

Exploring the Role of Language in Two Systems for Categorization

Kayleigh Ryherd, PhD
University of Connecticut, 2019

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Exploring the Role of Language in Two Systems for Categorization

Kayleigh Ryherd

B.A., The George Washington University, 2014
M.S., University of Connecticut, 2016

A Dissertation
Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy
at the
University of Connecticut
2019

Copyright by
Kayleigh Ryherd

2019

iii

APPROVAL PAGE

Doctor of Philosophy Dissertation

Exploring the Role of Language in Two Systems for Categorization
Kayleigh Ryherd, M.S.

Major Advisor: _____
Nicole Landi

Associate Advisor: _____
Letitia Naigles

Associate Advisor: _____
James Magnuson

Associate Advisor: _____
Eiling Yee

Associate Advisor: _____
Clinton Johns

2019

Acknowledgments

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Contents

1	General Introduction	1
1.1	Dual-systems model for category learning	1
1.2	Vocabulary/labels and category learning	1
1.3	Executive function and category learning	1
2	Experiment 1	2
2.1	Methods	2
2.1.1	Participants	2
2.1.2	Category Learning Task	2
2.1.3	Behavioral Measures	2
2.2	Procedure	2
2.3	Results	2
2.4	Discussion	2
3	Experiment 2	3
3.1	Methods	3
3.1.1	Participants	3
3.1.2	Category Learning Tasks	3
3.1.3	Behavioral Measures	3
3.2	Procedure	3
3.3	Results	3
3.4	Discussion	3
4	General Discussion	4
5	Appendix A: Statistical Density Calculations	5
5.1	Statistical Density Formulae	5
5.2	Statistical Density Calculations – Sparse	6
5.3	Statistical Density Calculations – Dense	7
	References	9

1 General Introduction

Blablabla said Nobody (? , ?).

1.1 Dual-systems model for category learning

1.2 Vocabulary/labels and category learning

1.3 Executive function and category learning

2 Experiment 1

2.1 Methods

2.1.1 Participants

2.1.2 Category Learning Task

2.1.3 Behavioral Measures

2.2 Procedure

2.3 Results

2.4 Discussion

3 Experiment 2

3.1 Methods

3.1.1 Participants

3.1.2 Category Learning Tasks

3.1.3 Behavioral Measures

3.2 Procedure

3.3 Results

3.4 Discussion

4 General Discussion

5 Appendix A: Statistical Density Calculations

5.1 Statistical Density Formulae

Statistical density is the method that Sloutsky and colleagues use to define categories (Sloutsky, 2010). Dense categories have multiple intercorrelated features, while sparse categories have few relevant features. Statistical density can vary between 0 and 1. Higher values (closer to 1) are dense, while lower values (closer to 0) are sparse. We calculate statistical density (D) with the following formula, where H_{within} is the entropy within the category and $H_{between}$ is the entropy between the category and contrasting categories.

$$D = 1 - \frac{H_{within}}{H_{between}}$$

To find total entropy(H), we sum entropy due to varying dimension and entropy due to varying relations among dimensions.

$$H = H^{dim} + H^{rel}$$

This equation is the same whether you are calculating within-category entropy or between-category entropy. To find entropy due to dimensions, you use the following formulas, where M is the total number of varying dimensions, w_i is the attentional weight of a particular dimension (assumed to be 1), and p_j is the probability of value j on dimension i .

$$H_{within}^{dim} = \sum_{i=1}^M w_i \left[\sum_{j=0,1} within(p_j \log_2 p_j) \right]$$

$$H_{between}^{dim} = \sum_{i=1}^M w_i \left[\sum_{j=0,1} between(p_j \log_2 p_j) \right]$$

To find entropy due to relations, you use a similar set of formulas, where O is the total number of possible dyadic relations among the varying dimensions, w_k is the attentional weight of a relation (assumed to be 0.5), and p_{mn} is the probability of the co-occurrence of values m and n on dimension k .

$$H_{within}^{rel} = - \sum_{k=1}^O w_k \left[\sum_{\substack{m=0,1 \\ n=0,1}} within(p_{mn} \log_2 p_{mn}) \right]$$

$$H_{between}^{rel} = - \sum_{k=1}^O w_k \left[\sum_{\substack{m=0,1 \\ n=0,1}} between(p_{mn} \log_2 p_{mn}) \right]$$

All categories have 7 dimensions. For dense categories, 6 of these dimensions are correlated. The seventh dimension is allowed to vary randomly. For sparse categories, 6 of the dimensions vary randomly. The seventh dimension is category-relevant and defines the category. All dimensions have two levels (e.g., for hair shape in aliens – curly and straight).

5.2 Statistical Density Calculations – Sparse

First, we calculate the entropy due to dimensions. We have 7 dimensions, so $M = 7$. Between categories (i.e., across all categories), each level of each dimension has a 0.5 probability of being present.

$$\begin{aligned} H_{between}^{dim} &= -7 * 1(2 * 0.5 \log_2 0.5) \\ H_{between}^{dim} &= -7 \log_2 0.5 \\ H_{between}^{dim} &= 7 \end{aligned}$$

Within categories, the relevant dimension does not vary – thus it does not contribute to the entropy. Its value goes to zero, leading to the following calculations.

$$\begin{aligned} H_{within}^{dim} &= -6 * 1(2 * 0.5 \log_2 0.5) \\ H_{within}^{dim} &= -6 \log_2 0.5 \\ H_{within}^{dim} &= 6 \end{aligned}$$

To find the entropy due to relations, we start by calculating O .

$$\begin{aligned} O &= \frac{M!}{(M-2)! * 2!} \\ O &= 21 \end{aligned}$$

Between categories, all dyadic relations have the same probability of co-occurrence (0.25). For each relation between dimensions, there are 4 possible combinations of the levels of those dimensions. They’re all equally probable. Recall that for relations, we use an attentional weight of 0.5. So, we end up with the following.

$$\begin{aligned} H_{between}^{rel} &= -21 * 0.5(4 * 0.25 \log_2 0.25) \\ H_{between}^{rel} &= -10.5 \log_2 0.25 \\ H_{between}^{rel} &= 21 \end{aligned}$$

Within the target category, 15 of the dyadic relationships don’t include the relevant feature. Thus, their probability of co-occurrence is .25. For 6

of the dyadic relations (any including the relevant feature), there is perfect co-occurrence: probability is either 0 or 1. This makes these terms go to zero, because $\log_2 1 = 0$, and anything multiplied by zero is zero.

$$\begin{aligned} H_{within}^{rel} &= -15 * 0.5(4 * 0.25 \log_2 0.25) \\ H_{within}^{rel} &= -7.5 \log_2 0.25 \\ H_{within}^{rel} &= 15 \end{aligned}$$

Now, we use these calculated values to find entropy between and within categories.

$$\begin{aligned} H_{within} &= 6 + 15 \\ H_{within} &= 21 \\ H_{between} &= 7 + 21 \\ H_{between} &= 28 \end{aligned}$$

Finally, we use the within- and between-category entropy to calculate the density.

$$\begin{aligned} D &= 1 - \frac{21}{28} \\ D &= 0.25 \end{aligned}$$

5.3 Statistical Density Calculations – Dense

The between category entropy for dense categories is the same as for sparse categories. $H_{between} = 28$

Next, we will consider within-category entropy due to dimensions. Six of the seven dimensions do not vary, so they do not contribute to the entropy. Their value goes to zero.

$$\begin{aligned} H_{within}^{dim} &= -1 * 1(2 * 0.5 \log_2 0.5) \\ H_{within}^{dim} &= -\log_2 0.5 \\ H_{within}^{dim} &= 1 \end{aligned}$$

Entropy due to relations is similar. Within the target category, 6 of the dyadic relationships don't include the relevant feature. Thus, their probability of co-occurrence is .25. For 15 of the dyadic relations, there is perfect co-occurrence, so their values go to zero.

$$\begin{aligned}
H_{between}^{rel} &= -6 * 0.5(4 * 0.25 \log_2 0.25) \\
H_{between}^{rel} &= -3 \log_2 0.25 \\
H_{between}^{rel} &= 6
\end{aligned}$$

Next, we calculate the within-category entropy.

$$\begin{aligned}
H_{within} &= 1 + 6 \\
H_{within} &= 7
\end{aligned}$$

Finally, we use the within- and between-category entropy to calculate the density.

$$\begin{aligned}
D &= 1 - \frac{7}{28} \\
D &= 0.75
\end{aligned}$$

References

- Sloutsky, V. M. (2010). From Perceptual Categories to Concepts: What Develops? *Cognitive Science*, *34*(7), 1244–1286. doi: 10.1111/j.1551-6709.2010.01129.x