Supporting information for

Neural signals predict information sharing across cultures

Supplementary Method

Participants and task

Inclusion criteria were fluency in English, right-handedness, and being a student or recent graduate of the university. Exclusion criteria were presence of any irremovable nonferrous metallic objects (i.e., implanted medical devices), pregnancy or breastfeeding, history of substance abuse, major mental health diagnosis, or recent psychotropic medication use.

With respect to gender, participants identified as the following: 49.0% men, 49.0% women, 1% non-binary, and 1% gender fluid. For race and ethnicity, participants identified themselves as the following: 54.1% White, 13.5% Hispanic or Latina/Latino/Latinx, 8.3% Black or African American, 5.2% East Asian, 3.1% Southeast Asian, 3.1% more than one race, 1% American Indian or Alaskan Native, 10.4% as another race ("other"), and 1% preferred not to say.

The majority of participants (40 out of 50 Dutch participants and all 44 US participants) read and rated 24 articles, pseudo-randomly chosen from the pool of 96, in a single run. Another 10 Dutch participants read and rated all 96 articles in 3 runs (i.e., 32 articles per run). In accordance with our pre-registration, we combined data from the short and long versions of the task.

News articles and population sharing measure

Articles about specific time-sensitive issues or events and articles with highly US-centric topics were excluded since scanning took place over the course of multiple months and in both US and Dutch samples. We chose articles of varying population sharing across the topics of health (n = 48; range of share counts of the news articles on Facebook pages: 58–22039) and climate change (n = 48; range of share counts: 44–25091), while making sure the length for each abstract was suitable for presentation during fMRI (word count mean = 31.9, s.d. = 5.5, range = 21–44). The sharing metric was accessed in May 2021 via CrowdTangle. (A secondary measure – click-through counts via bit.ly short links – yielded similar findings.)

The stimulus set in Scholz et al. (ref. 5) were 80 news abstracts of New York Times articles published during 2012–2013. We note that in the original Scholz et al. (ref. 5) paper, population sharing was operationalized differently as the number of times each article was shared within the New York Times website. Since the New York Times does not provide this metric anymore through its application programming interface (API), in order to facilitate comparison with the current results, we collected the Facebook share metric via CrowdTangle of these articles (mean = 690, s.d. = 1550, range = 2–8823) for re-analysis and building prediction models.

Neuroimaging data collection and extraction

At the US site, data was collected on a Siemens 3-Tesla scanner with a 64 channel head coil. At the Dutch site, data was collected on a Philips 3-Tesla scanner with a 32 channel head coil. The scanning protocol was the same at both sites: We started by collecting a high-resolution 3D T1-weighted scan for anatomical reference (US site: FOV = 192 × 256 × 224 mm; TR = 2.15s; TE = 0.00388s; FA = 8°; voxel size = 1 × 1 × 1 mm; Dutch site: FOV = 242 × 242 × 220 mm; TR = 0.0082s; TE = 0.0037s; FA = 8°; voxel size: 0.945 × 0.945 × 1 mm). While participants read news abstracts, functional scans were collected using a multiband echo-planar imaging (EPI) technique (multiband factor = 4; FOV = 240 × 240 × 118.8 mm; TR = 0.55s; TE = 30ms; FA = 55°; voxel size = 3 × 3 × 3 mm; slice gap = 0.3mm).

As detailed in the preregistration, MRI data were preprocessed using fMRIPrep 20.0.6. Briefly, anatomical images were segmented and normalized to MNI space using FreeSurfer; functional images were susceptibility distortion corrected using the ANTs symmetric normalization (SyN) technique, realigned, and coregistered to the normalized anatomical images. Data quality check (no signal outliers of +/- 3SD and less than 10% of volumes with framewise displacement > .75mm) was then performed, resulting in no exclusion of participants. Preprocessed functional data were smoothed using a 6mm full-width at half maximum smoothing kernel.

First level general linear model (GLM) was created for each participant in order to extract single-trial beta images during exposure to news abstract (i.e., activity during the bolded frame in Figure 1B). Regressors of no interest include: euclidean translation, euclidean rotation, translation derivatives, rotation derivatives, dummy-coded motion artifact outliers (framewise displacement > .75mm), and cerebrospinal fluid average signals.

For the Scholz et al. (ref. 5) dataset, we applied the same preprocessing pipeline, then extracted single-trial beta images using the same GLM parameters.

Hypothesis testing

Neurosynth association maps, which indicate how much more likely activation is reported in a region given a term is present in the neuroscientific literature, were used to create binarized regions of interest (ROIs) by thresholding at $p < .01$ (FDR corrected) with the following terms – self-related ROI: “self referential”; social-related ROI: “mentalizing”; value-related ROI: “value”. To increase dissociability, we first removed all overlapping voxels from each mask (i.e., retaining voxels that remain only appear in one of the three masks). These ROI masks are available in the preregistration. Voxel signals were then averaged within each ROI for each single-trial beta image and then z-scored within the participant.

We note that in the original Scholz et al. (ref. 5) study, self-, social- and value-related signals were extracted from functional ROIs defined in extant literature. For this paper, we re-analysed the data using the Neurosynth ROIs instead.

Prediction model training

The training dataset came from the neuroimaging data of Scholz et al. (ref. 5). Same as Scholz et al. (ref. 5), some trials that required participants to rate not reading intention but sharing intention after reading the news abstracts were excluded for model training.

The ROI-based model features average signals within binarized masks derived from nine Neurosynth association maps (thresholded at $p < .01$, FDR corrected), thought to be theoretically relevant to information sharing: perceptual (“visual”, “language”), affective (“arousal”, “emotion”), executive (“attention”, “memory”) and higher-order functions (“self referential”, “mentalizing”, “value”). No removal of overlapping voxels was done. It should be noted that this set of ROIs was not meant to be exhaustive, but we reasoned that together they should offer sufficiently informative neural signals.

The voxel-based model consists of voxels identified to be related to preference in the Scholz et al. (ref. 5) data. Specifically, we estimated a GLM with self-report reading intention as parametric modulation (using the same set of regressors of no interest as above), and chose voxels that significantly tracked reading intention at $p < .005$ (uncorrected) level (5,137 voxels in total). The voxel clusters appear mainly in medial frontal cortex, left inferior frontal gyrus, and visual cortex.

We used ridge regression for building both models (i.e., the ROI-based model had 9 features, while the voxel-based model had 5,137 features). The L2 penalty was determined after a search between 10^{-7} and 10^{7} under 8-fold cross-validation. Afterwards, the median coefficients across the eight cross-validation folds with the chosen L2 penalty were selected as feature weights.

Finally, we combined ROI-based and voxel-based prediction scores, as well as self-report ratings of reading intention to estimate an ensemble model with ridge regression using the same procedure as above. That is, we assigned weights to ROI-based prediction scores, voxel-based prediction scores, as well as self-report ratings so that we could combine these three features into an ensemble score.

Study preregistration

Prior to neuroimaging data collection, we preregistered (DOI: 10.17605/OSF.IO/JCVZ7) task stimuli, ROI masks used in hypothesis testing and prediction model training, feature weights in the ROI-based, voxel-based and ensemble models, as well as the analysis plan.