

1 Creating something out of nothing: Symbolic and non-symbolic  
2 representations of numerical zero in the human brain

3

4 Benjy Barnett<sup>1,2\*</sup>, Stephen M. Fleming<sup>1,2,3</sup>

5

6 <sup>1</sup> *Wellcome Centre for Human Neuroimaging, University College London, 12 Queen Square,*  
7 *London WC1N 3AR, UK.*

8 <sup>2</sup> *Department of Experimental Psychology, University College London, 26 Bedford Way, London*  
9 *WC1H 0AP, UK.*

10 <sup>3</sup> *Max Planck UCL Centre for Computational Psychiatry and Ageing Research, University*  
11 *College London, London, UK.*

12 \*Corresponding Author

13

14 **Correspondence:** [benjy.barnett.20@ucl.ac.uk](mailto:benjy.barnett.20@ucl.ac.uk)

15

## 16 **Summary**

17 Representing the quantity zero is considered a unique achievement of abstract human thought.  
18 Despite considerable progress in understanding the neural code supporting natural numbers, how  
19 numerical zero is encoded in the human brain remains unknown. We find that both non-symbolic  
20 empty sets (the absence of dots on a screen) and symbolic zero (“0”) occupy ordinal positions  
21 along graded neural number lines within posterior association cortex. Neural representations of  
22 zero are partly independent of numerical format, exhibiting distance effects with countable  
23 numerosities in the opposing (symbolic or non-symbolic) notation. Our results show that format-  
24 invariant neural magnitude codes extend to judgements of numerical zero, and offer support to  
25 theoretical accounts in which representations of symbolic zero are grounded in more basic  
26 representations of sensory absences.

27

28 **Keywords:** zero, numerosity, magnitude codes, sensory absence, perception, decoding,  
29 representational similarity analysis, MEG

## 30 **Introduction**

31 The number zero plays a central role in science, mathematics, and human culture (Kaplan, 1999;  
32 Nieder, 2016) and is considered a unique property of abstract human thought (Bialystok & Codd,  
33 2000; Nieder, 2016). The psychological basis of zero is unusual: while natural numbers  
34 correspond to the observable number of countable items within a set (e.g., one bird; three  
35 clouds), an empty set does not contain any countable elements. To conceptualise zero, one must  
36 instead abstract away from the (absence of) sensory evidence to construct a representation of  
37 numerical absence: creating ‘something’ out of ‘nothing’ (Butterworth, 1999; Nieder, 2016;  
38 Wellman & Miller, 1986). Given these differences, it remains an open question as to how zero is  
39 represented in relation to other numbers.

40

41 In contrast to zero, the neural representation of natural numbers is better understood. Distinct  
42 neural populations are selective for specific numerosities, exhibiting overlapping tuning curves  
43 with neighbouring populations tuned to adjacent numerosities (Kutter et al., 2018; Piazza et al.,  
44 2004). This architecture underpins a so-called distance effect (Dehaene et al., 1998), where  
45 numbers close together in numerical space have similar neural representations. For instance,  
46 neural responses to numbers one and two are more similar than neural responses to one and ten  
47 (Borghesani et al., 2019; Luyckx et al., 2019; Piazza et al., 2004). Importantly, a component of  
48 this neural code is thought to be invariant to numerical format (Damarla et al., 2016; Eger et al.,  
49 2003, 2009; Piazza et al., 2007; Teichmann et al., 2018) such that, for example, neural  
50 representations of ‘six’ are shared across symbolic and non-symbolic notations (e.g., both the  
51 Arabic numeral “6” and six dots; although see Cohen Kadosh et al. (2007)). In humans, these

52 format-invariant representations of numerical magnitude have been localised to the parietal  
53 cortex (Damarla et al., 2016; Eger et al., 2009; Piazza et al., 2007), with topographic maps  
54 underpinning numerosity perception more broadly being found across association cortex (Harvey  
55 et al., 2013; Harvey & Dumoulin, 2017).

56

57 Compared to natural numbers, zero is associated with distinct behavioural and developmental  
58 profiles. For instance, the reading times of human adults are increased for zero compared to non-  
59 zero numbers (Brysbaert, 1995), and zero concepts emerge later in children than those for natural  
60 numbers (Krajcsi et al., 2021; Merritt & Brannon, 2013; Wellman & Miller, 1986). Distinct  
61 behavioural characteristics associated with zero are not unsurprising given the heightened degree  
62 of abstraction required to conceptualise numerical absence. In turn, it is plausible that neural  
63 representations of zero are distinct to the scheme that has been discovered for natural numbers  
64 (e.g Schubert et al., 2020). Initial research in non-human animals has indicated that numerical  
65 zero shares some neural properties with natural numerosities, such as overlapping tuning curves  
66 and associated distance effects, along with invariance to particular stimulus properties  
67 (Kirschhock et al., 2021; Okuyama et al., 2015; Ramirez-Cardenas et al., 2016). However, it  
68 remains unknown whether the symbolic, human conceptualisation of numerical zero, which  
69 emerged independently of natural numbers (Ifrah, 1985; Kaplan, 1999), engenders  
70 representations of zero that are both distinct from other numbers and which studies in non-human  
71 animals may have failed to reveal.

72

73 Investigating neural representations of both non-symbolic and symbolic zero can also shed light  
74 on how a symbolic concept of numerical absence emerged in human thought. One intriguing idea

75 is that, across phylogeny and ontogeny, low-level perceptual representations tracking an absence  
76 of sensory stimulation (e.g. Merten & Nieder, 2012; Goh et al., 2023) gave rise to conceptual  
77 representations of numerical zero (Nieder, 2016). Specifically, an ability to represent perceptual  
78 absences is hypothesised to support an ability to quantify ‘none’ as less than one, which, in turn,  
79 may have given rise to our ability to reason and calculate with an abstract symbol – zero –  
80 denoting lack of quantity (Nieder, 2016). Asking whether symbolic neural representations of  
81 zero share variance with non-symbolic empty sets provides an initial empirical test of this  
82 hypothesis.

83

84 We tackled these questions by employing two qualitatively different numerical tasks in humans  
85 while leveraging methodological advances to reveal the representational content of neural  
86 responses to numerical stimuli in MEG data (Kriegeskorte & Diedrichsen, 2019; Luyckx et al.,  
87 2019). We assay both neural representations of non-symbolic numerosities (dot patterns),  
88 including zero (empty sets), and symbolic numerals, including symbolic zero. We show how  
89 neural representations of zero are situated along a graded neural number line shared with other  
90 natural numbers. Notably, symbolic representations of zero generalised to predict non-symbolic  
91 empty sets – consistent with a hypothesis that human zero is grounded in perceptual absence  
92 (Nieder, 2016). We go on to localise abstract representations of numerical zero to posterior  
93 association cortex, extending the purview of parietal cortex in human numerical cognition to  
94 encompass representations of zero (Harvey & Dumoulin, 2017; Piazza et al., 2007).

95

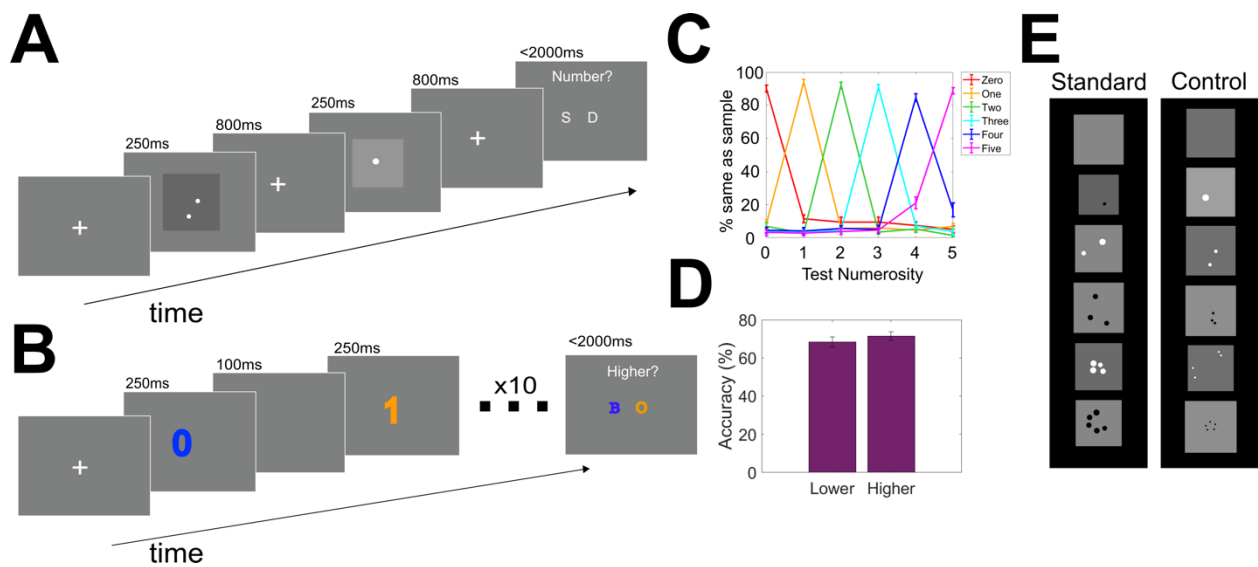
## 96 **Results**

97 29 human participants (24 after exclusions; see Methods for details) took part in a  
98 magnetoencephalography (MEG) experiment involving two numerical tasks. The first was a non-  
99 symbolic match-to-sample task (Figure 1A) where participants observed two sequentially  
100 presented dot patterns that ranged in number from zero dots (empty set) to five dots (Ramirez-  
101 Cardenas et al., 2016). Participants were asked to report whether the patterns contained the same  
102 number of dots or not. We employed two sets of dot patterns: a standard set which randomised  
103 the size of dots within each pattern, and a control set which kept size, density, and luminance  
104 constant across numerosities (Figure 1D). The second task was a symbolic averaging task  
105 (Figure 1B; Luyckx et al., 2019). Here, participants observed a rapid serial presentation of 10  
106 symbolic numerals from zero to five (0, 1, 2, 3, 4, 5), divided into orange and blue sets (5  
107 numbers in each). Participants were asked to report the set of numbers with the higher or lower  
108 average. Decision type (higher or lower) was counterbalanced across participants. The use of two  
109 different tasks (match-to-sample, averaging), and two different decision types in the averaging  
110 task ensured neural patterns induced by the perception of zero are unlikely to be driven by  
111 specific task features or calculation requirements.

112

113 In the non-symbolic match-to-sample task, participants accurately determined whether dot  
114 patterns had the same or different numbers of dots ( $Mean_{accuracy}$ : 0.92,  $SE$ : 0.16). Plotting  
115 behavioural tuning curves revealed near-ceiling performance across all numerosities (Figure 1C),  
116 with the exception of five-dot patterns which were more often confused with four-dot patterns  
117 than three-dot patterns ( $t(23) = 4.97$ ,  $p < .001$ ) – consistent with numerosity tuning curves

118 becoming wider as number increases (Dehaene et al., 1998). In the symbolic task, participants  
119 could reliably perform the task regardless of whether they were reporting the higher  
120 ( $Mean_{accuracy}$ : 0.71,  $SE$ : 0.23) or lower ( $Mean_{accuracy}$ : 0.68,  $SE$ : 0.27) average (Figure 1D), and  
121 there was no difference between performance across decision types ( $t(22) = -0.88, p = 0.39$ ).  
122



124 **Figure 1. Experimental Procedure.** *A. Trial structure for the non-symbolic match to sample*  
125 *task.* Participants observed a sample dot pattern followed by a test dot pattern before reporting  
126 whether the two patterns had the same or different numbers of dots. *B. Trial structure for the*  
127 *symbolic averaging task.* Participants observed a sequence of blue and orange numerals before  
128 reporting which set of numerals had the higher or lower average. *C. Behavioural tuning curves in*  
129 *the non-symbolic task.* Each curve reflects the percentage of trials that participants judged the  
130 test numerosity to be the same as the sample numerosity. Each colour represents trials with  
131 specific sample numerosities. The peak of each curve illustrates correct performance when the  
132 sample and test numerosities matched. Data points either side of the peak represent non-match  
133 trials. Error bars indicate SEM. *D. Accuracy in the symbolic task split across participants who*

134 judged which set of numbers was higher, and those who judged which was lower. *E. Stimulus*  
135 *sets for dot task.* Dot size was pseudorandomised in the standard set, while low level properties  
136 of the dots including size, density, and luminance were held constant in the control set.

137

### 138 Identifying Neural Representations of Number

139 We next asked whether neural patterns recorded by MEG were sensitive to numerosity, by  
140 timelocking our data to the presentation of the dot pattern/symbolic numeral stimuli. Multiclass  
141 decoders were trained to classify different numerosities (zero to five) in both the non-symbolic  
142 and symbolic tasks. The frequency with which the decoders confused numerosities for one  
143 another is illustrated in Figure 2A. Here, individual panels represent trials where a particular  
144 numerosity was presented to the classifier, and the coloured lines indicate the proportion of those  
145 trials where the classifier predicted each one of the possible classes (zero to five) over the trial  
146 epoch. For example, the “NS-one” panel shows that when one dot is presented in the non-  
147 symbolic task, the classifier predominantly and correctly labels this stimulus as numerosity one  
148 (yellow curve), with the next most likely error being a misclassification as the number two  
149 (green curve). Across all numbers and both formats, the classifiers successfully predicted the  
150 numerosity participants were viewing from their neural data, including zero numerosities.

151

152 We next leveraged temporal generalisation analysis to ask whether numerosity representations  
153 were stable over time (King & Dehaene, 2014) (Figure 2B). When training and testing on all  
154 combinations of time points, stable time-windows where numerical information could be  
155 decoded above chance level were identified in both tasks from shortly after stimulus presentation



156 up until the end of the analysed time window (non-symbolic: 70ms – 800ms; symbolic 56.7ms –  
157 800ms). These time windows in which stable numerosity representations were identified were  
158 used to create time-averaged data for use in subsequent population tuning curve (Figure 2D) and  
159 multidimensional scaling (Figure 2E) analyses.

160

### 161 A Neural Number Line from Zero to Five

162 A fundamental feature of neural codes for natural numbers is a distance effect, whereby numbers  
163 closer together in numerical space are closer together in representational space (Dehaene et al.,  
164 1998; Nieder & Dehaene, 2009). Here we asked whether numerical zero exhibits similar distance  
165 effects with other numbers, consistent with it sharing a neural number line with countable  
166 numerosities. A Representational Similarity Matrix (RDM) describing a distance effect from  
167 zero to five successfully predicted neural data across both non-symbolic and symbolic numerical  
168 formats (Figure 2A). In the non-symbolic task, an RDM generalising numerical information  
169 across the two non-symbolic stimulus sets significantly predicted neural responses throughout  
170 the trial, indicating that neural correlates of number were independent of the physical properties  
171 of the dot stimuli (Supplemental Figure 1). Multidimensional scaling of neural representations  
172 of numerosity in turn illustrates a distance effect (Figure 2E), with the numbers zero to five  
173 occupying positions along a single, ordered dimension, while a second dimension loosely  
174 distinguished intermediate numerosities (one to four) from the extremes (zero and five).

175

176 A stronger test of a distance effect in neural data is furnished by examining the confusability  
177 between neighbouring numerosities using population tuning curves (Figure 2D). These plots are

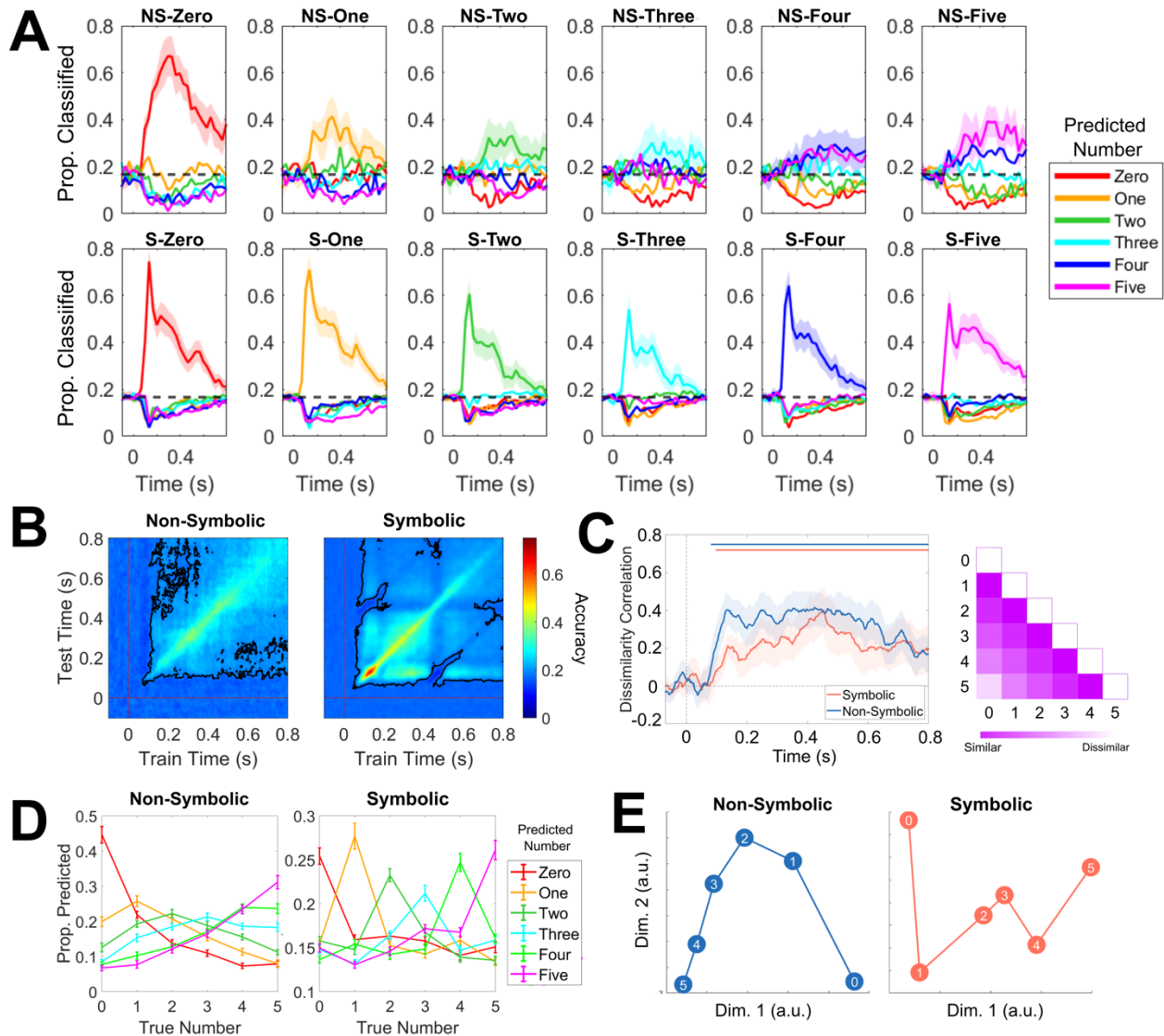
178 time-averaged versions of the classifier confusion matrices in Figure 2A, i.e. the proportion of  
179 trials where the classifier predicted a particular numerosity as a function of the true numerosity  
180 within the time window in which numerical information could be reliably decoded (Figure 2B).  
181 For example, the red curve in Figure 2D indicates that the proportion of trials predicted as being  
182 zero peaks when the numerosity seen by the decoder was also zero, is next highest when the  
183 numerosity seen by the decoder was “one”, and so on.

184

185 In the non-symbolic task (Figure 2D, left), the classifier confuses zero with one  
186 ( $Mean_{proportion\ predicted} = 0.218$ ) more often than it confuses zero with two  
187 ( $Mean_{proportion\ predicted} = 0.138$ ) ( $t(23) = 6.23, p < .001$ ). Similarly, it confuses one with two  
188 ( $Mean_{proportion\ predicted} = 0.206$ ) more often than with three ( $Mean_{proportion\ predicted} =$   
189  $0.155$ ) ( $t(23) = 4.76, p < .001$ ). This pattern of results is indicative of a gradedness in the  
190 representation of numerical magnitude across non-symbolic numerosities. In contrast, in the  
191 symbolic task (Figure 2D, right), the multiclass classifier does not confuse zero with one  
192 ( $Mean_{proportion\ predicted} = 0.159$ ) significantly more than it confuses zero with two  
193 ( $Mean_{proportion\ predicted} = 0.163$ ) ( $t(23) = -0.61, p = 0.54$ ), nor does it confuse one with two  
194 ( $Mean_{proportion\ predicted} = 0.153$ ) significantly more than it confuses one with three  
195 ( $Mean_{proportion\ predicted} = 0.143$ ) ( $t(23) = 1.67, p = 0.11$ ). This difference in distance effects  
196 between non-symbolic and symbolic formats was statistically significant for both zero ( $t(23) =$   
197  $5.45, p < .001$ ) and one ( $t(23) = 3.48, p = .002$ ), and is suggestive of more gradedness in the  
198 representation of non-symbolic than symbolic numerosities, consistent with previous work  
199 describing narrower tuning curves for symbolic numerals (Eger et al., 2009; Kutter et al., 2018).

200

201



202

203 **Figure 2. A Neural Number Line from Zero to Five.** *A.* Across-time confusion matrices for

204 multiclass decoders classifying non-symbolic (top) and symbolic numerosities (bottom).

205 Individual panels represent trials where particular numerosities were presented to the classifier.

206 Coloured lines indicate the proportion of those trials where the classifier predicted each

207 numerosity. *B.* Temporal generalisation of multiclass decoders trained to decode numerosities

208 zero to five in the non-symbolic (left) and symbolic (right) task reveals stable numerical  
209 representations over time in both tasks emerging shortly after stimulus presentation. Black lines  
210 illustrate timepoints where decoding was significantly above chance ( $p < .05$ , corrected for  
211 multiple comparisons). These stable time windows were used in the time-averaged analyses  
212 depicted in panels D and E. *C*: A model representational dissimilarity matrix (RDM) describing  
213 a distance effect from zero to five significantly predicted neural data in both non-symbolic and  
214 symbolic tasks. The diagonal of the RDM was not included in this analysis, preventing the self-  
215 similarity of each number from trivially explaining our results. Shaded areas indicate 95%  
216 confidence intervals. Horizontal lines show clusters of time where dissimilarity correlations were  
217 significantly above 0,  $p < .05$  corrected for multiple comparisons. *D*. *Population level tuning*  
218 *curves derived from decoder confusion matrices*. Each curve represents the proportion of trials  
219 the classifier predicted a particular numerosity (indicated by the curve's colour) as a function of  
220 the numerosity the decoder actually saw. For example, the red curve illustrates how the  
221 prediction of numerosity zero is distributed across different presented numerosities. For non-  
222 symbolic numerosities, the classifier confused numbers as a function of their numerical distance,  
223 consistent with a graded representation of numerical magnitude. In the symbolic task,  
224 representations were more categorical than graded. Error bars represent SEM. *E*.  
225 Multidimensional scaling of numerical representations in both tasks revealed a principal  
226 dimension which tracks numerical magnitude and a second dimension which loosely  
227 distinguishes extreme values from intermediate values.

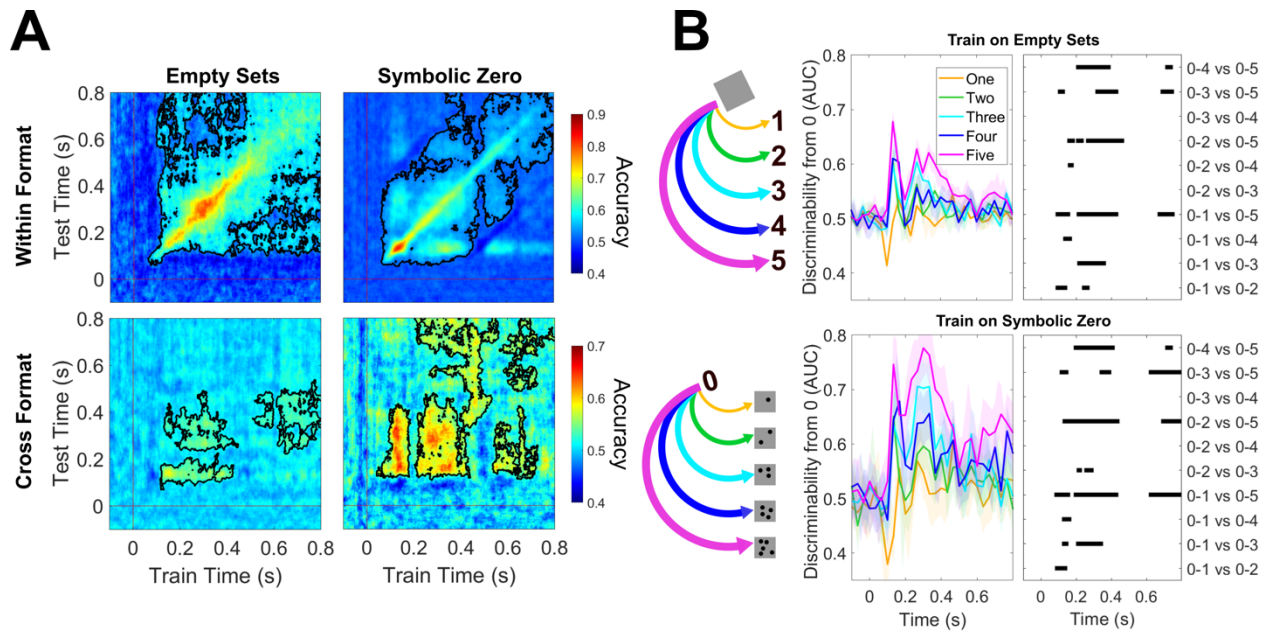
228

229 Representations of Zero are Shared Between Symbols and Empty Sets

230 Together, our previous analyses establish that neural representations of zero are graded  
231 (especially for non-symbolic numerosities) and situated within a number line spanning other  
232 countable numerosities from 1 to 5. We next asked whether representations of zero were format-  
233 and task-independent – generalising across non-symbolic (empty set) and symbolic (“0”) stimuli,  
234 and across the same/different and averaging tasks. To test this hypothesis, we performed further  
235 decoding analyses focused on dissociating numerical zero from non-zero numerosities. If a  
236 binary classifier trained to distinguish zero from non-zero numerosities in one numerical format  
237 is subsequently able to separate zero from non-zero numerosities in another numerical format,  
238 this furnishes evidence for an abstract neural representation of numerical absence that is common  
239 to both formats.

240

241 Decoders trained to distinguish numerical absence within each format separately revealed stable  
242 representations of numerical zero from approximately 100ms to 450ms after stimulus  
243 presentation, before exhibiting a more dynamic temporal profile until the end of the trial epoch  
244 (Figure 3A, top). Crucially, these decoders could also successfully classify representations of  
245 zero in the opposing format to which they had been trained (Figure 3A, bottom) – both when  
246 generalising from empty sets to the decoding of symbolic numerosities, and when generalising  
247 from symbolic zero to non-symbolic dot stimuli. This cross-decoding was successful over the  
248 initial 350ms period where the within-format decoders identified stable representations of  
249 numerical absence, although generalisation was generally stronger when generalising from  
250 symbolic zero to empty sets than vice-versa.



251

252

253 **Figure 3. Abstract and Graded Representations of Numerical Absence. A. Representations**

254 *of numerical absence generalise over numerical format. Top: A decoder trained to decode zero*

255 *from natural numbers reveals stable representations of zero up to ~450ms after stimulus*

256 *presentation for both non-symbolic (left) and symbolic (right) formats, with more dynamic /*

257 *unstable representations observed towards the end of the epoch. Bottom-left: A decoder trained*

258 *to decode empty sets also distinguished symbolic zero from non-zero symbolic numerals.*

259 *Bottom-right: A decoder trained to distinguish symbolic zero from non-zero symbolic numerals*

260 *also distinguished empty sets from non-symbolic numerosities. Black lines indicate clusters of*

261 *significantly above chance decoding,  $p < .05$ , corrected for multiple comparisons. B. Left:*

262 *Illustration of the hypothesis that abstract representations of numerical absence are situated on a*

263 *graded number line that generalises across format, with empty sets represented as more similar to*

264 *symbolic numeral one than numeral five (top), and symbolic zero as more similar to one dot than*

265 *five dots (bottom). Centre: Training a classifier to decode non-symbolic empty sets from non-*

266 symbolic numerosities and testing it on symbolic numbers in a pairwise manner revealed  
267 increasing discriminability as distance from zero increased (top). The same cross-format distance  
268 effect is observed when training a classifier on symbolic zero and testing it on non-symbolic  
269 numerosities (bottom). Shaded areas represent 95% CIs. *Right*: clusters of significant differences  
270 between different numerosities' discriminability from zero,  $p < .05$ , corrected for multiple  
271 comparisons. An increase in discriminability for numbers further from zero reveals a cross-  
272 format distance effect.

273

#### 274 Graded Representations of Zero are Invariant to Numerical Format

275 Next, to establish the representational structure of cross-format representations of zero, we  
276 leveraged the numerical distance effect already identified for within-format representations  
277 (Figure 2). To test for such effects, we computed the discriminability between zero and each  
278 non-zero numerosity in the alternative numerical format (Figure 3B, middle). Strikingly, neural  
279 representations of symbolic zero (“0”) were more often confused with one or two dots in the  
280 non-symbolic task, than they were with four or five dots (Figure 3B, middle). Similarly, neural  
281 representations induced by non-symbolic zero (empty sets) were more often confused with the  
282 symbolic numeral 1 or 2 than they were with symbolic numerals 4 or 5. Pairwise tests comparing  
283 the discriminability of different non-zero numerosities from zero revealed clusters of significant  
284 differences in discriminability (Figure 3B, right), with an increased distance from zero increasing  
285 discriminability. Together, these cross-format analyses support a hypothesis that an approximate,  
286 graded representation of numerical absence is engaged not only by symbolic zero (“0”) but also  
287 by lower-level perceptual absences (empty set stimuli).

288



289 We also sought to test a more stringent hypothesis that abstract, format-independent neural  
290 representations of zero are themselves situated within a cross-format neural number line –  
291 essentially extending the question of format-independence to now include all numerosities from  
292 0 to 5. A representational dissimilarity matrix situating abstract numerosity representations  
293 within a graded number line significantly predicted the neural data (Supplemental Figure 2, left).  
294 Testing for cross-format distance effects between all numerosities using RSA also revealed a  
295 qualitative distance effect, although this did not reach statistical significance (Supplemental  
296 Figure 2, right). Finally, multidimensional scaling of neural representations induced by symbolic  
297 and non-symbolic numerosities in a shared space corroborated evidence for a distance effect for  
298 zero across tasks (Supplemental Figure 3).

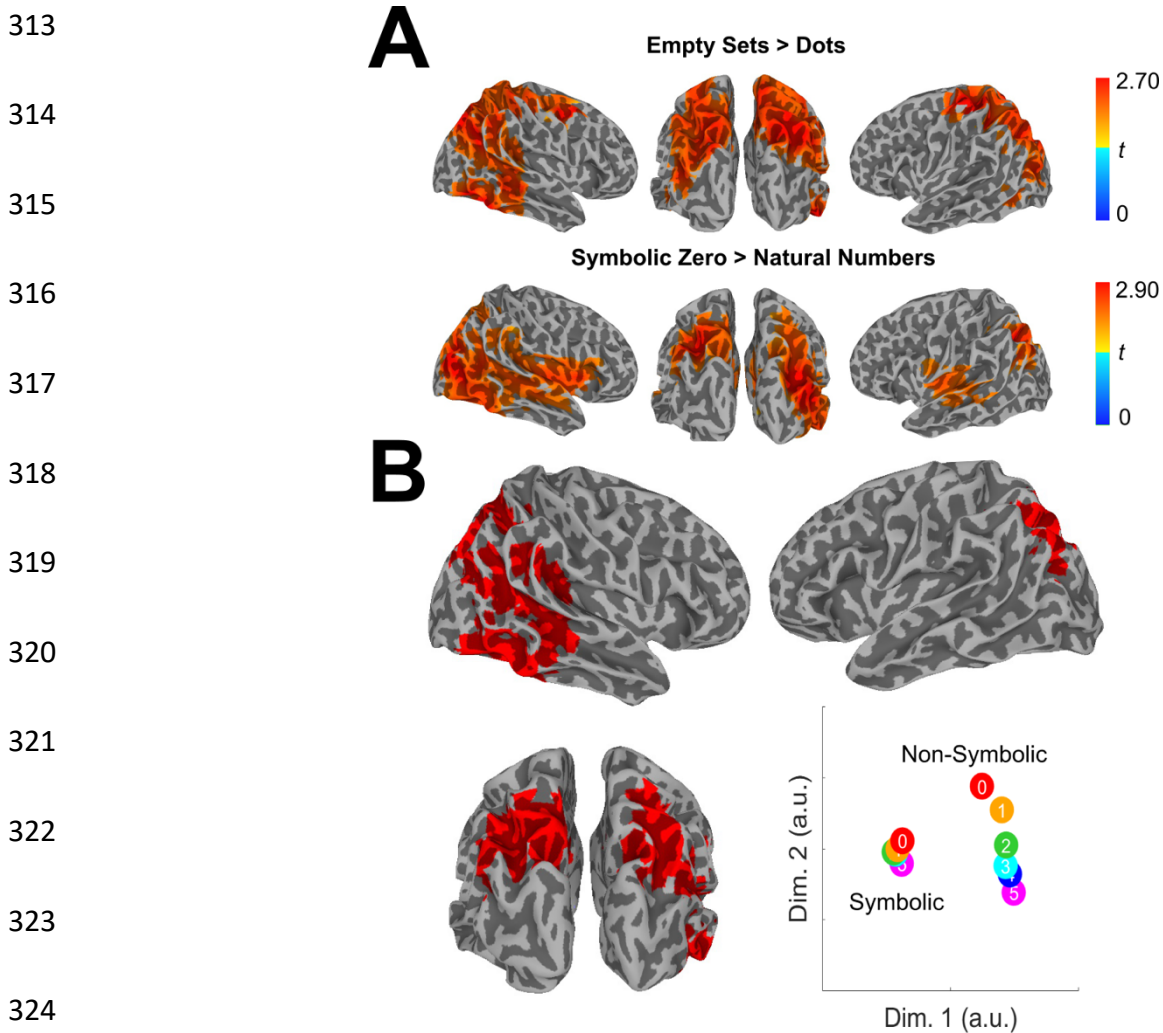
299

### 300 Abstract Representations of Numerical Zero are Localised to Posterior Association Cortex

301 Finally, we sought to reconstruct and compare source-level neural activity for zero and non-zero  
302 numerosities in both the non-symbolic and symbolic tasks. By performing mass-univariate  
303 contrasts of broadband source power (zero > non-zero numerosities) in both the non-symbolic  
304 (Figure 4A top; peak voxels ( $xyz$ ): left hemisphere = -36, -24, 56; right = 60, -64, -24) and  
305 symbolic (Figure 4A bottom; peak voxels ( $xyz$ ): left hemisphere = -28, -56, 32; right = 28, -72, 8)  
306 tasks and computing the conjunction between these two contrasts (Figure 4B; peak voxels within  
307 conjunction ( $xyz$ ): non-symbolic task: left hemisphere = -20, -64, 32; right = 60, -64, -24;  
308 symbolic task: left hemisphere = -28, -56, 32; right = 20, -48, -64), we are able to show that  
309 neural activity induced by numerical absence is distributed across the posterior association cortex  
310 (Figure 4B). Neural responses to zero within this conjunction map were again situated within a



311 number line populated by non-zero numbers, with numerical magnitude increasing along a single  
312 dimension that was similar for both symbolic and non-symbolic formats (Figure 4, bottom-right).



325 **Figure 4. Neural activity induced by numerical zero localised to posterior association**

326 **cortex.** *A*: Mass univariate contrasts of source power revealed regions more active following

327 presentations of zero vs. non-zero numerosities in non-symbolic (top) and symbolic (bottom)

328 tasks. Colour represents t-value and only clusters significant at  $p < .05$  are presented, corrected

329 for multiple comparisons. *B*: A conjunction of zero > non-zero contrasts in both numerical

330 formats yielded a map identifying broad regions of the posterior association cortex as

331 representing numerical absence across numerical formats. Multidimensional scaling of each

332 numerosity's neural pattern within these regions revealed a graded representational structure of  
333 numerical magnitude along a single dimension that was similar for both formats.

334

## 335 **Discussion**

336 The number zero is associated with unique psychological properties compared to natural  
337 numbers. Here, we characterise the neural representation of numerical zero in the human brain.  
338 We describe how numerical zero occupies a slot at the lower end of neural number lines for both  
339 symbolic and non-symbolic numerical formats. Strikingly, we show that a component of this  
340 representation is both task- and format-independent, such that empty sets – the absence of dots –  
341 generalised to predict the neural profiles and distance effects observed for symbolic zero. These  
342 abstract, format-invariant representations of zero were situated at the lower end neural code for  
343 number that was localised across the posterior association cortex.

344

345 That zero is situated at the lower end of a neural number line in the human brain is consistent  
346 with an emerging body of work examining representations of zero in non-human animals  
347 (Kirschhock et al., 2021; Okuyama et al., 2015; Ramirez-Cardenas et al., 2016). Across two  
348 different studies, single neurons selective for non-symbolic empty sets were found in the parietal  
349 and prefrontal cortex of non-human primates (Okuyama et al., 2015; Ramirez-Cardenas et al.,  
350 2016). In line with the present results, many of these neurons – but not all – were found to  
351 exhibit distance effects with non-zero numbers. When comparing non-symbolic and symbolic  
352 instances of zero, we found symbolic instances were more discrete and less graded than non-  
353 symbolic instances, consistent with work describing sharper tuning curves for symbolic number

354 representations (Eger et al., 2009; Kutter et al., 2018). Recent single-cell recordings in the human  
355 medial temporal lobe have also identified discrete symbolic zero-selective neurons that did not  
356 exhibit graded activations in relation to non-zero symbolic numerals (Kutter et al., 2023).  
357 Strikingly, however, the majority of our analyses revealed a graded representation of zero that  
358 generalised across both symbolic and non-symbolic formats, in keeping with behavioural  
359 findings that situate zero at the lower end of a graded psychological number line in humans  
360 (Merritt & Brannon, 2013).

361

362 Our finding that representations of numerical absence have a format-invariant component  
363 extends previous work documenting neural representations of numerosity that generalise across  
364 countable non-symbolic elements and their symbolic counterparts (Eger et al., 2009; Libertus et  
365 al., 2007; Piazza et al., 2007). Here we show how neural representations of non-symbolic empty  
366 sets, which do not contain any countable items, also share variance with symbolic zero (Figure  
367 3). These abstract representations of zero were localised to regions of the posterior association  
368 cortex that have previously been associated with numerical processing (Figure 4B; Arsalidou &  
369 Taylor, 2011; Eger et al., 2003; Harvey & Dumoulin, 2017; Piazza et al., 2007). It remains  
370 debated whether findings of format-invariant numerical codes are explained by single neurons  
371 coding for the same numerosities across formats, or whether they reflect the recruitment of  
372 neighbouring format-specific neural populations that are interdigitated within a numerosity map  
373 (Cohen Kadosh & Walsh, 2009). Future intracranial recording studies will be required to  
374 determine whether single cells in the human brain code for numerical zero in both non-symbolic  
375 and symbolic formats. However, our finding of cross-format distance effects is more consistent  
376 with a shared neural code, as it is less likely that spatially overlapping but format-specific neural

377 codes would also generalise to exhibit cross-format distance effects with more distant  
378 numerosities.

379

380 Finding shared neural representations of non-symbolic empty sets and symbolic zero lends  
381 weight to recent suggestions that representations of numerosity zero may have emerged from  
382 more fundamental representations of sensory absence (Nieder, 2016). On this account, low-level  
383 perceptual representations indicating an absence of sensory stimulation (e.g. Merten & Nieder,  
384 2012; Goh et al., 2023) provide the perceptual grounding for a conceptual representations of  
385 numerical zero (Nieder, 2016) – consistent with a broader principle that the human brain co-opts  
386 basic sensory and motor functions in the service of more complex cognitive abilities (Dehaene &  
387 Cohen, 2007). Such a hypothesis is consistent with similar behavioural signatures for the  
388 processing of absence across perceptual and numerical domains. For instance, reading times are  
389 increased for number zero compared to non-zero numbers (Brysbart, 1995), whilst reaction  
390 times for deciding a stimulus is absent are higher than for deciding a stimulus is present (Mazor  
391 et al., 2020, 2021). Additionally, judgements about the absence of features mature later in  
392 children than judgements about presence (Coldren & Haaf, 2000; Sainsbury, 1971), a  
393 developmental pattern mirrored by the late mastery of numerical zero in children (Krajcsi et al.,  
394 2021; Merritt & Brannon, 2013; Wellman & Miller, 1986). We note however that the neural  
395 responses recorded in our study to empty-set stimuli were still within the context of a numerical  
396 task – and, as such, only provide initial evidence for a perceptual grounding of zero. A stronger  
397 test of this hypothesis would examine shared representations of numerical and perceptual  
398 absence – for instance, classifying a stimulus as absent in a non-numerical task (Mazor et al.,  
399 2020; Merten & Nieder, 2012; Barnett et al., 2023).

400

401 We took care to ensure that the neural representations of zero identified in our data were not  
402 trivial consequences of zero being classified as the “lowest” stimulus in our tasks. The concern  
403 here is that if our tasks required participants to adopt a particular mathematical attitude towards  
404 zero, then decoding of this task-dependent concept would confound any results aimed at  
405 identifying task-invariant representations of numerical absence. We consider this explanation of  
406 our results as unlikely, however, as, by design, the symbolic and non-symbolic tasks required  
407 adopting qualitatively distinct mathematical attitudes towards zero stimuli: the match-to-sample  
408 task necessitated deciding whether two dot stimuli were the same or different, whereas the  
409 symbolic task required maintenance of condition-specific numerical averages. Because the non-  
410 symbolic task did not require participants to order stimuli, any format-invariant representations  
411 of zero cannot be explained by a generic requirement to identify lower vs. higher numerosities.

412

413 The adoption of the number zero has enabled great advances in science and mathematics  
414 (Kaplan, 1999). Here, we show that the human brain represents this unique number by  
415 incorporating representations of numerical absence into a broader neural coding scheme that also  
416 supports countable numerosities. Representations of numerical zero were found to be format-  
417 invariant and graded with respect to non-zero numerosities, and were localised to regions of the  
418 posterior association cortex previously implicated in numerical cognition. Our results  
419 demonstrate that neural number lines include zero, and, more importantly, provide initial  
420 evidence that the abstract concept of symbolic zero is linked to representations of non-symbolic  
421 empty sets. Our study lays the foundations for a deeper understanding of how the human ability

422 to represent the number zero may be grounded in perceptual capacities for detecting an absence  
423 of sensory stimulation.

424

## 425 **Materials and Methods**

### 426 Participants

427 29 participants ( $M_{age}$ : 29.27 years,  $SD_{age}$ : 10.69) took part in the MEG experiment at the  
428 Wellcome Centre for Human Neuroimaging, University College London. 5 participants either  
429 failed to follow task instructions (chance performance on one or more tasks) or did not complete  
430 the experiment and were therefore excluded from analysis. All analysis was performed on the  
431 remaining sample of 24 participants. Informed consent was given before the experiment and  
432 ethical approval was granted by the Research Ethics Committee of University College London  
433 (#1825/005).

434

### 435 Stimuli

436 Numerical dot stimuli were created using custom MATLAB (Mathworks) scripts and consisted  
437 of different numbers of dots (from zero to five) on grey backgrounds (Figure 1C). There were  
438 two sets of dot stimuli, a standard set and control set. In the standard set, dot size was  
439 pseudorandomly specified, while in the control set, low level visual properties of the stimuli (dot  
440 size, density, luminance) were constant across numerosities. Empty set stimuli contained only a  
441 grey background in both stimulus sets. To help prevent participants relying on low level visual  
442 cues in identifying empty set stimuli, the background luminance was varied within and across  
443 stimulus sets, the background square size was randomly varied across all stimuli, and 50% of

444 dots were white rather than black. A control analysis confirmed that numerical information was  
445 extracted from the stimuli independently from physical features (see Representational Similarity  
446 Analysis; Supplemental Figure 1).

447

#### 448 Experimental Procedure

449 The tasks were presented to subjects using MATLAB (Mathworks) and the Psychophysics  
450 Toolbox (Brainard, 1997; Kleiner et al, 2007). Participants practiced the tasks on a computer  
451 before the MEG session. In the MEG scanner, the tasks were performed in alternating  
452 miniblocks with 35 symbolic trials and 54 non-symbolic trials per MEG recording block. The  
453 order of the tasks would swap on each block, and the starting order was counterbalanced across  
454 participants. There were 9 MEG blocks in total, resulting in 315 symbolic numeral trials and 486  
455 non-symbolic dot trials across the whole experiment. Participants responded using two buttons  
456 and their right thumb.

457

#### 458 *Non-symbolic Task*

459 Participants performed a match to sample task on dot stimuli (Kirschhock et al., 2021; Ramirez-  
460 Cardenas et al., 2016). On each trial, participants saw a sample image containing between zero  
461 and five dots for 250ms followed by a fixation cross for 800ms. A test image, also containing  
462 between zero and five dots, was then presented for 250ms, followed before another 800ms  
463 fixation period (Figure 1A). Within a trial, a single stimulus set was used for both the sample and  
464 test image. Participants reported whether the number of dots in the test stimulus matched that of  
465 the sample stimulus, or not. The response was followed by feedback in the form of a coloured

466 rectangle surrounding the response options, with green and red used to indicate correct and  
467 incorrect answers, respectively. Response options were positioned randomly on each trial to  
468 eliminate any correlation between the decision and motor response. Intertrial intervals were also  
469 sampled randomly from a uniform distribution between 500-1000ms.

470

#### 471 *Symbolic Task*

472 We adapted the symbolic numeral averaging task introduced by Luyckx *et al.* (2019) to include  
473 the number zero. In one trial, ten numerals ranging from zero to five were presented in a random  
474 order (Figure 1B). Five of the numerals were blue and five were orange. Each numeral was  
475 displayed for 250ms with an interstimulus interval of 100ms. The numerals were randomly  
476 selected on each trial to obey the constraint that the mean of the blue numerals could not equal  
477 the mean of the orange numerals. The response required at the end of each trial was  
478 counterbalanced across subjects, with half of the subjects reporting which set of numerals  
479 (orange or blue) had the highest average, and the other half reporting the set with the lowest  
480 average. Participants had 2000ms to respond, after which they were given feedback in the form  
481 of a green (correct) or red (incorrect) rectangle surrounding the response options. Again, to  
482 disentangle participants' decisions from motor responses, response options were positioned  
483 randomly on each trial. Intertrial intervals were randomly sampled from a uniform distribution  
484 between 500-1000ms.

485

#### 486 MEG Preprocessing



487 MEG data were analysed using FieldTrip (Oostenveld et al., 2011). MEG was recorded  
488 continuously at 600Hz using a 273-channel axial gradiometer system (CTF Omega, VSM  
489 MedTech) while participants sat upright inside the scanner. To remove line noise, the raw MEG  
490 data were preprocessed with a Discrete Fourier Transform and bandstop filter at 50Hz and its  
491 harmonics. The numeral task was segmented into epochs of -500ms to 4000ms relative to trial  
492 onset. For the dot task, the segments were from -200ms to 2500ms. Baseline correction was  
493 performed where, for each trial, activity in the pre-trial window was averaged and subtracted  
494 from the entire epoch per channel. The data were downsampled to 300Hz to conserve processing  
495 time and improve signal to noise ratio. During artefact rejection, trials with high kurtosis were  
496 visually inspected and removed if they were judged to contain excessive artefacts. To assist in  
497 removing eye-movement artefacts, an independent components analysis was carried out on the  
498 MEG data, and the components with the highest correlation with eye-tracking data were  
499 discarded after visual inspection. Components showing topographic and temporal signatures  
500 typically associated with cardiac artefacts were also removed by eye. This procedure was  
501 performed separately for the numeral and dot task. Finally, a second stage of epoching was  
502 performed to generate trials of individual numerosities. In the numeral task, trials were  
503 segmented into -100ms to 800ms epochs around each numeral onset. Trials were then baseline  
504 corrected again using the pre-stimulus window. In the dot task, trials were segmented into two  
505 different -200ms to 800ms epochs with respect to the onsets of the sample and test stimuli. All  
506 analyses used the sample images only. Finally, all analyses focusing on shared representations  
507 across notational formats were performed on the shared timepoints of -100ms to 800ms relative  
508 to stimulus presentation.

509

## 510 Representational Similarity Analysis

511 Representational Similarity Analysis (RSA) allows us to test specific hypothesis about how  
512 neural representations are structured (Kriegeskorte & Kievit, 2013). Here, we tested for the  
513 existence of a distance effect across numerosities (Figure 2C). To do this, we defined a model  
514 representational dissimilarity matrix (RDM) that describes the dissimilarity of two numerosities  
515 as a function of their numerical distance. To compare this model dissimilarity matrix with the  
516 neural data we first created a neural dissimilarity matrix that represents the similarity in neural  
517 patterns associated with each numerosity. To do this, we first ran a linear regression on the MEG  
518 data with dummy coded predictors for each of the six numerosities (trial numerosity coded with  
519 a 1, alternative numerosities coded with a 0). This produced a coefficient weight for each  
520 numerosity at each time point and sensor. These weights were then combined into a vector,  
521 representing the multivariate neural response for each numerosity, averaged over trials. To create  
522 the neural RDM, we computed the Pearson distance between each pair of condition weights over  
523 sensors, resulting in a 6x6 neural RDM reflecting the pairwise similarity of neural patterns  
524 associated with each numerosity. These neural RDMs were smoothed over time via convolution  
525 with a 60ms uniform kernel. To compare the neural and model RDMs, at every time point we  
526 correlated the lower triangle of each matrix (excluding the diagonal) using Kendall's Tau rank  
527 correlation (Nili et al., 2014).

528

529 Cross-task RSA was performed in the same manner, except here there were 12 predictors in the  
530 linear regression (0-5 symbolic, 0-5 non-symbolic). This resulted in a 12x12 neural RDM, of  
531 which we used the quadrant representing the cross-task pairwise similarities between  
532 numerosities when comparing with the model RDM (Supplemental Figure 2). The whole

533 quadrant including the diagonal was used in this analysis. This is because here the diagonal does  
534 not contain redundant information, but rather the similarity of the same numerosity across two  
535 different notations, and cells in the upper triangle represent different pairwise similarities to  
536 those in the lower triangle.

537

538 Finally, to test whether numerical information was decodable from non-symbolic stimuli over  
539 and above the physical features of the stimuli, we ran a cross-stimulus set RSA in the same  
540 manner as above, except now we tested exclusively within the non-symbolic task (Supplemental  
541 Figure 1). As such, the 12 predictors were: 0-5 from the standard set and 0-5 from the control set.  
542 This RSA established whether representations of numerical magnitude generalised across  
543 stimulus set, and therefore went beyond information that could be derived solely from physical  
544 features of the stimuli.

545

#### 546 Decoding Analyses

547 To examine the representational structure of the number zero more specifically across symbolic  
548 and non-symbolic formats, we employed different decoding techniques using both multiclass and  
549 binary decoders. First, to reveal the temporal profile of numerosity representations, we trained a  
550 multiclass Linear Discriminant Analysis (LDA) decoder to decode numerosities zero to five  
551 (Figure 2B). This was performed in a temporal generalisation procedure, whereby the classifier  
552 was trained on each time point and tested on all other time points (King & Dehaene, 2014). This  
553 process results in a train time x test time decoding accuracy matrix, which illustrates how stable  
554 representations of numerosity are over time.

555

556 We performed both within-notation and cross-notation decoding procedures. Within-notation  
557 decoding involved training and testing a classifier to identify numerosities on trials from one  
558 notation (e.g. numerals or dots). In cross-notation decoding, we trained the classifier on one  
559 notation and tested it on the other (e.g., training on symbolic trials and testing on non-symbolic  
560 trials, and vice versa). For the within-notation approach, we implemented a 5-fold cross-  
561 validation strategy. Prior to decoding, five trials per numerosity were averaged and the resulting  
562 average trials was balanced per numerosity. It is worth noting that cross-validation is not  
563 required in cross-notation decoding because the test data is never seen by the classifier during  
564 training, and thus there is no risk of overfitting. Cross-notation decoding allows us to empirically  
565 assess whether the neural patterns associated with numerals share a common neural code across  
566 notations.

567

568 To complement our RSA analyses and isolate the representational structure underpinning  
569 numerical zero specifically, we extracted the confusion matrices from the decoders (Figure 2A,  
570 2D). Confusion matrices indicate how often different stimulus classes (i.e., numerosities) are  
571 confused for one another, and this information can be used to infer the organisation of neural  
572 representations. For example, a decoder that confuses zero with the number one more than the  
573 number two displays evidence for a numerical distance effect. The data used to train the decoders  
574 from which these confusion matrices were extracted was time-averaged over the timepoints  
575 where the initial multiclass decoder could decode numerosity significantly above chance (non-  
576 symbolic: 70ms – 800ms, symbolic 56.7ms – 800ms; Figure 2B). We also computed confusion  
577 matrices across time (Figure 2A).

578

579 To examine whether representations of zero could reliably be dissociated from numerosities  
580 presented in the alternative format, we created a decoding procedure using a binary LDA  
581 classifier to decode zero vs. non-zero numerosities (Figure 3A). Within this training regime, the  
582 number of trials per non-zero numerosities was kept equal, and the number of zero trials vs. non-  
583 zero numerosity trials was also balanced. The resulting ‘zero’ decoder was uniquely trained to  
584 identify neural representations of numerical zero in symbolic or non-symbolic notation and was  
585 tested on the other format to identify format-invariant representations of zero.

586

587 Finally, to reveal whether abstract representations of numerical zero exist on a graded number  
588 line, or whether they are categorically distinct from other numbers, we ran a new cross-format  
589 decoding analysis using binary classifiers. Here, we trained the decoders to discriminate zero vs.  
590 all non-zero numerosities (one to five) separately, and then tested these binary decoders on the  
591 corresponding numerosities in the opposite notation. This resulted in five different classifiers per  
592 notation. Specifically, we trained five different decoders to dissociate: symbolic zero vs symbolic  
593 one, symbolic zero vs symbolic two, symbolic zero vs symbolic three, symbolic zero vs  
594 symbolic four, and symbolic zero vs symbolic five. We then tested these decoders on empty sets  
595 vs one dot, empty sets vs two dots, empty sets vs three dots, empty sets vs four dots, and empty  
596 sets vs five dots, respectively. This was also done in the reverse direction: training on non-  
597 symbolic trials and testing on symbolic numerals. We used the area under the receiver operating  
598 characteristic (AUROC) as a metric for discriminability between each pair of classes. In line  
599 with the hypothesis that format-invariant representations of zero exist on a graded, abstract  
600 neural number line, we expected the discriminability to improve as the numerical distance from

601 zero increased (Figure 3B). To statistically test whether this was the case, we performed one-  
602 tailed, paired comparisons between the discriminability of successive numbers with zero (e.g., by  
603 comparing 0-2 vs. 0-1, 0-3 vs. 0-2, etc.; Figure 3B).

604

605 For all decoding analyses, we utilized multiclass or binary LDA decoders in conjunction with the  
606 MVPA-light toolbox (Treder, 2020) integrated with FieldTrip. To improve the robustness of the  
607 classifier, we applied L1-regularization to the covariance matrix, and the shrinkage parameter  
608 was automatically determined using the Ledoit-Wolf formula within each training fold (Ledoit &  
609 Wolf, 2004).

610

### 611 Source Reconstruction

612 Both FieldTrip's template single shell head model and its standard volumetric grid (8mm  
613 resolution) were warped to participants' individual fiducial points, generating a subject-specific  
614 forward model aligned in MNI space. Source reconstruction was performed using a linearly  
615 constrained minimum variance (lcmv) beamformer (Van Veen et al., 1997) which applies spatial  
616 filters to the MEG data to generate source-level time courses. To reduce the impact of noise on  
617 the source estimates, we used a regularisation parameter of  $\lambda = 5\%$ . For each task, spatial  
618 filters were calculated by combining the leadfield matrix with the data covariance matrix across  
619 all numerosities and the timepoints coinciding with the stable cluster of significantly above-  
620 chance decoding in the zero vs. non-zero cross-task classifier (100 – 450ms). These spatial filters  
621 were then applied to zero trials and non-zero trials separately, generating reconstructed maps of  
622 source activity for these two trial types. We contrasted the broadband source power of zero >

623 non-zero trials in a mass-univariate procedure across subjects for each task separately (Figure  
624 4A) with an alpha parameter of  $p < .05$ , corrected for multiple comparisons. For binary LDA  
625 classifiers, this is equivalent to localising the classifier weights (Haufe et al., 2014), and therefore  
626 gives an indication of which brain regions drove our decoding results. We computed the  
627 conjunction of these two contrasts, revealing the voxels where zero stimuli were dissociable from  
628 other numbers in both symbolic and non-symbolic notations (Figure 4B).

629

630 Multidimensional scaling of source space activity was performed using the same beamforming  
631 parameters to calculate spatial filters over combined non-symbolic and symbolic trials. Using  
632 these filters, virtual channels were created for each source location within the map defined by the  
633 conjunction analysis. The virtual channels were then used to create a cross-task representational  
634 dissimilarity matrix in the same manner as described for the cross-task RSA sensor-level  
635 analysis. This was then submitted to MATLAB's `cmdscale` function for multidimensional  
636 scaling.

637

### 638 Statistical Inference

639 Across sensor and source level analyses, cluster-based permutation testing was used to  
640 statistically test hypotheses and correct for multiple comparisons (Maris & Oostenveld, 2007).  
641 For all analyses (decoding, RSA, and source-level contrasts), 1000 permutations were used with  
642 cluster-forming alpha parameter of .05 and a significance threshold of .05. It is important to  
643 emphasize that this cluster-based permutation testing approach does not provide information  
644 about when neural representations emerge. This limitation arises because the statistical inference

645 process does not focus on individual time points; instead, it relies on cluster-level statistics that  
646 encompass multiple time points (Sassenhagen & Draschkow, 2019).

647



## 648 **References**

- 649 Arsalidou, M., & Taylor, M. J. (2011). Is  $2+2=4$ ? Meta-analyses of brain areas needed for  
650 numbers and calculations. *NeuroImage*, *54*(3), 2382–2393.  
651 <https://doi.org/10.1016/j.neuroimage.2010.10.009>
- 652 Bialystok, E., & Codd, J. (2000). Representing quantity beyond whole numbers: Some, none,  
653 and part. *Canadian Journal of Experimental Psychology*, *54*(2), 117–128.  
654 <https://doi.org/10.1037/h0087334>
- 655 Borghesani, V., de Hevia, M. D., Viarouge, A., Pinheiro-Chagas, P., Eger, E., & Piazza, M.  
656 (2019). Processing number and length in the parietal cortex: Sharing resources, not a  
657 common code. *Cortex*, *114*, 17–27. <https://doi.org/10.1016/j.cortex.2018.07.017>
- 658 Brysbaert, M. (1995). Arabic Number Reading: On the Nature of the Numerical Scale and the  
659 Origin of Phonological Recoding. *Journal of Experimental Psychology: General*, *124*(4),  
660 434–452. <https://doi.org/10.1037/0096-3445.124.4.434>
- 661 Butterworth, B. (1999). *The mathematical brain* (1. publ). Macmillan.
- 662 Cohen Kadosh, R., Cohen Kadosh, K., Kaas, A., Henik, A., & Goebel, R. (2007). Notation-  
663 Dependent and -Independent Representations of Numbers in the Parietal Lobes. *Neuron*,  
664 *53*(2), 307–314. <https://doi.org/10.1016/j.neuron.2006.12.025>
- 665 Cohen Kadosh, R., & Walsh, V. (2009). Numerical representation in the parietal lobes: Abstract  
666 or not abstract? *Behavioral and Brain Sciences*, *32*(3–4), 313–328.  
667 <https://doi.org/10.1017/S0140525X09990938>
- 668 Coldren, J. T., & Haaf, R. A. (2000). Asymmetries in infants' attention to the presence or  
669 absence of features. *Journal of Genetic Psychology*, *161*(4), 420–434.  
670 <https://doi.org/10.1080/00221320009596722>

- 671 Damarla, S. R., Cherkassky, V. L., & Just, M. A. (2016). Modality-independent representations  
672 of small quantities based on brain activation patterns. *Human Brain Mapping*, 37(4),  
673 1296–1307. <https://doi.org/10.1002/hbm.23102>
- 674 Dehaene, S., & Cohen, L. (2007). Cultural Recycling of Cortical Maps. *Neuron*, 56(2), 384–398.  
675 <https://doi.org/10.1016/j.neuron.2007.10.004>
- 676 Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in  
677 the animal and human brain. *Trends in Neurosciences*, 21(8), 355–361.  
678 [https://doi.org/10.1016/S0166-2236\(98\)01263-6](https://doi.org/10.1016/S0166-2236(98)01263-6)
- 679 Eger, E., Michel, V., Thirion, B., Amadon, A., Dehaene, S., & Kleinschmidt, A. (2009).  
680 Deciphering Cortical Number Coding from Human Brain Activity Patterns. *Current*  
681 *Biology*, 19(19), 1608–1615. <https://doi.org/10.1016/j.cub.2009.08.047>
- 682 Eger, E., Sterzer, P., Russ, M. O., Giraud, A.-L., & Kleinschmidt, A. (2003). A supramodal  
683 number representation in human intraparietal cortex. *Neuron*, 37(4), 719–725.  
684 [https://doi.org/10.1016/s0896-6273\(03\)00036-9](https://doi.org/10.1016/s0896-6273(03)00036-9)
- 685 Goh, R. Z., Phillips, I. B., & Firestone, C. (2023). The perception of silence. *Proceedings of the*  
686 *National Academy of Sciences*, 120(29), e2301463120.  
687 <https://doi.org/10.1073/pnas.2301463120>
- 688 Harvey, B. M., & Dumoulin, S. O. (2017). A network of topographic numerosity maps in human  
689 association cortex. *Nature Human Behaviour*, 1(2). [https://doi.org/10.1038/s41562-016-](https://doi.org/10.1038/s41562-016-0036-0)  
690 0036
- 691 Harvey, B. M., Klein, B. P., Petridou, N., & Dumoulin, S. O. (2013). Topographic  
692 Representation of Numerosity in the Human Parietal Cortex. *Science*, 341(6150), 1123–  
693 1126. <https://doi.org/10.1126/science.1240405>

- 694 Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J. D., Blankertz, B., & Bießmann, F.  
695 (2014). On the interpretation of weight vectors of linear models in multivariate  
696 neuroimaging. *NeuroImage*, *87*, 96–110.  
697 <https://doi.org/10.1016/j.neuroimage.2013.10.067>
- 698 Ifrah, G. (1985). From one to zero: A universal history of numbers. (*No Title*).
- 699 Kaplan, R. (1999). *The Nothing that Is: A Natural History of Zero*. Oxford University Press.
- 700 King, J.-R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: The  
701 temporal generalization method. *Trends in Cognitive Sciences*, *18*(4), 203–210.  
702 <https://doi.org/10.1016/J.TICS.2014.01.002>
- 703 Kirschhock, M. E., Ditz, H. M., & Nieder, A. (2021). Behavioral and Neuronal Representation of  
704 Numerosity Zero in the Crow. *The Journal of Neuroscience*, *41*(22), 4889–4896.  
705 <https://doi.org/10.1523/jneurosci.0090-21.2021>
- 706 Krajcsi, A., Kojouharova, P., & Lengyel, G. (2021). Development of Preschoolers’  
707 Understanding of Zero. *Frontiers in Psychology*, *0*, 3169.  
708 <https://doi.org/10.3389/FPSYG.2021.583734>
- 709 Kriegeskorte, N., & Diedrichsen, J. (2019). Peeling the Onion of Brain Representations. *Annual*  
710 *Review of Neuroscience*, *42*(1), 407–432. [https://doi.org/10.1146/annurev-neuro-080317-](https://doi.org/10.1146/annurev-neuro-080317-061906)  
711 [061906](https://doi.org/10.1146/annurev-neuro-080317-061906)
- 712 Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: Integrating cognition,  
713 computation, and the brain. *Trends in Cognitive Sciences*, *17*(8), 401–412.  
714 <https://doi.org/10.1016/J.TICS.2013.06.007>

- 715 Kutter, E. F., Bostroem, J., Elger, C. E., Mormann, F., & Nieder, A. (2018). Single Neurons in  
716 the Human Brain Encode Numbers. *Neuron*, *100*(3), 753-761.e4.  
717 <https://doi.org/10.1016/j.neuron.2018.08.036>
- 718 Kutter, E. F., Dehnen, G., Borger, V., Surges, R., Mormann, F., & Nieder, A. (2023). Distinct  
719 neuronal representation of small and large numbers in the human medial temporal lobe.  
720 *Nature Human Behaviour*, 1–10. <https://doi.org/10.1038/s41562-023-01709-3>
- 721 Libertus, M. E., Woldorff, M. G., & Brannon, E. M. (2007). Electrophysiological evidence for  
722 notation independence in numerical processing. *Behavioral and Brain Functions*, *3*.  
723 <https://doi.org/10.1186/1744-9081-3-1>
- 724 Luyckx, F., Nili, H., Spitzer, B., & Summerfield, C. (2019). Neural structure mapping in human  
725 probabilistic reward learning. *eLife*, *8*. <https://doi.org/10.7554/eLife.42816>
- 726 Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data.  
727 *Journal of Neuroscience Methods*, *164*(1), 177–190.  
728 <https://doi.org/10.1016/j.jneumeth.2007.03.024>
- 729 Mazor, M., Friston, K. J., & Fleming, S. M. (2020). Distinct neural contributions to  
730 metacognition for detecting, but not discriminating visual stimuli. *eLife*, *9*, e53900.  
731 <https://doi.org/10.7554/eLife.53900>
- 732 Mazor, M., Moran, R., & Fleming, S. M. (2021). Metacognitive asymmetries in visual  
733 perception. *Neuroscience of Consciousness*, *2021*(1), 1–15.  
734 <https://doi.org/10.1093/nc/niab025>
- 735 Merritt, D. J., & Brannon, E. M. (2013). Nothing to it: Precursors to a zero concept in  
736 preschoolers. *Behavioural Processes*, *93*, 91–97.  
737 <https://doi.org/10.1016/j.beproc.2012.11.001>

- 738 Merten, K., & Nieder, A. (2012). Active encoding of decisions about stimulus absence in primate  
739 prefrontal cortex neurons. *Proceedings of the National Academy of Sciences of the United*  
740 *States of America*, 109(16), 6289–6294. <https://doi.org/10.1073/pnas.1121084109>
- 741 Nieder, A. (2016). Representing Something Out of Nothing: The Dawning of Zero. *Trends in*  
742 *Cognitive Sciences*, 20(11), 830–842. <https://doi.org/10.1016/j.tics.2016.08.008>
- 743 Nieder, A., & Dehaene, S. (2009). Representation of number in the brain. *Annual Review of*  
744 *Neuroscience*, 32, 185–208. <https://doi.org/10.1146/annurev.neuro.051508.135550>
- 745 Nili, H., Wingfield, C., Walther, A., Su, L., Marslen-Wilson, W., & Kriegeskorte, N. (2014). A  
746 Toolbox for Representational Similarity Analysis. *PLoS Computational Biology*, 10(4).  
747 <https://doi.org/10.1371/journal.pcbi.1003553>
- 748 Okuyama, S., Kuki, T., & Mushiake, H. (2015). Representation of the Numerosity ‘zero’ in the  
749 Parietal Cortex of the Monkey. *Scientific Reports*, 5(1), Article 1.  
750 <https://doi.org/10.1038/srep10059>
- 751 Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software  
752 for advanced analysis of MEG, EEG, and invasive electrophysiological data.  
753 *Computational Intelligence and Neuroscience*, 2011.  
754 <https://doi.org/10.1155/2011/156869>
- 755 Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for  
756 approximate numerosity in the human intraparietal sulcus. *Neuron*, 44(3), 547–555.  
757 <https://doi.org/10.1016/j.neuron.2004.10.014>
- 758 Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A Magnitude Code Common to  
759 Numerosities and Number Symbols in Human Intraparietal Cortex. *Neuron*, 53(2), 293–  
760 305. <https://doi.org/10.1016/j.neuron.2006.11.022>

- 761 Ramirez-Cardenas, A., Moskaleva, M., & Nieder, A. (2016). Neuronal Representation of  
762 Numerosity Zero in the Primate Parieto-Frontal Number Network Article Neuronal  
763 Representation of Numerosity Zero in the Primate Parieto-Frontal Number Network.  
764 *Current Biology*, 26, 1285–1294. <https://doi.org/10.1016/j.cub.2016.03.052>
- 765 Sainsbury, R. (1971). The “feature positive effect” and simultaneous discrimination learning.  
766 *Journal of Experimental Child Psychology*, 11(3), 347–356.  
767 [https://doi.org/10.1016/0022-0965\(71\)90039-7](https://doi.org/10.1016/0022-0965(71)90039-7)
- 768 Sassenhagen, J., & Draschkow, D. (2019). Cluster-based permutation tests of MEG/EEG data do  
769 not establish significance of effect latency or location. *Psychophysiology*, 56(6), 1–8.  
770 <https://doi.org/10.1111/psyp.13335>
- 771 Schubert, T. M., Rothlein, D., Brothers, T., Coderre, E. L., Ledoux, K., Gordon, B., &  
772 McCloskey, M. (2020). Lack of awareness despite complex visual processing: Evidence  
773 from event-related potentials in a case of selective metamorphopsia. *Proceedings of the*  
774 *National Academy of Sciences*, 117(27), 16055–16064.  
775 <https://doi.org/10.1073/pnas.2000424117>
- 776 Teichmann, L., Grootswagers, T., Carlson, T., & Rich, A. N. (2018). Decoding Digits and Dice  
777 with Magnetoencephalography: Evidence for a Shared Representation of Magnitude.  
778 *Journal of Cognitive Neuroscience*, 30(7), 999–1010.  
779 [https://doi.org/10.1162/jocn\\_a\\_01257](https://doi.org/10.1162/jocn_a_01257)
- 780 Treder, M. S. (2020). MVPA-Light: A Classification and Regression Toolbox for Multi-  
781 Dimensional Data. *Frontiers in Neuroscience*, 0, 289.  
782 <https://doi.org/10.3389/FNINS.2020.00289>

- 783 Van Veen, B. D., van Drongelen, W., Yuchtman, M., & Suzuki, A. (1997). Localization of brain  
784 electrical activity via linearly constrained minimum variance spatial filtering. *IEEE*  
785 *Transactions on Bio-Medical Engineering*, 44(9), 867–880.  
786 <https://doi.org/10.1109/10.623056>
- 787 Wellman, H. M., & Miller, K. F. (1986). Thinking about nothing: Development of concepts of  
788 zero. *British Journal of Developmental Psychology*, 4(1), 31–42.  
789 <https://doi.org/10.1111/j.2044-835X.1986.tb00995.x>
- 790

## 791 Supplemental Figures

792

793

794

795

796

797

798

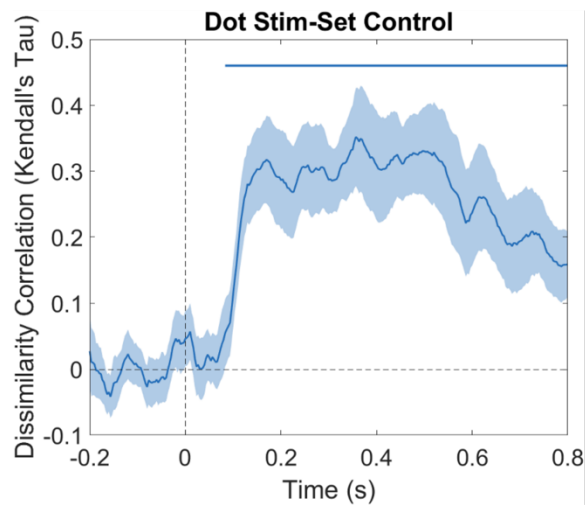
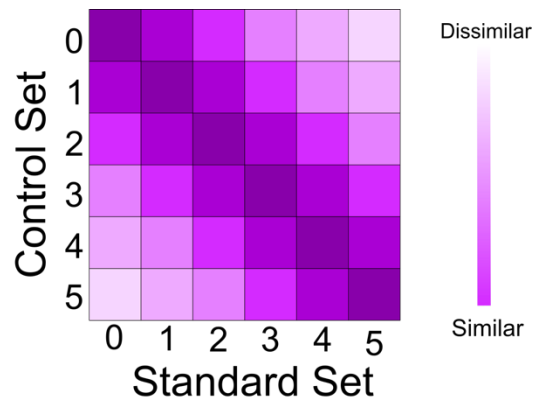
799

800

801

802

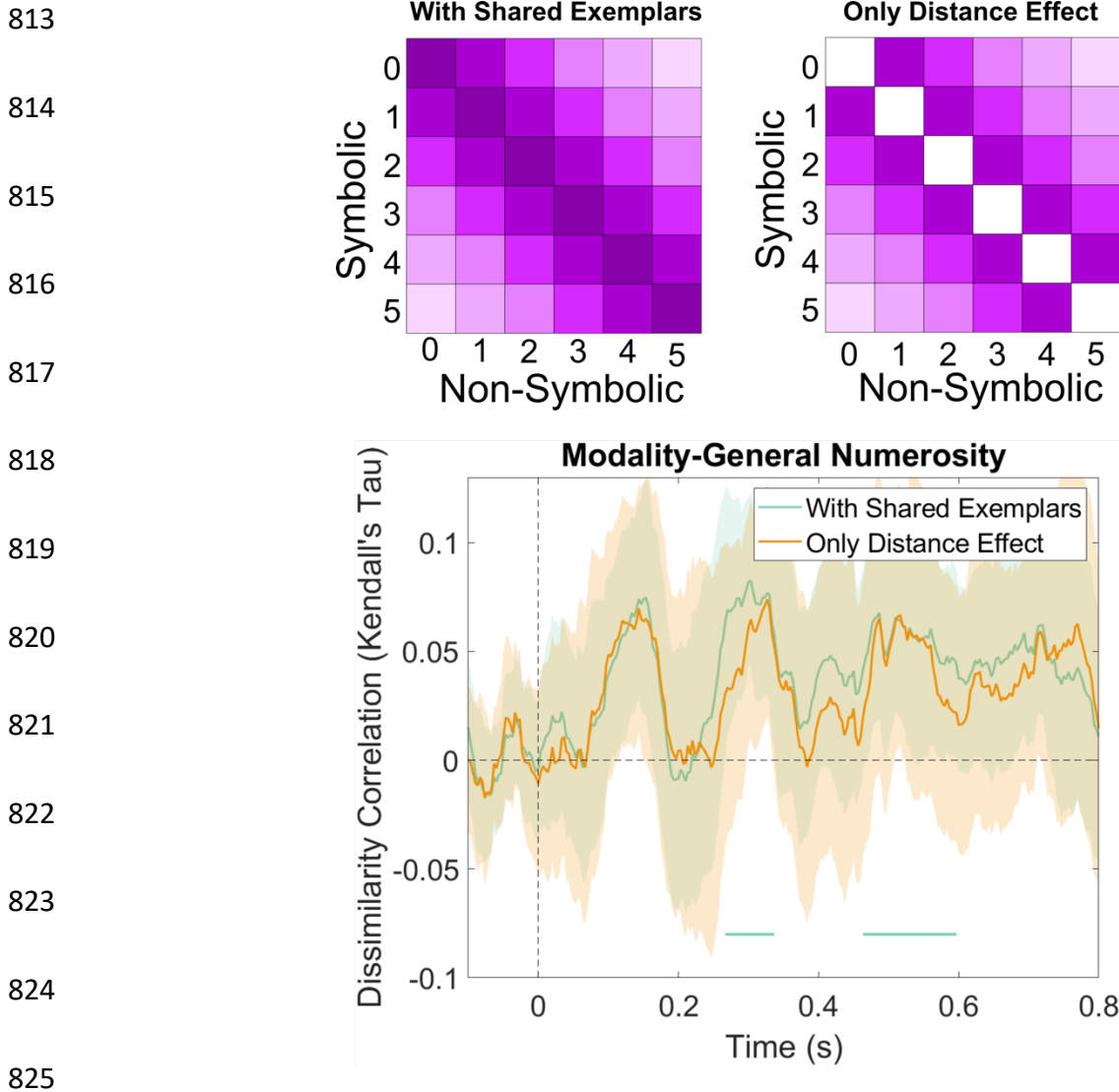
803



804 **Supplemental Figure 1.** *Cross-stimulus set representations of numerosity in non-symbolic*  
805 *stimuli.* A representational dissimilarity matrix (RDM) was constructed that modelled the  
806 distance effect between non-symbolic numerosities across stimulus sets (top). This model tests  
807 whether numerical information is shared across the standard and control set, independently of  
808 their unique physical features. Numerical distance effects could be extracted from the neural data  
809 independently of the stimulus set soon after stimulus presentation and for the remainder of the  
810 epoch (bottom). Horizontal line represents time points where the correlation of the model RDM



811 with the neural data was significantly above zero with an alpha of  $p < .05$ , corrected for multiple  
812 comparisons.



826 **Supplemental Figure 2.** *Cross-Task RSA reveals a format-invariant neural code for number. An*  
827 *RDM modelling numerical information as shared between numerical format successfully*  
828 *predicted our neural data at two different timepoints. Removing the diagonal from this RDM*  
829 *removes the shared exemplars from the model (empty sets and zero, one dot and symbolic one,*  
830 *etc.) providing a strong test of the hypothesis that abstract numerical information also exhibits a*  
831 *distance effect. This model showed a similar pattern of prediction to the full model with shared*  
832 *exemplars yet failed to reach statistical significance. Broad confidence intervals, represented by*  
833 *the shaded area, suggest this may be an issue of limited statistical power. Horizontal lines*

834 indicate clusters where the model RDM correlated with the neural data significantly more than

835 zero,  $p < .05$ , corrected for multiple comparisons. Shaded areas represent 95% confidence

836 intervals.

837

838

839

840

841

842

843

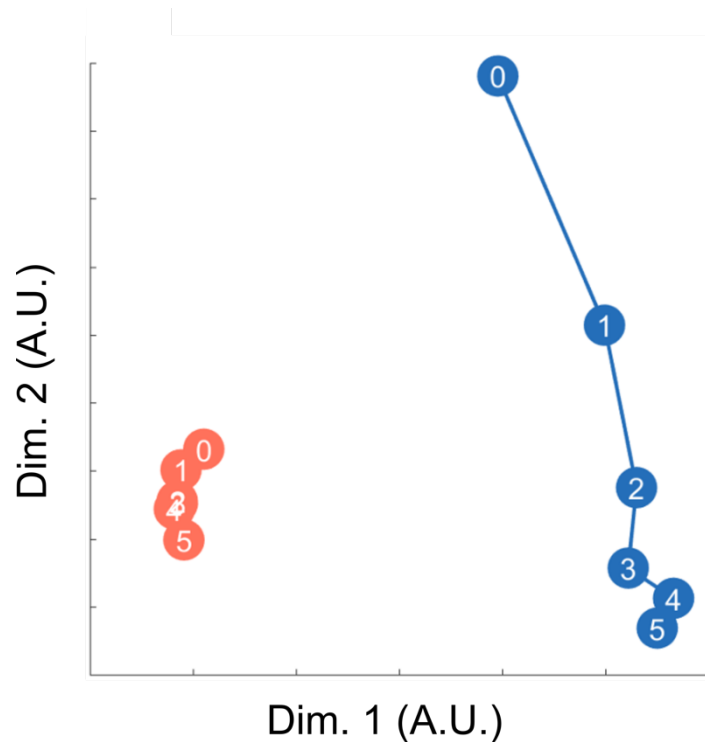
844

845

846

847

848



849

**Supplemental Figure 3.** *Multidimensional scaling of numbers across format.* Performing

850

multidimensional scaling on numerosity representations in a shared space revealed alignment

851

along an axis defining numerical magnitude (dimension two). This illustrates the cross-task

852

distance effect, where empty sets (blue zero) are represented more closely to symbolic one (red

853

one) than symbolic five (red five), and vice versa. Dimension one discriminates between the two

854

tasks.