Origins of Shared conceptual Reprsentation in a ViSUAL COmmunicative System

ANONYMOUS AUTHOR 1\*1,2 and ANONYMOUS AUTHOR 21

\*Corresponding Author: name@domain.com

1This Department, University X, City, Country

2That Department, University Y, City, Country

To communicate, two people must have someoverlapping representation of meaning. Yet, there is ample evidence that cultures can vary in ways both dramatic and subtle in their semantic structure (). Little is known, however, about the source of this variability: Why might some cultures have more similar semantics to each other and others more different?

We explore this question using a dataset of 50 million drawings of common objects (e.g., “bread”) from 15 million participants worldwide (QuickDraw: quickdraw.withgoogle.com/data). Like spoken language, visual drawings are a communicative representational system. But, unlike language, drawings are analog rather than symbolic; the sign (the drawing) could be iconically related to the cognitive representation. This analogic representation allows us to more directly observe representations of meaning in a communicative system, as well as their variability cross-culturally.

In our study, we sought to quantify cross-cultural drawing similarity and examine predictors of variability. As one measure of similarity, we adopted a measure from the object recognition literature, Haussdorf Distance (HD), which quantifies image similarity as the minimum Euclidean distance between two sets of points (Huttenlocher, Klanderman, & Rucklidge, 1993). As a second measure, we took the prediction weights (layer FC2) for each of our drawings from a neural net model trained on ImageNet (Deng, et al., 2009), and then calculated the cosine distance (CD) on weights between pairs of images.

We selected two items – “bread” and “tree” – and sampled 1500 pairs of images for each item across the 72 countries in the dataset. For each pair, we calculated both HD and CD (HD: *M* = 93.87, *SD* = .39; CD: *M* = 0.99984, *SD* = .0005). These two measures were uncorrelated (*r* = -.02, *p* = .29), suggesting that they captured different aspects of visual similarity.

We then validated our similarity measures using human judgements. We selected 20 pairs from each HD decile for each item (Fig. 1), and asked participants to rate the similarity of the objects in the drawings using a 1 (almost identical) to 7 (very different) Likert scale. Each participant (*N* = 100) rated 50 pairs of a single item.

Human judgements of similarity were highly correlated with HD (*r* = .29, *p* < .0001; Fig. 2) as well as CD (*r* = -0.20, *p* < .0001). In a mixed effect model with both HD and CD as fixed effects, both measures of similarity were reliable predictors of human similarity judgements (HD: *B* = .35; *t* = 12.39; CD: *B* = -.26; *t* = -9.95).

Next, we asked if we could predict cross-country variability in similarity ratings using three measures: physical distance, linguistic distance, and socio-political distance. We quantified physical distance as the distance in meters between the centroid of each pair of countries (GNS Database, 2012). We quantified linguistic distance as the overlap in ASJP vocabulary (Bakker, et al., 2009). Finally, we quantified socio-political distance by taking the count (log) of the number of socio-political events between country pairs (e.g., “making a statement;” Boschee, et al., 2015), and normalizing by the mean population of each country. In mixed effect models, both physical distance (*B* = 0.14, *t* = 2.21) and socio-political distance (*B* = -0.17, *t* = 2.30) were reliable predictors of ratings, but linguistic distance was not (*B* = 0.08, *t* = 1.13). In an additive model with all three predictors, only socio-political distance was a reliable predictor of ratings (*B* = -.21; *t* =-2.29): Countries that have less socio-political distance (more events) have more similar drawings.

These data suggest that visual communication systems reflect culturally-specific representations of meaning, and that these similarities derive from physical and social proximity.

pair_example.pdf

Fig. 1: Sample airs of “bread” drawings used in the human similarity judgement experiment. Pairs are shown from Haussdorf decile 1 (top), 5 (middle), and 10 (bottom)

experiments/conceptviz_1/analyses/distance_decile.pdf

Fig. 2: Similarity ratings from the human similarity judgement experiment, as a function of Hausdorff similarity distances. Ranges are bootstrapped 95% confidence intervals.

Bakker, D., Müller, A., Velupillai, V., Wichmann, S., Brown, C. H., Brown, P., ... & Holman, E. W. (2009). Adding typology to lexicostatistics: A combined approach to language classification. *Linguistic Typology*, *13*(1), 169-181.

Boschee, E., Lautenschlager, J., O'Brien, S., Shellman, S. Starz, J. Ward, M. (2015). "ICEWS Coded Event Data", doi:10.7910/DVN/28075, Harvard Dataverse, V20

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 248–255).

Huttenlocher, D. P., Klanderman, G. A., & Rucklidge, W. J. (1993). Comparing images using the Hausdorff distance. *IEEE Transactions on pattern analysis and machine intelligence*, *15*(9), 850-863.

GNS database, <http://geonames.nga.mil/gns/html/>