Indiana Undergraduate Journal of Cognitive Science 5 (2010) 3-14

Simultaneous Noun and Category Learning via Cross-Situational Statistics

Tarun Gangwani, George Kachergis

Department of Psychological & Brain Sciences / Cognitive Science Program

Indiana University

Abstract

Previous research shows that people can acquire an impressive number of word-referent pairs after viewing a series of ambiguous trials by accumulating co-occurrence statistics (e.g., Yu & Smith, 2006). The present study extends the cross-situational word learning paradigm, which has previously dealt only with noun acquisition, and shows that humans can concurrently acquire nouns and adjectives (i.e., a natural category with a distinctive, unifying feature). Furthermore, participants are able to learn ad hoc categories of referents consistently co-occurring with a label, while simultaneously learning instance labels. Thus, humans demonstrate an impressive ability to simultaneously apprehend regularities at multiple levels in their environment.

Keywords

Categorization, concept learning, statistical learning, language acquisition, cross-situational learning

Introduction

Many objects have hierarchically organized labels. For example, one may refer to one's pet with the basic-level name "dog," or with a subordinate-level name, e.g. "golden retriever." Labels can also describe a particular category instance: e.g., one's golden retriever may be named "Rex." Hence, objects can be referred to using specific names (e.g., Rex) or using a category label at any level of specificity (e.g., golden retriever, dog, mammal, etc.).

To succeed in learning a natural language, people must learn labels at many levels of abstraction: instance, subordinate, basic, superordinate. These labels are often consistent with physical or abstract referent properties that make that label salient (e.g., all mammals have hair or fur). Although some categories have members that share a number of perceptual properties (e.g., **dogs** typically bark, have four legs, and are furry), other categories may have members that share no perceptual characteristic (e.g., **things in the kitchen**), but belong to a common context or share abstract properties. Categories that share unifying perceptual features are dubbed *natural* categories, and those that share abstract properties (or context) are dubbed *ad hoc* categories.

For a learner to acquire knowledge of an object's names at different levels, it likely requires many co-occurrences of each label and the object in many different situations. The theory that a word's meaning may be learned by experiencing multiple, individually ambiguous naming events is dubbed *cross-situational statistical word learning* (Pinker, 1984), and has proven to be an effective word learning mechanism in adults and children (Smith & Yu, 2008; Yu & Smith, 2007). Previous studies have investigated the effects of word-object pair frequency and the contextual diversity of the learning situations (Kachergis, Yu, & Shiffrin, 2009a), the effects of prior knowledge on learning additional pairs (Klein, Yu, & Shiffrin, 2008), and the mutual exclusivity bias (Yurovsky & Yu, 2008). In the present study, we investigate whether learners can make use of cross-situational statistics to acquire both 1-to-1 (i.e., subordinate-level) and 1-to-many (i.e., basic-level) labels when they are presented concurrently.

In the cross-situational word learning paradigm (Yu & Smith, 2007), subjects are exposed to multiple words and multiple objects on each of a series of training trials and then asked to identify which object each word refers to. A typical training trial consisted of four objects (*A*, *B*, *C* and *D*). Four pseudowords are played on each slide (*a, b, c,* and *d*), where each refers to one object out of four on each slide (*A*-*a*, *B*-*b*, etc.). The correct pairings on any given training trial remain ambiguous, as the word-presentation order is randomized, and thus does not systematically correspond to any object's location. No word to referent pairing ever appeared in two consecutive trials, as this would trivialize the learning of that pairing (see Kachergis, Yu, & Shiffrin, 2009b). In a typical training block, participants would attempt to learn 18 word-referent pairings from 27 12-second trials, during which each stimulus pair appeared six times. At test, learning was assessed by asking participants to identify which object from a subset of four of the 18 objects corresponded to a pseudoword trained on. On average, participants learned 9 to 10 out of 18 pairings, but a couple subjects learned every pairing (Yu & Smith 2007). The regular co-occurrence of a word and an object is sufficient for people to learn to pair that word and referent.

In this study, there will be only two objects on each training trial, but three words will be played. Two of these words (the instance labels) name the displayed referents, and the other word (the category label) refers to both objects on the screen as well as two other objects for each block. A total of 12 word-to-referent (1-to-1) associations and 12 category-to-referent (1-to-many) associations will be shown in each block. In two of the three experimental blocks, the objects used are uncommon items (*e.g.* strange tools) and the category label refers to four randomly determined objects (the *ad-hoc category* condition). In the other block, the *natural category* condition, the objects are uncommon shapes with one of three distinct features on a set of four objects. Each category label will always refer to the same feature for each object. So, if objects *A, B, C, D* have a hook shape, category label *X* will denote objects with a hook and *a, b, c,* and *d* will name each object, respectively. Therefore, an example training trial would be objects *A* and *C* on the screen and words *c, a,* and *X* played. At the end of training, there will be two separate tests to measure learning of 1-to-1 and 1-to-many relations. For the 1-to-1 relation, participants will need to choose which object from all the objects trained on matches a word played using 12-alternative forced choice testing (12AFC). For the 1-to-many relation, participants must choose from three objects, where each object is a representative of each category that the word played belongs to (3AFC). Additionally, subjects will also be asked to do a similar test, but with novel objects, to gauge participants' ability to generalize.

Participants are not informed of the hierarchical nature of the mappings: they are simply instructed to learn which words belong with which objects. From their perspective, there will be some words (the category labels) that appear more often, and with more referents. It likely will take some time for subjects to realize that these labels are not simply noise, but are consistently appearing with four referentsIn the Yu & Smith (2007) study, there were 18 1-to-1 relations (18 objects and 18 referents) to be learned. In addition to 12 1-to-1 relations (12 objects and 12 referents), the present study also tests the participant to learn 12 1-to-many relations (which includes 4 stimuli in each category), even though there are fewer total stimuli in memory.

Even though we have seen that people are good at learning 1-to-1 relationships, the main question to be answered is whether people can learn both 1-to-1 and 1-to-many relationships simultaneously. In order to test this, the experiment was organized in three blocks. Block 1 is an ad-hoc condition focusing on simultaneous category and name learning ability. The ad-hoc condition is first to gauge whether subjects can notice and learn the category relationship even if there is no perceptual feature to clue them into the relation. Block 2 is a natural category condition which, in distinction to the previous block, focuses on categorization using an external feature and generalization to novel objects. With the second block, subjects should have a better chance of grasping the 1-to-many relationships. Block 3 is another ad-hoc condition to investigate the effects of having a natural category condition in the previous block. While the function of the category label becomes clearer in the 2nd and 3rd blocks, there still may be difficulty in keeping track of both types of associations. In the present study, we mainly investigate whether people can learn both relationships simultaneously, but one condition also tests the effect of having a shared, external feature amongst the objects.

Experiment Overview

The experiment utilized the cross-situational paradigm by presenting the subjects with a set of words and pictures across several training slides. Each training slide displayed two objects from a set of 12 stimuli, which were presented simultaneously with three nonsense words from a set of 15 words played through a speaker. The stimuli were presented in a random order, and for each block different stimuli were used. Subjects were trained on several pairings of 1-to-1 or word-to-referent associations as well as several pairings of 1-to-many or category associations. The 1-to-1 relationship is a word that co-occurs most often with a single referent. In contrast, the 1-to-many relationship is a word that co-occurs equally often with four referents. For example, the word **dog** is associated with many referents (e.g. puppies, golden retrievers, etc.), so its relationship with its associated stimuli is 1-to-many.

On each training trial, three words and two pictures were presented, where one of these words represented a 1-to-many label (the category label).

The subjects were trained on two different conditions (in three blocks):
Block 1 was an *ad-hoc category* condition, in which the objects had no obvious shared perceptual features.
Block 2 was a *natural category* condition, in which the objects in each category's objects share a salient feature (e.g., a hook or arrow shape).
Block 3 was another *ad-hoc category* condition (with different stimuli) to gauge attention shift after learning natural categories.

The subject, after repeated pairings of the words and the stimuli, was then tested on their knowledge of the category label relationship as well as the name of the object. To test the 1-to-1 relationships, subjects heard each word they were trained on and chose one object from all the 12 objects presented in that block. To test the 1-to-many relationship, subjects heard a word they were trained on and chose one stimulus from three stimuli on screen, chosen at random as a representative member of each category. Since the relationship between the category label and the stimuli that are associated with it were not explicit in the ad-hoc condition, one can posit that performance will not be as strong as in the natural category condition.

Subjects

Participants were 24 undergraduates at Indiana University who received course credit for participating. None had participated in other cross-situational experiments.

Stimuli

Each training trial consisted of two novel objects shown on a computer screen. For the ad-hoc conditions, these objects were mostly strange tools. For the natural category condition, these objects were odd shapes with different textures filling the surface and one of three features protruding from the shape. These objects on screen were shown concurrently with 3 pseudowords, spoken sequentially. The 45 pseudowords generated by computer are phonotactically-probable in English (e.g. "stigson"), and were spoken by a monotone, synthetic voice. The 36 pairs of training stimuli were randomly assigned to three sets of 12 word-object pairings, one set for each block. Additionally, four word-object pairings were randomly assigned to 3 sets of 4 word-object pairings in each block to distinguish each category set. In the natural category condition, an additional 12 pairs of testing stimuli were used for a generalization task.

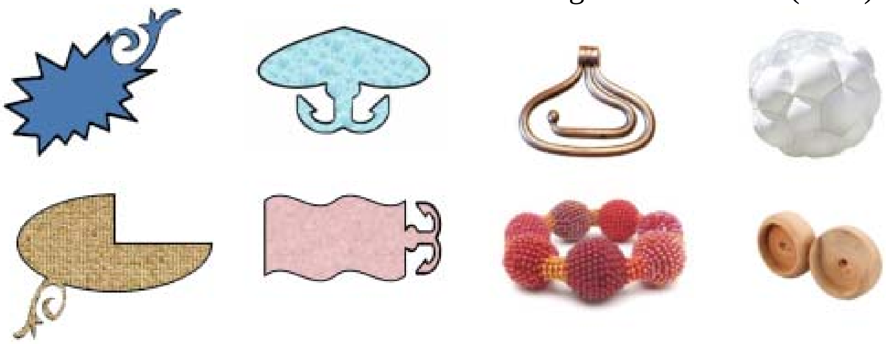

Figure 1: Examples of objects used in experiment. Each object was assigned at random a label to name it and a label to unify it into one of three categories. **Left**: In the natural category condition, objects with multiple types of textures and three different protruding shapes were used in training. **Right**: In the ad hoc condition, objects had no apparent unifying feature.

Referents \ Words	X (12)				Y (12)				Z (12)			
	A	B	C	D	E	F	G	H	I	J	K	L
a	6	2	2	2	0	0	0	0	0	0	0	0
b	2	6	2	2	0	0	0	0	0	0	0	0
c	2	2	6	2	0	0	0	0	0	0	0	0
d	2	2	2	6	0	0	0	0	0	0	0	0
e	0	0	0	0	6	2	2	2	0	0	0	0
f	0	0	0	0	2	6	2	2	0	0	0	0
g	0	0	0	0	2	2	6	2	0	0	0	0
h	0	0	0	0	2	2	2	6	0	0	0	0
i	0	0	0	0	0	0	0	0	6	2	2	2
j	0	0	0	0	0	0	0	0	2	6	2	2
k	0	0	0	0	0	0	0	0	2	2	6	2
l	0	0	0	0	0	0	0	0	2	2	2	6

Figure 2: The completed association matrix after each block. Words indicate the words played and referents indicate the objects associated with each word by letter. Each word and referent association was seen 6 times, as indicated on the diagonal. Additionally, each referent appeared with every other referent 2 times, but never appeared with referents outside its category (as indicated by the shaded regions corresponding to each grouping). Each category label was played to correspond with its respected shaded grouping 12 times.

Training for each condition consisted of 36 trials. Each training trial began with the appearance of two objects, which remained visible for the entire trial. After two seconds of initial silence, each word was heard (randomly ordered, duration of one second) followed by two additional seconds of silence, for a total of 9 seconds per trial.

After each training phase was completed, participants were tested for knowledge of noun labels (i.e., word-object mappings) and category labels (i.e., word-shared feature mappings). A single word—an instance or a category label—was played on each test trial. On noun test trials, all 12 referents were displayed. On category test trials, three referents were displayed; one exemplar from each of the three categories defined by the unifying perceptual feature.

Procedure

Participants were informed that they would experience a series of trials in which they would hear some words and see some pictures. They were also told that their knowledge of which words belong with which objects would be tested at the end. After training, their knowledge was assessed using 12-alternative forced choice (12AFC) and 3-alternative forced choice (3AFC) testing: on each test trial a single word was played, and the participant was instructed to choose the appropriate object from a display of all 12 or 3 objects representative of 3 categories of objects. Condition order was counter balanced.

Results: Block 1 – Ad Hoc Category

Figure 3 displays the learning performance for the two types of associations (1-to-1 and 1-to-many) the subjects were tested on. Subjects performed well on identifying the explicit 1-to-1 relations, performing significantly better above chance during testing ($M = .52$ $t(22) = 10.219$). Subjects' performance on recognizing the category relation was also above chance but the results were not as notable.

Figure 4 displays the correlation between 1-to-1 and 1-to-many performance across all the subjects tested. There is no trend present when subjects are first presented with an ad hoc condition ($r^2 = .197$).

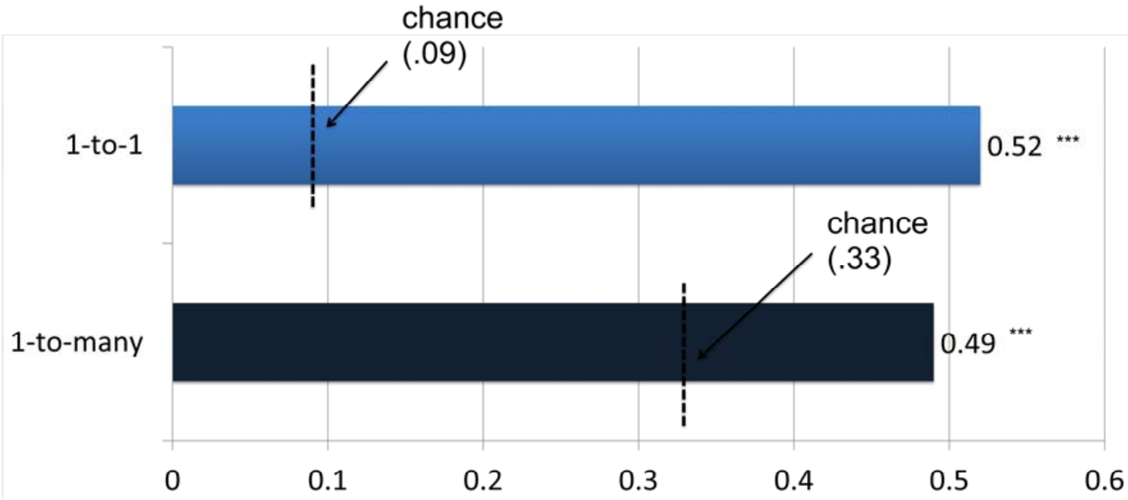

Figure 3: Learning performance for 1-to-1 and 1-to-many mappings in Block 1. Participants performed far more above chance (9%, 12AFC) for 1-to-1 pairings than in for 1-to-many pairings (33%, 3AFC).

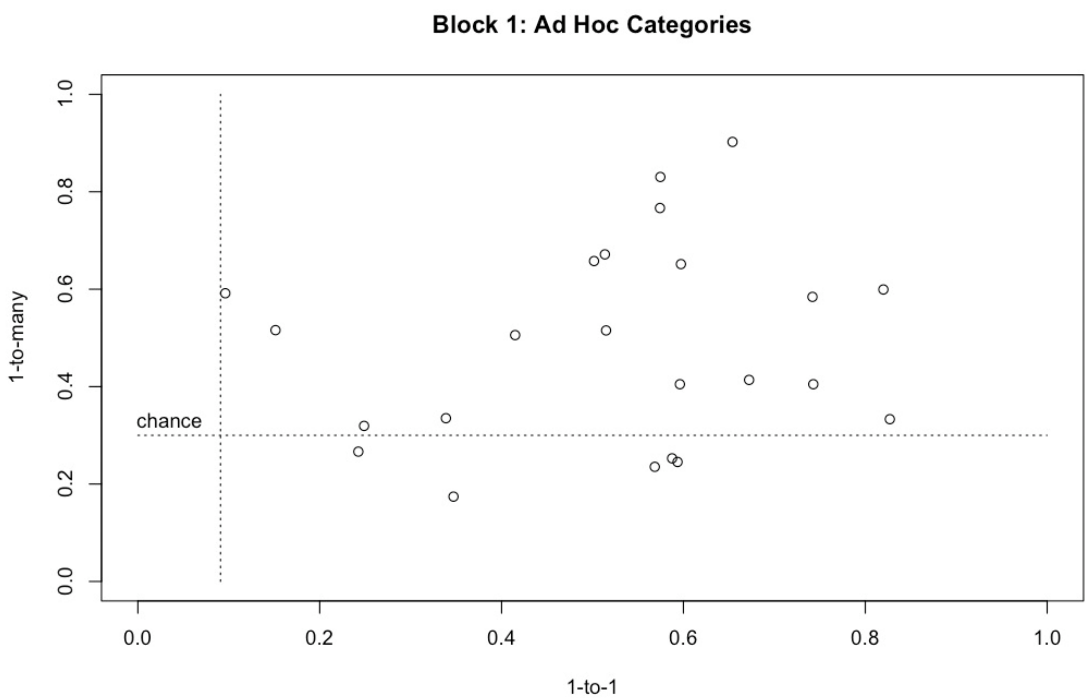

Figure 4: Correlation of 1-to-1 performance to 1-to-many performance across all subjects. No significant trend is present upon first presentation of an ad hoc condition.

Results: Block 2 – Natural Category

Figure 5 displays the learning performance for the two types of associations tested as well as performance on generalizing the explicit category feature trained on to new stimuli at test. The test slides for generalization were the same as the test for the 1-to-many test slides except that the only previously-seen parts of the stimuli were the distinct, unifying features (e.g., a hook) that were seen in training to distinguish the different category types.

In contrast to Block 1, the 1-to-1 performance was decreased (*M* = .33), and 1-to-many learning reached a higher level of performance: the blatant category similarity shifted their attention (*M* = .59). In addition, their ability to generalize the explicit feature is reflected in their significant performance above chance (*M* = .53).

Figure 6 displays the correlation between the 1-to-1 performance and the 1-to-many performance. Again, no significant trend is present (r^2 = -.095).

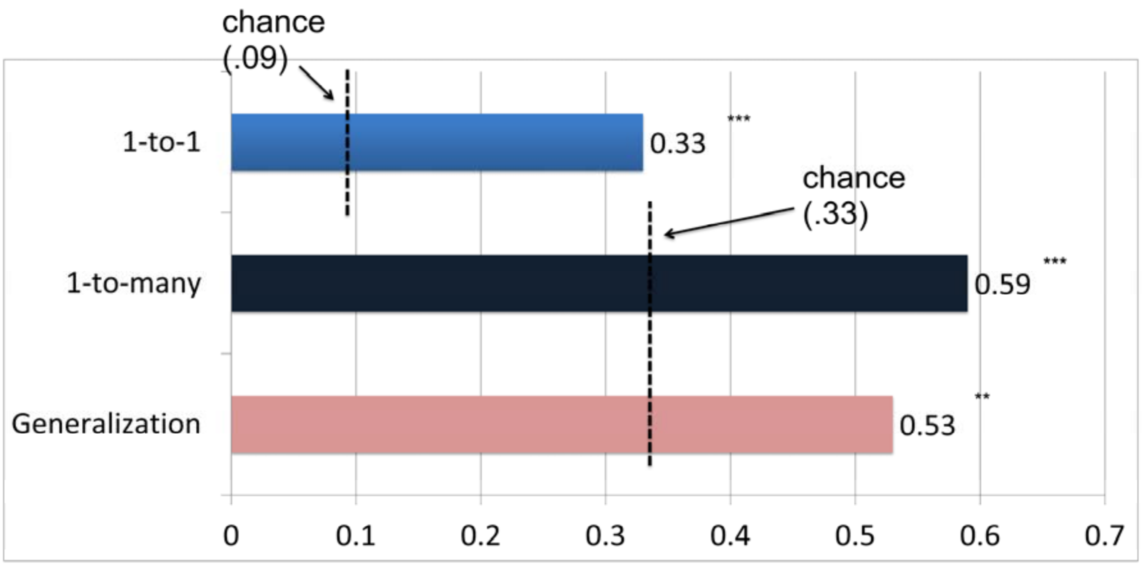

Figure 5: Performance on 1-to-1 and 1-to-many associations for Block 2. Subjects performance contrasted sharply from Block 1 as their performance on 1-to-1 testing decreased but their 1-to-many performance increased. Subjects also performed well at the generalization task, solidifying their ability to learn category relations.

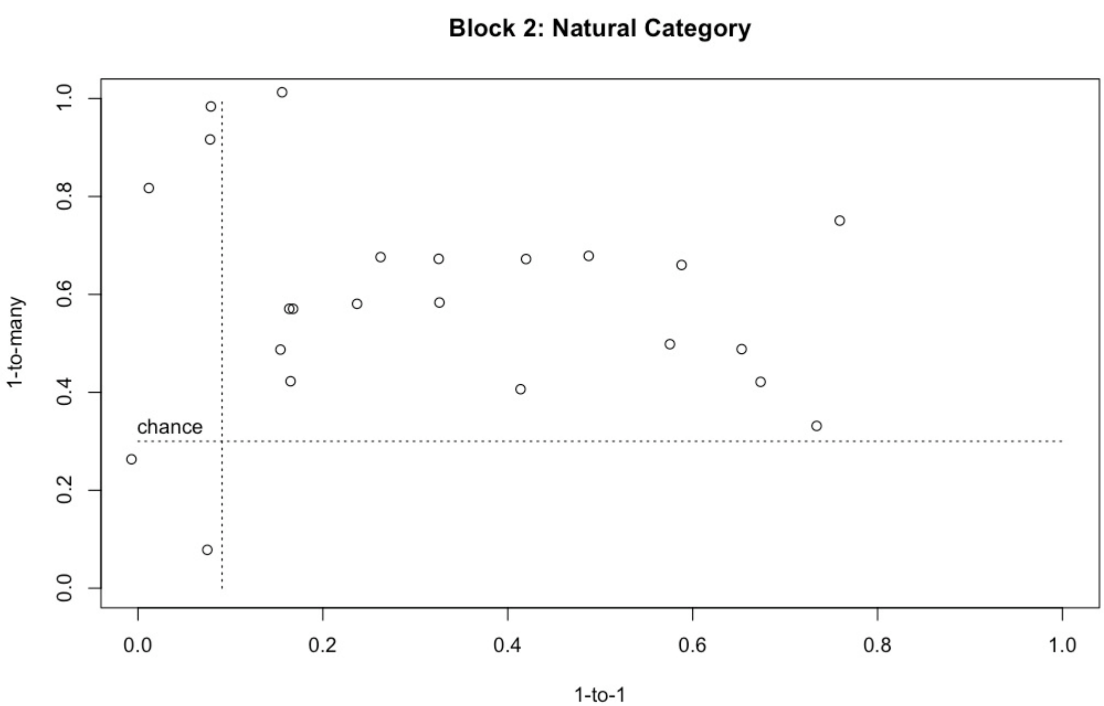

Figure 6: Correlation of 1-to-1 performance to 1-to-many performance across all subjects. No significant trend is present upon first presentation of an ad hoc condition.

Results: Block 3 – Ad Hoc Category (itera)

Figure 7 shows the learning performance for the two associations to serve as a comparison to the results in Block 1 (Figure 1) as this Block is the same condition as Block 1. Here, the performance on 1-to-1 testing is much lower than that of Block 1 (*M* = .34, t(22) = 4.875), and performance on 1-to-many testing is much higher than Block 1 (*M* = .53, t(22) = 4.289). It is important to note that these results occurred *after* the natural category condition in Block 2.

Figure 8 shows a positive correlation between 1-to-1 and 1-to-many performance after Blocks 1 and 2 (r^2 = 0.59).

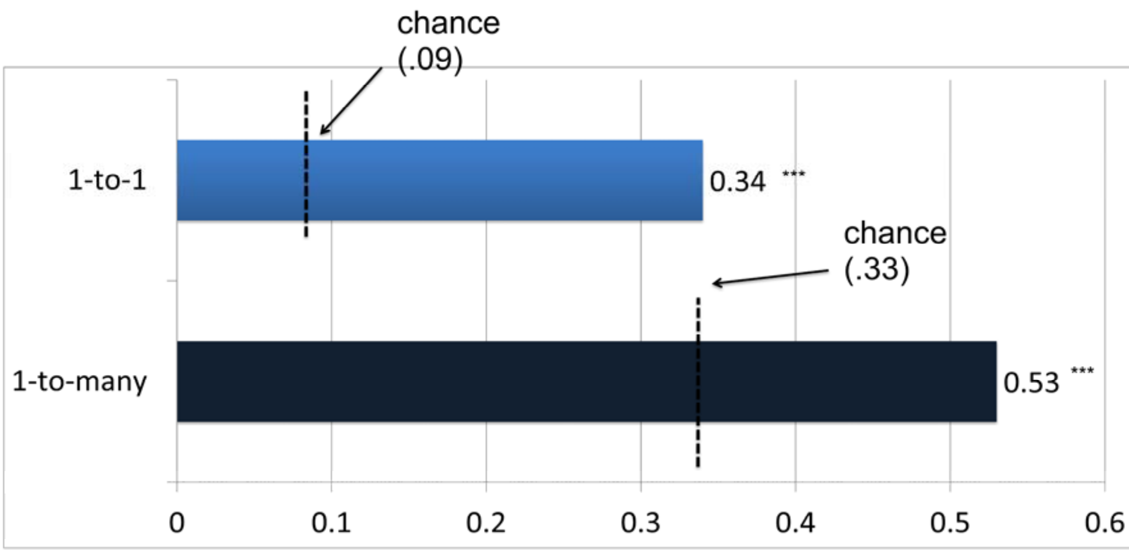

Figure 7: Performance on 1-to-1 and 1-to-many associations for Block 3. In comparison to Block 1 (the first ad-hoc condition test), subject performance on 1-to-1 relations decreased and subject performance on 1-to-many relations increased.

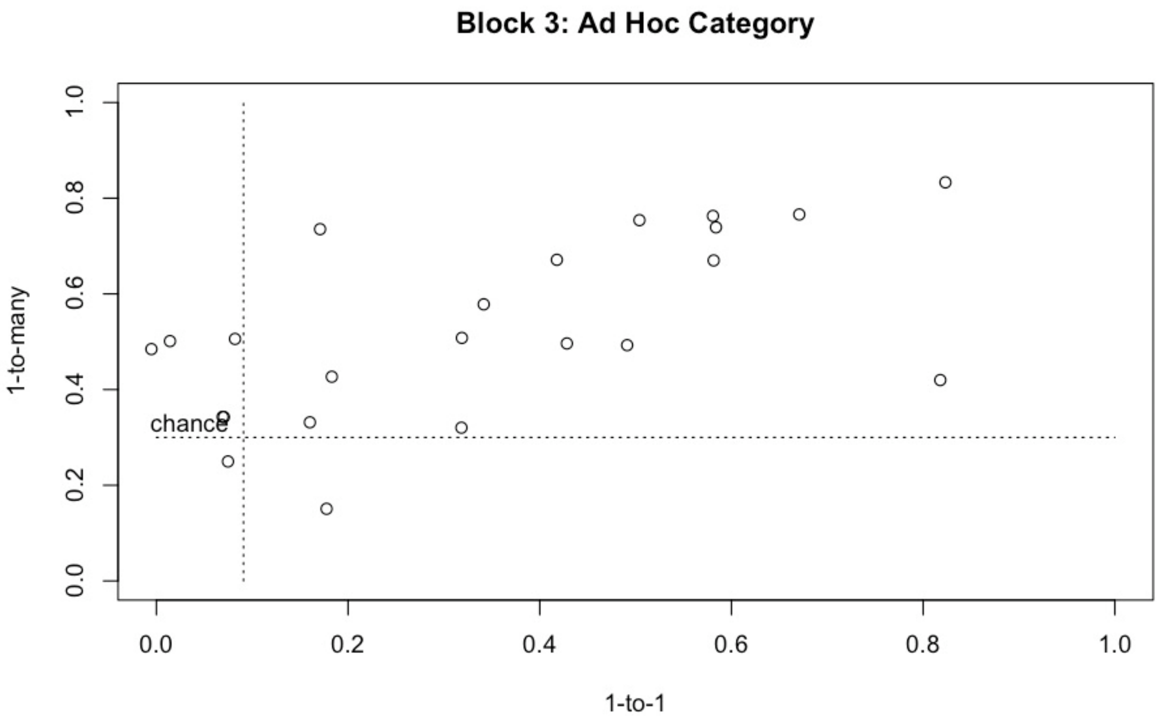

Figure 8: Correlation of 1-to-1 performance to 1-to-many performance across all subjects. There is a strong positive correlation of 1-to-1 to 1-to-many performance.

Discussion

Objects are frequently associated with multiple labels. The participants were tested to see if they could learn both the 1-to-1 and 1-to-many label for each referent. In one block, participants were also tested to see if they could generalize their knowledge of a concept to novel objects. Subjects at first did not grasp the need to learn the category label and thus likely disregarded it as noise (1-to-many, $M = .49$) but did very well on the 1-to-1 relationship ($M = .52$). After the second block with the natural categories, subjects' performance on category learning markedly improved ($M = .59$), but the performance on the explicit 1-to-1 relationship dropped ($M = .33$). Therefore, in this experiment there was an apparent trade-off: if a participant focuses on learning which word names which picture, their performance is markedly well in this type of learning. If the participant focuses on the category relation, then their performance is improved in the 1-to-many association instead of the 1-to-1 association.

Specifically, Block 1 demonstrates that subjects at first have a strong bias to naming objects and a need to ignore certain information unless they are clearly indicated of its importance beforehand. Block 2 suggests that after the explicit feature is presented, subjects may shift their attention to this explicit feature, which overall had decreased their learning of 1-to-1 relations. Intriguingly, Block 3 shows that subjects can still recognize a categorical relationship, even with no extrinsic feature present. But, their learning performance for 1-to-1 associations did not improve even after being experimented on both of the conditions in this study. One of the conclusions of Rosch & Mervis' (1975) study was that category members need not share the same feature, but only need to resemble each other. In our study, Block 3 suggests that category members may not even resemble each other, as performance was still well above chance for the category relationship. However, their performance did slightly decrease as compared to the results in Block 2 (Block 2: M = .59; Block 3: M = .53), suggesting that having a natural feature present does facilitate learning.

Why was it that, on average, a participant could not perform equally well in both associations? Even after Block 2, the performance on the 1-to-1 association did not increase by a significant amount. However, mean performance alone may not be a correct measure of simultaneous 1-to-1 and 1-to-many learning. In Figure 8, the correlation between 1-to-1 vs. 1-to-many performance is shown, and there is a positive trend (r^2 = .59): a subjects performance on one type of relationship does affect their performance on the other type. As further evidence for simultaneous learning, for incorrect 1-to-1 answers given on the 12AFC test at the end of each Block, subjects were at chance for naming an object in the same category (M = .39). Therefore, a majority of participants did learn both types of associations

For future studies, we would like to see whether subjects can retain the category relation over a longer period of time, i.e. one or two weeks. Another possibility is to use more complex structures for category learning. For example, two words and two referents are played on each trial, with only one word naming a referent and another word naming a category only, but not the other referent. For example, words *a*, *X* are played and referents *A*, *C* are displayed. In later trials, subjects could be trained on either 1-to-1 or 1-to-many, but not both types of relations. This could further complicate the 1-to-1 bias participants may have and block the ability to learn certain referent names.

Acknowledgements

Thanks to Brenden Sewell, editor of the Undergraduate Journal for Cognitive Science at Indiana University, for allowing me to publish, as well as the IU Undergraduate Cognitive Science Conference for inviting me to do so. Special thanks to George Kachergis for mentoring me through the entire process, as well as lab directors Chen Yu and Richard Shiffrin.

References

Kachergis, G., Yu, C., & Shiffrin, R. M. (2009a). Frequency and Contextual Diversity Effects in Cross-Situational Word Learning. In N. Taatgen, H. van Rijn, J. Nerbonne, & L. Schomaker (Eds.) *Proceedings of the 31st Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Kachergis, G., Yu, C., & Shiffrin, R. M. (2009b). Temporal Contiguity in Cross-Situational Statistical Learning. In N. Taatgen, H. van Rijn, J. Nerbonne, & L. Schomaker (Eds.)

Proceedings of the 31st Annual Conference of the Cognitive Science Society. Austin, TX: Cognitive Science Society.

Klein, K. A., Yu, C., & Shiffrin, R. M. (2008). Prior knowledge bootstraps cross-situational learning. *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 1930-5). Austin, TX: Cognitive Science Society.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Rosch, E. & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*. 573-605.

Smith, L. & Yu, C. (2008). Infants rapidly learn wordreferent mappings via cross-situational statistics. *Cognition,* 106, 1558-1568.

Yu, C. & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science,* 18, 414-420.

Yurovsky, D. & Yu, C. (2008). Mutual exclusivity in cross-situational statistical learning. *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 715-720). Austin, TX: Cognitive Science Society.