbox3 (refs. 65–67).

Behavioural analyses

First, we plotted the behavioural measures (error rates and RTs, averaged across participants, $n = 17$) as a function of numerical value of the stimulus. This function is characterized by a shallow, near-zero slope for small numbers, and a steeper slope for numbers beyond the subitizing range. The discontinuity point, in which the slope of this function changes, defines the upper boundary of the subitizing range. To quantify this boundary, we applied the algorithm for calculating individual subitizing ranges[18], that is, we first fitted a sigmoid (logistic) function to the behavioural data:

$$\mathrm{BM} = L + (U - L) \cdot \frac{1}{1 + \exp(-x - \mathrm{IP})}$$

where BM is the behaviour exhibited in response to the presentation of numerical value x. The model coefficients lower bound L, upper bound U and inflection point IP were estimated in the fitting process. We then applied the Levenberg–Marquardt algorithm to solve this non-linear least-squares curve-fitting problem. Next, we fitted two linear functions to the sigmoid curve. The subitizing line is equivalent to the lower bound L where the sigmoid curve crosses the y axis; the tangent line is fitted to the tangent at the inflection point IP. The intersection point of these two linear fits is then used as a proxy for the upper boundary of the subitizing range.

Neuronal analyses

Overall behavioural performance was high across all participants (mean ± standard deviation: 86.4 ± 3.1%). Errors occurred mainly for larger numbers. Because of the low error rate and the need to have balanced numbers of trials across numerosities, we included both correct and incorrect trials into the analyses.

Tuning characteristics

Spike trains were smoothed trial-wise (Gaussian kernel with $\sigma = 150$ ms) for each unit within the trial window −300 to 1,200 ms (fixation onset to 100 ms after delay offset). At every 20 ms step, instantaneous firing rates were subjected to a two-factor ANOVA with factors 'numerical value' (0–9) and 'protocol' (standard versus control) to detect tuning to numerical values, and a separate Mann–Whitney U test with factor 'parity' (even versus odd) to exclude neurons responsive to parity judgements (note that we could not apply a three-factor ANOVA as parity is not independent from the numerical value), resulting in a temporal sequence of P values for each of the three factors. A cluster permutation test[68] was then performed to identify time intervals of significant number encoding, thereby controlling for multiple comparisons across time ($\alpha_{clus} = 0.01$; $P_{rank} < 1\%$; $n_{perm} = 100$). A unit was termed 'exclusively number-selective' (NUM-ONLY) if a significant time interval for the factor 'numerical value' was observed between 0 ms and 1,000 ms (stimulus onset to 100 ms before delay offset), and there were no overlapping significant intervals for the factors 'protocol' or 'parity'. These units are henceforth referred to as 'number neurons'. Proportions of these number neurons were determined for each MTL region and subjected to a binomial test ($P_{chance} = 0.01$), Bonferroni-corrected for multiple comparisons across brain regions ($n = 4$), to evaluate whether the observed proportions were higher than expected by chance.

For each number neuron, we calculated tuning functions by averaging the firing rates during the significant time interval across trials for all numerical values. The numerical value eliciting the maximum response was defined as 'preferred numerosity'. To test for potential differences in the proportions of preferred numerosities, we applied the Mantel–Haenszel X^2 test[69,70], a generalized version of Pearson's X^2 test, for analysis of $2 \times 9 \times 17$ contingency tables, excluding zero as an outlier and stratified for different participants ($n = 17$).

Tuning functions were then standardized by z-scoring, that is, we subtracted the mean baseline activity elicited during the fixation periods (−300 to 0 ms) from all values and divided the difference by the standard deviation. In cases where multiple significant number intervals were identified within the same unit, we calculated separate tuning curves for each of these intervals (8/121 number neurons). Population tuning functions were then obtained by averaging across all units that preferred the same number.

To quantify surround suppression, we combined the firing rates to all non-preferred numbers for all units preferring the same numerical value and tested whether they were significantly smaller than spontaneous activity (that is, a z-score of 0) using a one-sided Wilcoxon signed rank test.

To estimate the tuning amplitude and width of all numerosity-selective neurons, we fitted a Gauss function, representing the standard symmetric distribution, to each individual tuning curve:

$$\mathrm{FR}(x) = a \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) + o$$

where FR is the z-scored firing rate elicited in response to the presentation of numerical value x. The mean μ was fixed to the preferred number; the model coefficients amplitude a, offset o, and standard deviation σ were estimated in the fitting process, thereby using the following bound constraints: $a = [0;\max(\mathrm{FR})]$, $o = [\min(\mathrm{FR});\max(\mathrm{FR})]$ and $\sigma = [0;\mathrm{Inf}]$, to avoid implausible fitting results. We then applied the Levenberg–Marquardt algorithm to solve this non-linear least-squares curve-fitting problem.

RSA

Pearson's correlation coefficient quantifies the strength of the linear relationship between two variables. To evaluate firing rate differences between different number conditions, we performed an RSA. We calculated a correlation matrix, showing the correlations between firing rates in response to number i and to number j, respectively, for all condition pairs, based on the z-scored tuning curves of all number neurons. We then determined the boundary that divided the data best into the categories of 'small' and 'large' numbers. For this, we defined eight potential boundaries (1|2, 2|3, …, 8|9). Due to obvious dissimilarity, zero was excluded from this categorization analysis. A boundary divided the correlation matrix into four squares. Correlation coefficients in the upper-left (within small-number category) and lower-right squares (within large-number category) of the matrix were then iteratively compared with the coefficients in the remaining upper-right and lower-left (across-category) matrix squares for different number boundaries. For each boundary, the difference between within-category and across-category elements was then quantified using a two-sided Mann–Whitney U test ($\alpha = 0.01$, Bonferroni-corrected for multiple comparisons of boundaries, $n = 8$). Note that the main diagonal was excluded as it reflects the correlation of each stimulus with itself, and that the correlation matrix is symmetric. Thus, only

values from the upper triangular portion of the correlation matrix were considered. For each boundary, we assessed the *P* value of the statistical test, and the average difference between within-category and across-category elements to evaluate how well the boundary divided the data into two categories. Pairs of correlation coefficients are not statistically independent due to their bivariate nature, which might bias the results of the Mann–Whitney *U* test. To account for this, we performed an additional permutation test by randomly shuffling the within-category and across-category labels 10,000 times and comparing the test statistics of the random data to the true ones, using again a two-sided Mann–Whitney *U* test[71].

SVM classification

For each unit, data were divided into ten classes according to the numerical value of the sample stimulus (16 trials per class) and spike trains per trial were smoothed (Gaussian kernel with $\sigma = 150$ ms) within the analysis window −300 to 1,200 ms (fixation onset to 100 ms after delay offset). A default multi-class SVM classifier[53] was then trained and tested on the instantaneous firing rates at every 20 ms step[72]. We applied Monte-Carlo cross-validation, that is, we created multiple splits of our dataset ($n_{repetitions} = 100$) by randomly sampling 50% of the trials as training set, balancing conditions within each split. The remaining 50% of the trials were used as test set. Thus, each training and test set comprised 80 trials. For each split, we standardized all firing rates by *z*-scoring (mean and standard deviation obtained from training data only), fitted the classifier to the training data, and assessed predictive accuracy by counting the instances for which a certain activity pattern of the test data was labelled correctly. The results were then averaged over all splits. To identify temporal clusters during which accuracy differed significantly from chance level (10% for ten classes), the analysis was repeated with randomly shuffled trial labels ($n_{perm} = 100$), and a cluster permutation test[63] was performed. In short, we identified temporal clusters of significant values by comparing the true accuracy values against the distribution of random ones ($\alpha_{clus} = 0.01$). The significance of these 'candidate clusters' was then evaluated by comparison with the clusters of the random data ($P_{rank} < 1\,\%$). Next, an SVM (with the same settings as above) was trained and tested on the firing rates obtained by averaging across the significant time window (60–1,200 ms). We assembled a confusion matrix, which counted the frequency at which a trial of a certain class was assigned different labels by the classifier.

Again, we analysed which boundary divided the data best into the categories of 'small' versus 'large' numbers. As before, we defined eight potential boundaries (1|2 to 8|9; excluding zero). Classification probabilities in the upper-left (within small-number category) and lower-right squares of the matrix (within large-number category) were iteratively compared with the classification probabilities in the remaining (across-category) matrix squares (lower left and upper right) for all number boundaries. The difference between both groups was then quantified using a non-parametric Mann–Whitney *U* test ($\alpha = 0.01$; Bonferroni-corrected for multiple comparisons of boundaries, $n = 8$). Note that the main diagonal was excluded as it reflects correct classifications, and that the confusion matrix is not symmetric (unlike the correlation matrix). For each boundary, we assessed the *P* value of the statistical test and the average difference between within-category and across-category elements to evaluate how well the boundary divided the data into two categories. The results of the Mann–Whitney *U* test were again verified using a permutation test.

Multi-dimensional state-space analysis

To analyse neural activity of a neuronal population, we defined an *n*-dimensional space, where each axis represents the instantaneous firing rate of a number-selective neuron. At any given time, the population activity is then characterized by a single point in this space, resulting in a neural trajectory as the activity evolves over time. In other words, we calculated the trajectories for the ten different numerosities in a 121-dimensional space after averaging across conditions and smoothing (Gaussian kernel with $\sigma = 150$ ms) spike trains for each number neuron. A Gaussian-process factor analysis model was then applied[73], and the resulting neural trajectories were orthonormalized to order the dimensions by the amount of data covariance they explain. For visualization, only the top three dimensions (in terms of covariance explained) were considered.

Next, the neural population state was calculated by averaging firing rates across the significant time window (60–1,200 ms) for all trials. The neural state space was then orthonormalized using principal component analysis. For visualization, only the top two dimensions were displayed. We then applied unsupervised *k*-means clustering to partition all trials ($n = 160$) into two clusters[33]. In short, *k*-means clustering iteratively partitions the data into *k* distinct non-overlapping clusters such that the distance between all elements of the cluster and every cluster's centroid is minimized. We used the squared Euclidean distance as a distance metric, that is, centroids are the arithmetic mean of the elements in that cluster and repeated the algorithm 50 times with new randomly chosen initial cluster centroid positions.

We applied two criteria to evaluate the optimal number of clusters for our data. First, we calculated the Caliński–Harabasz index[31], also called VRC, which is defined as the ratio between overall between-cluster variance and overall within-cluster variance. Maximizing the VRC value with respect to *k* classes yields the optimal number of classes. As a second criterion, we calculated the gap value[32]. It formalizes the heuristic 'elbow method', according to which the optimal number of clusters can be found by locating the most dramatic decrease in error measurement (the 'elbow' or 'gap'). Note that, unlike the Caliński–Harabasz criterion, the gap criterion is also defined for clustering solutions containing only one cluster. For cross-validation, the *k*-means clustering analysis was repeated 50 times, using only 75% randomly selected trials per condition for each cross-validation run.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data associated with this study are publicly available at https://github.com/EstherKutter/Distinct-Neuronal-Representation-Of-Small-And-Large-Numbers-In-The-Human-MTL.

Code availability

The custom code associated with this study is publicly available at https://github.com/EstherKutter/Distinct-Neuronal-Representation-Of-Small-And-Large-Numbers-In-The-Human-MTL.

References

1. Jevons, W. S. The power of numerical discrimination. *Nature* **3**, 281–282 (1871).
2. Kaufman, E. L., Lord, M. W., Reese, T. W. & Volkmann, J. The discrimination of visual number. *Am. J. Psychol.* **62**, 498–525 (1949).
3. Mandler, G. & Shebo, B. J. Subitizing: an analysis of its component processes. *J. Exp. Psychol. Gen.* **111**, 1–22 (1982).
4. Feigenson, L., Dehaene, S. & Spelke, E. Core systems of number. *Trends Cogn. Sci.* **8**, 307–314 (2004).
5. Anobile, G., Cicchini, G. M. & Burr, D. C. Number as a primary perceptual attribute: a review. *Perception* **45**, 5–31 (2016).
6. Cheyette, S. J. & Piantadosi, S. T. A unified account of numerosity perception. *Nat. Hum. Behav.* **4**, 1265–1272 (2020).
7. Tsouli, A. et al. The role of neural tuning in quantity perception. *Trends Cogn. Sci.* **26**, 11–24 (2022).
8. Piazza, M., Mechelli, A., Butterworth, B. & Price, C. J. Are subitizing and counting implemented as separate or functionally overlapping processes? *NeuroImage* **15**, 435–446 (2002).

9. Libertus, M. E., Woldorff, M. G. & Brannon, E. M. Electrophysiological evidence for notation independence in numerical processing. *Behav. Brain Funct.* **3**, 1 (2007).
10. Harvey, B. M., Klein, B. P., Petridou, N. & Dumoulin, S. O. Topographic representation of numerosity in the human parietal cortex. *Science* **341**, 1123–1126 (2013).
11. Fornaciai, M. & Park, J. Decoding of electroencephalogram signals shows no evidence of a neural signature for subitizing in sequential numerosity. *J. Cogn. Neurosci.* **33**, 1535–1548 (2021).
12. Cai, Y. et al. Topographic numerosity maps cover subitizing and estimation ranges. *Nat. Commun.* **12**, 3374 (2021).
13. Sathian, K. et al. Neural evidence linking visual object enumeration and attention. *J. Cogn. Neurosci.* **11**, 36–51 (1999).
14. Fink, G. R. et al. Deriving numerosity and shape from identical visual displays. *NeuroImage* **13**, 46–55 (2001).
15. Hyde, D. C. & Spelke, E. S. All numbers are not equal: an electrophysiological investigation of small and large number representations. *J. Cogn. Neurosci.* **21**, 1039–1053 (2009).
16. Kutter, E. F., Bostroem, J., Elger, C. E., Mormann, F. & Nieder, A. Single neurons in the human brain encode numbers. *Neuron* **100**, 753–761.e4 (2018).
17. Kutter, E. F., Boström, J., Elger, C. E., Nieder, A. & Mormann, F. Neuronal codes for arithmetic rule processing in the human brain. *Curr. Biol.* **32**, 1275–1284.e4 (2022).
18. Leibovich-Raveh, T., Lewis, D. J., Kadhim, S. A. R. & Ansari, D. A new method for calculating individual subitizing ranges. *J. Numer. Cogn.* **4**, 429–447 (2018).
19. Atkinson, J., Campbell, F. W. & Francis, M. R. The magic number 4±0: a new look at visual numerosity judgements. *Perception* **5**, 327–334 (1976).
20. Simon, T. J. & Vaishnavi, S. Subitizing and counting depend on different attentional mechanisms: evidence from visual enumeration in afterimages. *Percept. Psychophys.* **58**, 915–926 (1996).
21. Sengupta, R., Surampudi, B. R. & Melcher, D. A visual sense of number emerges from the dynamics of a recurrent on-center off-surround neural network. *Brain Res.* **1582**, 114–124 (2014).
22. Sengupta, R., Bapiraju, S. & Melcher, D. Big and small numbers: empirical support for a single, flexible mechanism for numerosity perception. *Atten. Percept. Psychophys.* **79**, 253–266 (2017).
23. Fias, W. Two routes for the processing of verbal numbers: evidence from the SNARC effect. *Psychol. Res.* **65**, 250–259 (2001).
24. Nuerk, H.-C., Iversen, W. & Willmes, K. Notational modulation of the SNARC and the MARC (linguistic markedness of response codes) effect. *Q. J. Exp. Psychol. A* **57**, 835–863 (2004).
25. Nieder, A. Representing something out of nothing: the dawning of zero. *Trends Cogn. Sci.* **20**, 830–842 (2016).
26. Schoups, A., Vogels, R., Qian, N. & Orban, G. Practising orientation identification improves orientation coding in V1 neurons. *Nature* **412**, 549–553 (2001).
27. Yang, T. & Maunsell, J. H. R. The effect of perceptual learning on neuronal responses in monkey visual area V4. *J. Neurosci.* **24**, 1617–1626 (2004).
28. Lee, S. H. et al. Activation of specific interneurons improves V1 feature selectivity and visual perception. *Nature* **488**, 379–383 (2012).
29. Nieder, A. & Miller, E. K. Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. *Neuron* **37**, 149–157 (2003).
30. Saxena, S. & Cunningham, J. P. Towards the neural population doctrine. *Curr. Opin. Neurobiol.* **55**, 103–111 (2019).
31. Caliński, T. & Harabasz, J. A dendrite method for cluster analysis. *Commun. Stat.* **3**, 1–27 (1974).
32. Tibshirani, R., Walther, G. & Hastie, T. Estimating the number of clusters in a data set via the gap statistic. *J. R. Stat. Soc. Ser. B* **63**, 411–423 (2001).
33. Lloyd, S. Least squares quantization in PCM. *IEEE Trans. Inf. Theory* **28**, 129–137 (1982).
34. Viswanathan, P. & Nieder, A. Differential impact of behavioral relevance on quantity coding in primate frontal and parietal neurons. *Curr. Biol.* **25**, 1259–1269 (2015).
35. Ramirez-Cardenas, A., Moskaleva, M. & Nieder, A. Neuronal representation of numerosity zero in the primate parieto-frontal number network. *Curr. Biol.* **26**, 1285–1294 (2016).
36. Viswanathan, P. & Nieder, A. Spatial neuronal integration supports a global representation of visual numerosity in primate association cortices. *J. Cogn. Neurosci.* **32**, 1184–1197 (2020).
37. Hartline, H. K., Wagner, H. G. & Ratliff, F. Inhibition in the eye of Limulus. *J. Gen. Physiol.* **39**, 651–673 (1956).
38. Isaacson, J. S. & Scanziani, M. How inhibition shapes cortical activity. *Neuron* **72**, 231–243 (2011).
39. Diester, I. & Nieder, A. Complementary contributions of prefrontal neuron classes in abstract numerical categorization. *J. Neurosci.* **28**, 7737–7747 (2008).
40. Ditz, H. M., Fechner, J. & Nieder, A. Cell-type specific pallial circuits shape categorical tuning responses in the crow telencephalon. *Commun. Biol.* **5**, 269 (2022).
41. Angelucci, A. & Bressloff, P. C. Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. *Prog. Brain Res.* **154**, 93–120 (2006).
42. Bair, W., Cavanaugh, J. R. & Movshon, J. A. Time course and time–distance relationships for surround suppression in macaque V1 neurons. *J. Neurosci.* **23**, 7690–7701 (2003).
43. Nassi, J. J., Lomber, S. G. & Born, R. T. Corticocortical feedback contributes to surround suppression in V1 of the alert primate. *J. Neurosci.* **33**, 8504–8517 (2013).
44. Ison, M. J. et al. Selectivity of pyramidal cells and interneurons in the human medial temporal lobe. *J. Neurophysiol.* **106**, 1713–1721 (2011).
45. Gast, H. et al. Burst firing of single neurons in the human medial temporal lobe changes before epileptic seizures. *Clin. Neurophysiol.* **127**, 3329–3334 (2016).
46. Mosher, C. P. et al. Cellular classes in the human brain revealed in vivo by heartbeat-related modulation of the extracellular action potential waveform. *Cell Rep.* **30**, 3536–3551.e6 (2020).
47. Railo, H., Koivisto, M., Revonsuo, A. & Hannula, M. M. The role of attention in subitizing. *Cognition* **107**, 82–104 (2008).
48. Vetter, P., Butterworth, B. & Bahrami, B. Modulating attentional load affects numerosity estimation: evidence against a pre-attentive subitizing mechanism. *PLoS ONE* **3**, e3269 (2008).
49. Burr, D. C., Turi, M. & Anobile, G. Subitizing but not estimation of numerosity requires attentional resources. *J. Vis.* **10**, 20 (2010).
50. Luck, S. J. & Vogel, E. K. The capacity of visual working memory for features and conjunctions. *Nature* **390**, 279–281 (1997).
51. Cowan, N. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* **24**, 87–114 (2001).
52. Piazza, M., Fumarola, A., Chinello, A. & Melcher, D. Subitizing reflects visuo-spatial object individuation capacity. *Cognition* **121**, 147–153 (2011).
53. Wang, X.-J., Tegnér, J., Constantinidis, C. & Goldman-Rakic, P. S. Division of labor among distinct subtypes of inhibitory neurons in a cortical microcircuit of working memory. *Proc. Natl Acad. Sci. USA* **101**, 1368–1373 (2004).
54. Störmer, V. S. & Alvarez, G. A. Feature-based attention elicits surround suppression in feature space. *Curr. Biol.* **24**, 1985–1988 (2014).
55. Kiyonaga, A. & Egner, T. Center-surround inhibition in working memory. *Curr. Biol.* **26**, 64–68 (2016).

56. Müller, N. G. & Kleinschmidt, A. The attentional 'spotlight's' penumbra: center-surround modulation in striate cortex. *NeuroReport* **15**, 977–980 (2004).
57. Hopf, J.-M. et al. Direct neurophysiological evidence for spatial suppression surrounding the focus of attention in vision. *Proc. Natl Acad. Sci. USA* **103**, 1053–1058 (2006).
58. Sundberg, K. A., Mitchell, J. F. & Reynolds, J. H. Spatial attention modulates center-surround interactions in macaque visual area v4. *Neuron* **61**, 952–963 (2009).
59. Anton-Erxleben, K., Stephan, V. M. & Treue, S. Attention reshapes center-surround receptive field structure in macaque cortical area MT. *Cereb. Cortex* **19**, 2466–2478 (2009).
60. Nieder, A., Freedman, D. J. & Miller, E. K. Representation of the quantity of visual items in the primate prefrontal cortex. *Science* **297**, 1708–1711 (2002).
61. Nieder, A. & Merten, K. A labeled-line code for small and large numerosities in the monkey prefrontal cortex. *J. Neurosci.* **27**, 5986–5993 (2007).
62. Ditz, H. M. & Nieder, A. Neurons selective to the number of visual items in the corvid songbird endbrain. *Proc. Natl Acad. Sci. USA* **112**, 7827–7832 (2015).
63. Ditz, H. M. & Nieder, A. Format-dependent and format-independent representation of sequential and simultaneous numerosity in the crow endbrain. *Nat. Commun.* **11**, 686 (2020).
64. Niediek, J., Boström, J., Elger, C. E. & Mormann, F. Reliable analysis of single-unit recordings from the human brain under noisy conditions: tracking neurons over hours. *PLoS ONE* **11**, e0166598 (2016).
65. Brainard, D. H. The Psychophysics Toolbox. *Spat. Vis.* **10**, 433–436 (1997).
66. Pelli, D. G. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* **10**, 437–442 (1997).
67. Kleiner, M., Brainard, D. & Pelli, D. What's new in Psychtoolbox-3? *Perception* **36**, 14 (2007).
68. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).
69. Mantel, N. & Haenszel, W. Statistical aspects of the analysis of data from retrospective studies of disease. *J. Natl Cancer Inst.* **22**, 719–748 (1959).
70. Somes, G. W. The generalized Mantel–Haenszel statistic. *Am. Stat.* **40**, 106–108 (1986).
71. Mormann, F. et al. Neurons in the human amygdala encode face identity, but not gaze direction. *Nat. Neurosci.* **18**, 1568–1570 (2015).
72. Chang, C.-C. & Lin, C.-J. LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **102**, 1–27 (2011).
73. Yu, B. M. et al. Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. *J. Neurophysiol.* **102**, 614–635 (2009).

Acknowledgements

We thank all patients for their participation. This research was supported by the German Research Council (Mo 930/4-2, SPP 2205, SPP 2411, SFB 1089; Ni 618/11-1, SPP 2205), the BMBF (031L0197B) and a NRW Network Grant (iBehave). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

A.N. and F.M. designed the study; R.S. and F.M. recruited patients; V.B. and F.M. implanted the electrodes; E.F.K. and G.D. collected the data; E.F.K. and A.N. analysed the data with contributions from F.M.; A.N., E.F.K. and F.M. wrote the paper. All authors discussed the results and commented on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-023-01709-3.

Correspondence and requests for materials should be addressed to Florian Mormann or Andreas Nieder.

Peer review information *Nature Human Behaviour* thanks Rakesh Sengupta and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

nature portfolio

Corresponding author(s): Andreas Nieder

Last updated by author(s): Aug 18, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Depth electrodes were furnished with bundles of nine microwires each (eight high-impedance recording electrodes, one low-impedance reference, AdTech, Racine, WI) protruding ~4mm from the electrode tips. Recording system: Neuralynx ATLAS system (Bozeman, MT), Neuralynx Cheetah software (Bozeman, MT) Spike-sorting: Combinato package, Version 1 (Niediek et al., PLOS ONE 11 (12): e0166598. 2016) (https://github.com/jniediek/combinato) Stimulus presentation: Psychtoolbox3 (www.psychtoolbox.org)
Data analysis	Matlab R2017a (The MathWorks Inc., Natick, MA)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Data and analysis scripts to reproduce the main results can be found in a GitHub repository (https://github.com/EstherKutter/Distinct-Neuronal-Representation-Of-Small-And-Large-Numbers-In-The-Human-MTL)

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Both sexes participated randomly in the study (five males, twelve females).
Population characteristics	All participants were right-handed. No population characteristics were included as covariates in the current study.
Recruitment	The current sample is an opportunity sample, based on patient influx for epilepsy treatment. Patients were included in the study if they consented to microwire implantation.
Ethics oversight	Informed written consent was obtained from each patient. All studies conformed to the guidelines of the Ethik-Kommission, Mediziniische Fakultät, Rheinische Friedrich-Wilhelms-Universität Bonn, i.e. in English: the Medical Institutional Review Board at the University of Bonn, Germany (license Nr. 146/19).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Seventeen human subjects (five males, twelve females, mean age 37.6 years) with medically refractory focal epilepsy undergoing invasive presurgical assessment participated in the study. The sample is representative of patients with intractable epilepsy. No a-priori sample-size calculation was performed. The number of recording sessions comfortably complies with or exceeds current standards in human microwire recordings.
Data exclusions	no data were excluded
Replication	not applicable due to the constraints of the intervention
Randomization	there was no allocation to experimental groups
Blinding	blinding was not possible due to the constraints of the intervention

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging