Tracing the Emergence of Context-Specific Conceptual Representations in the Brain

Sarah Solomon & Sharon Thompson-Schill

Department of Psychology, University of Pennsylvania

HYPOTHESIS

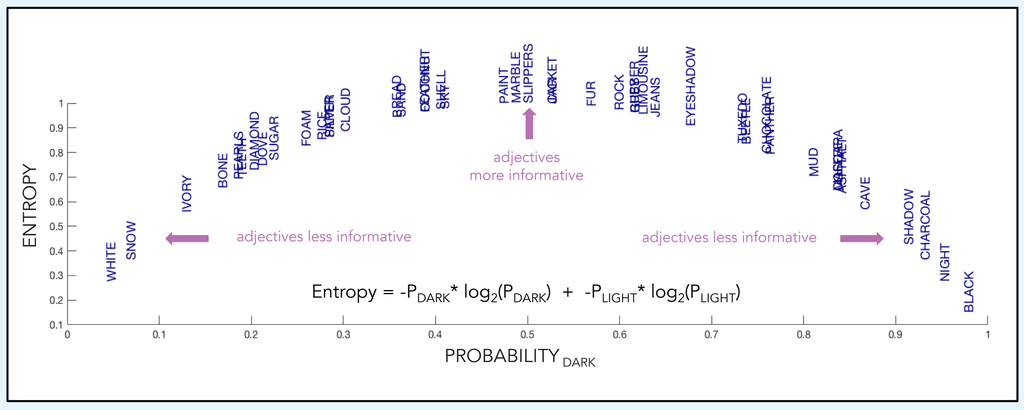
The **informational content of individual words is flexible** across contexts[1,2], requiring the cognitive system to generate context-specific representations during language comprehension. For example:

"DARK" has a large effect in "DARK PAINT", but not in "DARK CHARCOAL"

Our goal is to predict the context-specific representations in a simple model, using the test case of **adjective-noun combinations**. We use the adjectives "LIGHT" and "DARK" to modulate brightness levels of 45 concepts (e.g., SNOW, FUR, CHARCOAL), and try to predict the amount of brightness modulation in each case.

We calculated **brightness probability** for each concept, and transformed these probability values to **brightness entropy**, which is used in information theory to capture the informativity of a signal[3].

We **predicted that the entropy** (i.e., informativity) of an adjective in a given adjective-noun pair **should predict the amount of brightness modulation** that occurs during comprehension.



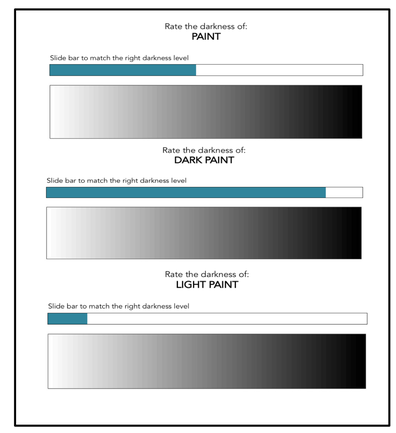
BEHAVIORAL RATINGS

1. Probability
   Participants on AMT (*N*=68) reported the extent to which each unmodified concept is "typically dark" on a scale from 1-5. These values were then scaled between 0 and 1 to create P_{DARK}. $P_{LIGHT} = 1 - P_{DARK}$.

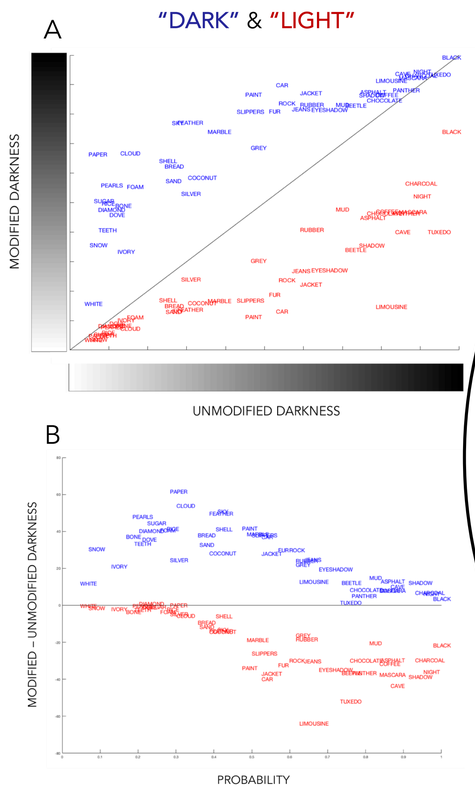
2. Entropy
   P_{DARK} and P_{LIGHT} were transformed into Entropy using the standard equation from information theory:

   $$Entropy = -P_{DARK} * \log_2(P_{DARK}) + -P_{LIGHT} * \log_2(P_{LIGHT})$$

3. Darkness Values
   One group of participants on AMT (*N*=118) made explicit darkness judgments on the 45 unmodified concepts, and another group (*N*=235) made explicit darkness judgments on the 90 combinations.



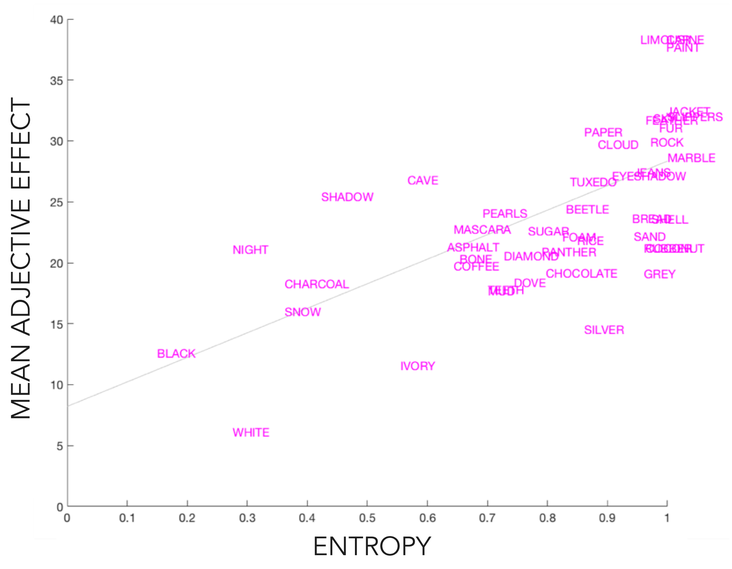
BEHAVIORAL RESULTS



(A) Unmodified vs. Modified brightness. Distance from line signifies amount of property modulation.

(B) Brightness modulation (MOD – UNMOD) as a function of P_{DARK}. Positive values mean the combination became darker.

INFORMATIVITY PREDICTS ADJECTIVE EFFECTS



The modulation caused by "DARK" and "LIGHT" was averaged for each concept, resulting in a mean adjective effect. This represents the extent to which the adjectives modulate the brightness of each concept. This measure is significantly predicted by Entropy, suggesting that the effects of adjectives are influenced by the concept-specific adjective informativity.

NEUROIMAGING RESULTS

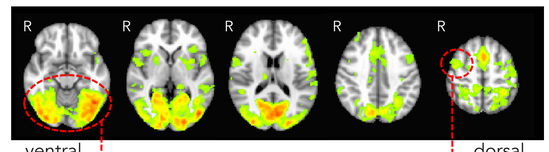
During comprehension of adjective-noun phrases, the brain must represent the individual constituents (e.g., "dark", "paint") and also must generate a context-specific representation.

What neural regions contain brightness representations that are modulated by the informativity of an adjective-noun phrase?

We collected fMRI data while participants (*N*=11) were presented with the same concepts and combinations and performed an orthogonal color judgment task. We isolated regions that were sensitive to the brightness of the unmodified concepts.

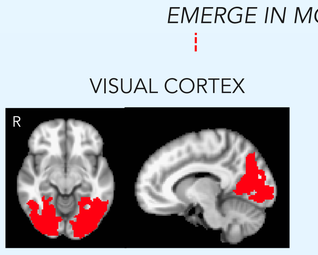
REGIONS SENSITIVE TO CONCEPTUAL BRIGHTNESS

We ran a GLM on the unmodified runs using $UNMOD_{DARK}$ as a parametric covariate. Voxels sensitive to conceptual brightness are found across the brain. We divided these voxels into ROIs, including one in visual cortex and one centered on the right middle frontal gyrus.

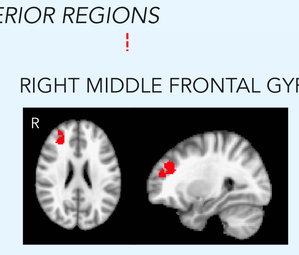


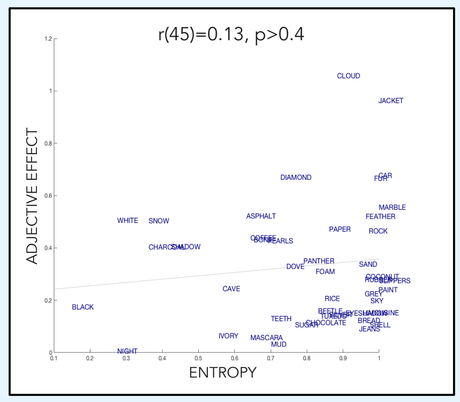
EFFECTS OF INFORMATIVITY ON REPRESENTATIONS OF BRIGHTNESS EMERGE IN MORE ANTERIOR REGIONS

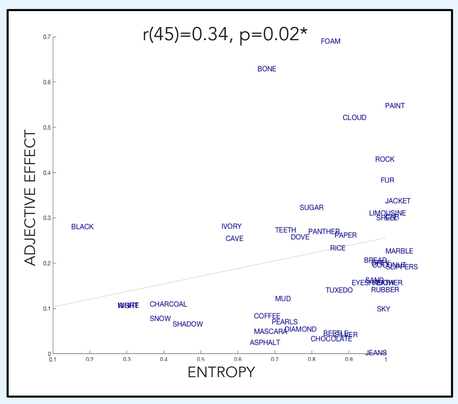
VISUAL CORTEX



RIGHT MIDDLE FRONTAL GYRUS







In each ROI, we averaged the MOD_{DARK} responses and MOD_{LIGHT} responses to each concept across subjects. The adjective effect for each concept is the absolute difference between the MOD_{DARK} and MOD_{LIGHT} responses : this is the amount of neural brightness modulation. Regions in which entropy correlates with adjective effects are interpreted to have representations of conceptual brightness that are influenced by adjective informativity.
Entropy predicts adjective effects in frontal cortex (*p*=0.02), but not visual cortex.

CONCLUSIONS

① Using explicit judgments of darkness, we find that the amount of brightness modulation in "light" and "dark" combinations is predicted by entropy, a measure from information theory that captures informativity.

② Using fMRI, we find that not all regions that represent brightness do so in a context-specific way: effects of adjective-informativity are found in frontal cortex, but not visual cortex.

③ These results begin to reveal the pathways by which conceptual information is transformed into context-specific representations.

Download me!

ACKNOWLEDGEMENTS

This research was supported by Grant R01 DC015359 awarded to STS and an NSF Graduate Research Fellowship awarded to SHS.

1. Yee & Thompson-Schill. (2016). Psychonomic Bulletin and Rev.
2. Halff, H. M., Ortony, A., & Anderson, R. C. (1976). Memory & Cog.
3. Shannon. (1948). Mathematical Reviews.