

Running Head: REDUNDANT LABELS FACILITATE CATEGORIZATION

Language is not just for talking: redundant labels facilitate learning of novel categories.

Gary Lupyan^{1,2}, David H. Rakison¹, James L. McClelland³

1. Department of Psychology, Carnegie Mellon University

2. Center for the Neural Basis of Cognition

3. Department of Psychology, Stanford University

Gary Lupyan

342C Baker Hall

Carnegie Mellon University

Pittsburgh, PA 15213

Tel: 917-843-4868

Fax: (412) 268-2798

Email: glupyan@cnbc.cmu.edu

WORD COUNT: 3999

In addition to their communicative functions, verbal labels may play a role in shaping concepts. Two experiments assessed whether the presence of redundant labels affected category formation. Participants learned to categorize “aliens” as those to be approached or those to be avoided. After receiving feedback on their response, a nonsense label was either presented or not. Providing nonsense category labels facilitated category learning even though the labels were redundant and all participants had equivalent experience with supervised categorization of the stimuli. A follow-up study investigated differences between learning verbal and nonverbal associations. The findings show that labels make category distinctions more concrete and bear directly on the language and thought debate.

The ability to form categories is ubiquitous in the animal kingdom, yet it is only humans who habitually use names for theirs.¹ Learning to name things in the world allows for linguistic communication, and it has been argued that words stabilize abstract ideas in working memory and make them available for inspection (Clark & Karmiloff-Smith, 1993; Clark, 1997; James, 1890; Rumelhart, Smolensky, McClelland, & Hinton, 1986; Vygotsky, 1962). Although a great deal of research has examined the communicative function of words (see Roy, 2005 for a novel approach) and the role of categorization in perception (Goldstone, 1994; Goldstone & Barsalou, 1998), to date few studies have addressed the role that category names themselves play in the learning of categories. Learning that two objects are both called “dax” may communicate to the learner that they share commonalities, which in turn may cause the objects to be grouped into a common category (Waxman & Markow, 1995). The presence of words may therefore turn an unsupervised learning task into a supervised one (Cabrera & Billman, 1996). But do words do more than alert the learner to group together similarly named objects? While non-human animals categorize non-linguistically, humans have the potential benefit of labels—category names—as they decide to which category an object belongs. The crucial question, then, is how does the presence of labels affect human categorization?

The question of the role of names in categorization is related to investigations of the effects of categories on perception. Numerous studies have shown that perception of initially-equidistant items is affected by categorization; most notably, categorizing items into different groups makes them more dissimilar, and this effect has been shown to be the result of representational change rather than similarity ratings being exaggerated by knowing that two objects belong to the same category (Goldstone, Lipka, & Shiffrin, 2001). Though relevant to the present studies, the question here is whether labels themselves facilitate category formation.

The specific question of how labels affect category-learning has been most thoroughly investigated in two contexts. First, work on children's language acquisition revealed that infants as young as 9 months of age more readily individuate labeled than unlabeled objects (Xu, 2002), and by 12 months of age have an expectation that words refer to object categories (Waxman & Hall, 1993; Waxman & Markow, 1995). These researchers have argued that words serve as "invitations for categories" with superordinate labels such as "vehicle" leading children to form the appropriate category (Waxman & Markow, 1995). When correlated with regularities in the world (e.g., shape is predictive of solid categories but not non-solid ones), labels have been shown to improve the learning of these regularities (Yoshida & Smith, 2005); in these experiments the labels are semantically empty. A different sort of labeling influence can be seen in children. For instance, learning to associate small-medium-large groupings with the labels "baby"- "mommy"- "daddy" facilitates performing relational judgments, enabling children to transfer the size relation to novel stimuli; here, the effect of the labels is tied to their semantics (Kotovsky & Gentner, 1996; Ratterman & Gentner, 1998).

Second, Miller and Dollard (1941), extending James's "law of dissociation by varying concomitants," (1890) hypothesized that associating different responses with otherwise similar stimuli should increase the difference between the stimuli. Object names were seen as a kind of motor response, and so associating undifferentiated stimuli with different names was hypothesized to increase the salience between the stimuli and facilitate categorization (Rossman & Goss, 1951). The hypothesis was tested with mixed results because the relevant experiments failed to control for stimulus familiarity (e.g., Arnoult, 1953; Battig, 1956; Rossman & Goss, 1951), and it was unclear whether increased discriminability or facilitated categorization arose

from the learned associations between stimuli and labels or merely from additional experience with the stimuli (Gibson & Gibson, 1955; Robinson, 1955).

It is important to be clear what we mean by the term “label.” As used here, a label is anything that is (1) consistently correlated with a category and (2) is used to refer to a category. The category can comprise objects, sounds, actions, spatial relations, etc. In principle, any cue can serve as a label and what counts as a label is a function of experience and environment. Individuals proficient in a sign language treat motor gestures as labels; our participants, being hearing college students have had an immense amount of experience treating words (both oral and written) as labels.

It is often through words that we come to know what categories are relevant. For instance, calling certain objects “chairs” suggests that chairs are a useful and relevant category. The question addressed by the present work was not whether labels facilitate category formation because they point out the relevant categories, but instead whether labeled categories are easier to acquire *because* they have a name—even when categorization can be performed without relying on labels. Thus, the present set of studies are the first to directly test the idea that labels make category differences “more concreted” (James, 1890, p.333). The two experiments reported here address two main questions: Are labeled categories easier to learn even when the labels are entirely redundant, contributing no additional information? How does associating stimuli with verbal labels compare to learning a non-linguistic category association?

Experiment 1

In this experiment, we measured performance of participants learning to associate novel objects with behavioral responses. Participants performed this task while also learning names for

the stimulus categories (*label-condition*) or not (*no-label-condition*). If it is easier to learn named categories, performance should be superior in the *label* condition.

Method

Participants. Forty-eight Carnegie Mellon University undergraduates (ages 18-24) participated in the experiment for course credit. The participants were randomly assigned to *label* and *no-label* groups. Data from 4 participants were excluded for failure to follow instructions. Data for the testing part of the task were not available for two of the participants due to experimenter error.

Materials. The stimuli were a subset of the YUFO stimulus set (Gauthier, James, Curby, & Tarr, 2003) (Figure 1). The difference between the two categories was one of configuration rather than any simple features. Items in one category (Figure 1-left) have flatter bases and a subtle ridge on their “heads.” The stimuli placed in the other category (Figure 1-right) have more rounded bases, and a smoother head. Judging by their responses on a post-study questionnaire following the category training, participants were overwhelmingly more attuned to the distinction in the heads and used adjectives like “pointy” versus “fat,” or “bumpy” versus “smooth” when asked to verbalize the difference between the two categories.²

The stimuli were presented on a black background on a 17” CRT screen and subtended 8° of visual angle. Responses were collected using a gamepad controller. For the *label* condition, the categories were associated with the nonsense labels “leebish” and “grecious” which were displayed in a white 16-point font.³

Training Procedure. Participants were told that they were an explorer on another planet learning about alien life forms. Their task was to determine which aliens they should approach and which they should move away from. On each training trial, one of the 16 aliens appeared in

the center of the screen. After 500ms, an outline of a character in a spacesuit (the “explorer”) appeared in one of four positions—left, right, top, or bottom of the alien. Participants were instructed to respond with the appropriate direction key depending on the category of the alien. For instance, if the explorer appeared to the top of the alien the participants would press the ‘down’ key to move towards the alien, or the ‘up’ key to move away, following which the explorer would move towards/away from the alien. Auditory feedback—a buzz for an incorrect response and a bell for a correct response—sounded 200ms after the explorer stopped moving. In the *label* condition, a printed label (“leebish” or “grecious”) appeared to the right of the alien 300ms after the feedback. After 1500ms, the screen was erased and a fixation cross marked the start of the next trial. The total trial duration was equal for both conditions. For all experiments the pairing of the label with the category (move away, move towards), and the perceptual stimuli (family₁, family₂), was counterbalanced across participants. Participants in the *label* condition were told that previous visitors to the planet have found it useful to name the two kinds of aliens, and that they should pay careful attention to the labels. All participants received the same number of categorization trials (9 blocks of 16 trials each) and had equal exposure to the stimuli. The only difference between the two conditions was the presence or absence of the verbal labels appearing after each response.

Testing procedure. Following the training trials, participants completed a test phase. On each trial, one of the aliens appeared in the center of the screen and participants were instructed to classify it as a kind to be approached or escaped from. There was no feedback and no presentation of the names learned in the *label* condition. To determine whether participants formed category representations rather than just memorized specific examples, the testing phase included, in addition to previously categorized stimuli, stimuli from the learned categories

unseen during training. There were 4 blocks of 24 trials each (16 old+8 new stimuli that tested generalization performance).

Results and Discussion

Training phase. All results were analyzed using repeated-measures mixed design analysis of variance (ANOVA) with label condition (*label* versus *no-label*) as a between-subjects factor and block and stimulus-type as within-subject factors. The analyses revealed that there were no differences between different pairings of stimuli/labels/responses, all $F_s < 1$. Performance improved over time, $F(8,42)=33.12$; $p_{rep} > .99$, $\eta_p^2 = .45$, with final performance in both groups asymptoting near ceiling by the end of the 144 training trials. The *label* group was significantly more accurate ($M = .88$, $SD = .11$) than the *no-label* group ($M = .80$, $SD = .17$), $F(1,42)=9.03$, $p_{rep} = .98$, $\eta_p^2 = .15$. The *label* group also learned to categorize significantly faster, as revealed by a significant condition $\times$ block interaction, $F(8,336)=2.59$, $p_{rep} = .95$ (Figure 2 left). For instance, participants in the *no-label* group reached 80% correct classification after approximately 72 trials; participants in the *label* group reached the same level of performance in approximately 30 trials.

There was no difference in RTs between the conditions: *label* ($M = 1229$ ms; $SD = 254$ ms), *no-label*: ($M = 1178$ ms; $SD = 270$ ms), $F(1,42) < 1$; the RT condition $\times$ block interaction was also not significant, $F_s < 1$.

Testing phase. In this phase, which immediately followed the training, no feedback or labels were provided. Overall generalization performance was far above chance ($M = .86$, $SD = .11$), though below the performance on the previously seen items ($M = .92$, $SD = .08$), $F(1,40)=15.41$, $p_{rep} > .99$. Training performance between the groups was statistically indistinguishable for the last 3 blocks of training, but once response feedback was removed

performance in the *label* condition ($M=.93$, $SD=.04$) remained significantly higher than in the *no-label* condition ($M=.87$, $SD=.10$), $F(1,40)=7.01$, $p_{rep}=.96$ (Figure 2 right). Reaction times did not differ between conditions ($M_{label}=904\text{ms}$, $SD=332\text{ms}$; $M_{no-label}=875\text{ms}$, $SD=253\text{ms}$), $F(1,40)<1$.

An important goal of the *testing phase* of the task was to see whether representations formed in the presence of labels were more robust. Support for this idea was found from analyzing the responses to the novel stimuli. Generalization performance of participants in the *label* improved with time, $F(3,60)=3.79$, $p_{rep}=.96$, which was not the case for participants who were not exposed to labels, $F(3, 60)<1$. This suggests that participants who previously learned the labels were better able to generalize the learned categories to unseen items even without feedback or labels to guide them. Performance of the *label* participants on the novel stimuli was superior to that of the *no-label* participants, but the condition $\times$ block interaction did not reach significance, $F(3,120)=1.63$, $p_{rep}=.83$.

Data from Experiment 1 show that associating initially meaningless labels with novel categories facilitated their categorization. This was true even though the names provided no additional category information. Participants who learned the categories in the context of labels not only learned to categorize the two families of aliens more quickly, but also showed superior category knowledge when response feedback and labels were removed.

Experiment 2

Experiment 1 had two potential shortcomings. First, it was possible that the placement of written labels may have attracted participants' attention to some part of the stimuli that was particularly informative of the category distinction. This was addressed by introducing an *auditory-label* condition that equated the visual information presented in the *label* and *no-label* conditions. Second, it remains unclear whether the facilitation observed is specific to labels, or if learning any additional association would produce the same type of facilitation. It is known that clusters of correlated cues reinforce each other (Billman & Knutson, 1996; Goldstone, 1998), and because the labels were associated both with the stimuli and with behavioral responses (approach/escape) they may have strengthened the association between the stimuli and responses. If this was the case then learning nonverbal associations should produce a similar effect. However, if there is something special about associating words with category exemplars, then facilitation due to labels should be greater than that due to learning a nonverbal association.

Methods

Participants. Seventy-five Carnegie Mellon University undergraduates participated in the experiment for course credit (ages 18-24). The participants were randomly assigned to four conditions: *no-label* (n=22), *written-label* (n=18), *auditory-label* (n=18), and a nonverbal association condition—*location* (n=17).

Materials and Procedure. The materials and procedure were identical to those used in Experiment 1 with the following exceptions: In the *auditory-label* condition, the written labels were replaced by auditory sound clips of a female saying “leebish” and “grecious.” In the *location* condition, participants were told that some aliens live on one side of the planet, and others live on the other side. Following the response (approach/escape), and auditory feedback,

the alien moved up or down to signal where it “lived.” The motion started 300ms after response feedback, and lasted approximately 400ms. The trial ended 1200ms after the alien stopped moving. The result of this was that the alien remained on the screen at least as long after onset of feedback in the motion condition as in the label condition.

To measure the degree to which participants learned the association between stimuli and labels/locations, we included verification trials. Verification trials were presented after a random 10% of training trials. For each verification trial, one of the aliens appeared randomly paired with a label, and a query asking: “Is this one leebish [grecious] yes/no?” (*label* conditions). For the *location* condition, the presented alien moved up or down and participants responded to the query “Is this correct? yes/no.” Participants were allowed to repeat the motion numerous times before making their response. No feedback was provided for the verification trials.

Results and Discussion

A mixed ANOVA with label-condition as a between-subjects factor and block as a within-subjects factor showed a significant main effect of condition, $F(3,71)=3.18$, $p_{rep}=.94$, $\eta_p^2=.08$ and a block $\times$ condition interaction, $F(24,568)=1.66$, $p_{rep}=.94$, $\eta_p^2=.07$ (Figure 3). Planned comparisons (simultaneous Tukey tests with CI=95%) showed that the two label conditions did not differ from each other, $F(1,34)=.26$, and that the *no-label* condition did not differ from the *location* condition $F(1,37)=0.00$ (Figure 3). Pooling the data from the two *label* conditions, and comparing it against the combined *no-label* and *location* conditions revealed a significant main effect, $F(1,73)=11.10$, $p_{rep}=.99$, $\eta_p^2=.15$, and a condition $\times$ block interaction, $F(8,584)=2.99$, $p_{rep}=.98$, $\eta_p^2=.04$. A direct comparison of the *written-label* condition to the *location* condition revealed a significant main effect, $F(1, 33)=4.11$, $p_{rep}=.92$, $\eta_p^2=.08$ and a significant condition $\times$ block interaction, $F(8, 264)=3.58$, $p_{rep}=.99$, $\eta_p^2=.09$. As in Experiment 1,

labels improved categorization. This improvement was seen both for written and spoken labels relative to a no-label condition as well as the *location* condition in which participants learned to associate alien categories with locations rather than nonsense words.⁴

Overall, verification accuracy was correlated with training accuracy $r(50)=.34, p=.02$. The three-way comparison between the *location*, *written-label*, and *auditory-label* conditions was significant, $F(2,50)=5.28, p_{rep}=.97, \eta_p^2=.17$. Planned comparisons revealed that accuracy in the *auditory-label* condition ($M=.92, SD=.07$) was significantly greater than the *location*, and *written-label* condition, but the latter two conditions did not differ (both $M=.82; SDs=.13$). Although spoken labels were learned better than written labels, there was no difference in categorization performance between the two groups. At the same time, participants learned the association between the stimuli and where they “live” just as well as they learned the written labels, yet categorization was facilitated by learning the labels but not the locations.

General Discussion.

In this work we posed two main questions: (1) Are labeled categories easier to learn than unlabeled categories? (2) How does associating stimuli with verbal labels compare to learning a non-linguistic category association? Given equal experience categorizing novel stimuli, participants learned correct classification more quickly when they also learned names for the stimuli even though the names constituted extra material that needed to be learned. Learning nonsense verbal labels facilitated categorization beyond that of learning a non-verbal association. There is also some evidence that categories learned with labels are more robust than those learned without labels. In Experiment 1, the advantage following training with labels was observed even when labels were no longer present. In a follow-up to the present studies, Lupyan (2006) reported that the presence of labels protected learned categories from interference when

novel stimuli were introduced during training, and facilitated incorporation of these novel stimuli into existing categories, further suggesting that learning labels results in more robust category representations.

These findings provide empirical support to an idea that has existed for over a century—words do more than communicate information between individuals (James, 1890; Whorf, 1956; Clark, 1997)—and bears directly on the language and thought debate (e.g., Gentner & Goldin-Meadow, 2003; Gleitman & Papafragou, 2005). It is shown here that even when exposure and categorization experience is equated, learning names for categories facilitates their acquisition. Consider, for instance, the practice of naming culturally important plants (Berlin, Breedlove, & Raven, 1966). The present results suggest that in addition to whatever communication benefits accrue from naming important plants, exposure to category names helps individuals to classify correctly. Given equal categorization experience (e.g., practice in finding a certain kind of berry), individuals who have a name for the category will have an advantage in classification.

The categories used in the current experiment were by no means incommensurate with linguistic coding by English-speaking adults. Indeed, many participants in the *no-label* condition reported self-generating labels during the learning task. It remains to be seen whether labeling can produce a still greater facilitation in learning categories for which self-generation of labels is more difficult or impossible.

What was it about the labels that facilitated category-learning? The present results are compatible with several accounts. The task required participants to learn a subtle perceptual distinction based on experience with individual category exemplars. In this context, the labels can be thought to provide perceptually-simple correlates to an otherwise perceptually complex task. The labels may have allowed participants to more easily represent the somewhat fuzzy

perceptual distinction between categories: “more rounded and smooth” versus “less rounded with ridges” in terms of a simpler verbal distinction: “leebish” versus “grecious.”

One explanation for the failure to find a facilitation from associating stimuli with categories seemingly just as discrete and simple (“lives above” versus “lives below”) is that adults have much more experience associating words (but not such facts as lives-above/below) with object categories (Colunga & Smith, 2005). Though the up/down distinction was an easily learned extra cue, the non-referential context failed to facilitate category learning. It remains to be seen whether non-word labels used in an explicitly referential context can facilitate categorization.

The general account that names cohere items from common categories because the names serve as reliable cues to class membership is sufficient to explain the present results. An alternate account is that rather than being fixed features, category names modulate item representations on-line through top-down feedback. Activation of the label activates features most relevant to the category, thus dynamically increasing the coherence of the items associated with the label. It is this activation of category-relevant features that results in more powerful category attractors. It is unclear whether once formed, the category attractors continue to depend on labels. Participants in the *label* condition of Experiment 1 continue to show facilitation when labels were removed, but this may have been due to continued implicit labeling.

If facilitation depends on activation of verbal representations then somehow inactivating labels may cause impairments in categorization. This prediction is supported by two sources of evidence. First, linguistic impairments may be linked to impairments in categorization (Goldstein, 1948). Indeed, some types of aphasia are reliably associated with nonlinguistic categorization impairments even in the absence of strong semantic impairments (De Renzi,

Faglioni, Savoiardo, & Vignolo, 1966; Hjelmquist, 1989; Roberson, Davidoff, & Braisby, 1999). Second, verbal, but not visual, interference removes categorical perception of colors and facial expressions in normal adults (e.g., Roberson & Davidoff, 2000; Gilbert, Regier, Kay, & Ivry, 2006).

Some have attributed cross-linguistic differences in performance on tasks such as arithmetic (Pica, Lemer, Izard, & Dehaene, 2004), spatial cognition (Levinson, Kita, Haun, & Rasch, 2002), and biases in memory (Boroditsky, Ham, & Ramscar, 2002) to differences in cultural emphasis and experience rather than differences in language (Li & Gleitman, 2002; Gleitman & Papafragou, 2005). The present findings show that when experience is equated, linguistic differences matter: it is easier to learn named rather than unnamed categories even when the names themselves are entirely redundant, suggesting that differences in language can indeed lead to differences in concept learning.

Acknowledgments.

This work was supported in part by a NSF Graduate Fellowship to the first author and by NIMH program project grant MH64445. We would like to thank Brian Mathias and Ashleigh Molz for help with data collection.

Reference List

- Arnoult, M. D. (1953). Transfer of Predifferentiation Training in Simple and Multiple Shape Discrimination. *Journal of Experimental Psychology*, 45, 401-409.
- Battig, W. F. (1956). Transfer from Verbal Pretraining to Motor-Performance As A Function of Motor Task Complexity. *Journal of Experimental Psychology*, 51, 371-378.
- Berlin, B., Breedlove, D. E., & Raven, P. H. (1966). Folk Taxonomies and Biological Classification. *Science*, 154, 273-&.
- Billman, D. & Knutson, J. (1996). Unsupervised concept learning and value systematicity: A complex whole aids learning the parts. *Journal of Experimental Psychology-Learning Memory and Cognition*, 22, 458-475.
- Boroditsky, L., Ham, W., & Ramscar, M. (2002). What is universal about event perception? Comparing English and Indonesian speakers. In *Proceedings of the 24th Annual Meeting of the Cognitive Science Society* Fairfax, VA.
- Cabrera, A. & Billman, D. (1996). Language-driven concept learning: Deciphering Jabberwocky. *Journal of Experimental Psychology-Learning Memory and Cognition*, 22, 539-555.
- Clark, A. (1997). *Being There: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.
- Clark, A. & Karmiloff-Smith, A. (1993). The Cognizer's Innards: A Psychological and Philosophical Perspective on the Development of Thought. *Mind & Language*, 8, 487-519.
- Colunga, E. & Smith, L. B. (2005). From the lexicon to expectations about kinds: A role for associative learning. *Psychological Review*, 112, 347-382.

- De Renzi, E., Faglioni, P., Savoiaro, M., & Vignolo, L. A. (1966). The influence of aphasia and of the hemispheric side of the cerebral lesion on abstract thinking. *Cortex*, 2, 399-420.
- Gauthier, I., James, T. W., Curby, K. M., & Tarr, M. J. (2003). The influence of conceptual knowledge on visual discrimination. *Cognitive Neuropsychology*, 20, 507-523.
- Gentner, D. & Goldin-Meadow, S. (2003). *Language in Mind: Advances in the Study of Language and Thought*. Cambridge, MA.: MIT Press.
- Gibson, J. J. & Gibson, E. J. (1955). Perceptual Learning - Differentiation Or Enrichment. *Psychological Review*, 62, 32-41.
- Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 489-494.
- Gleitman, L. & Papafragou, A. (2005). Language and thought. In K. Holyoak & B. Morrison (Eds.), *Cambridge Handbook of thinking and Reasoning* (pp. 633-661). Cambridge: Cambridge University Press.
- Goldstein, K. (1948). *Language and language disturbances*. New York: Grune & Stratton.
- Goldstone, R. (1994). Influences of Categorization on Perceptual Discrimination. *Journal of Experimental Psychology-General*, 123, 178-200.
- Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, 49, 585-612.
- Goldstone, R. L. & Barsalou, L. W. (1998). Reuniting perception and conception. *Cognition*, 65, 231-262.
- Goldstone, R. L., Lippa, Y., & Shiffrin, R. M. (2001). Altering object representations through category learning. *Cognition*, 78, 27-43.

- Hjelmquist, E. K. E. (1989). Concept-Formation in Non-Verbal Categorization Tasks in Brain-Damaged Patients with and Without Aphasia. *Scandinavian Journal of Psychology*, 30, 243-254.
- James, W. (1890). *Principles of Psychology. Vol. 1*. New York: Holt.
- Kotovsky, L. & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, 67, 2797-2822.
- Levinson, S. C., Kita, S., Haun, D. B. M., & Rasch, B. H. (2002). Returning the tables: language affects spatial reasoning. *Cognition*, 84, 155-188.
- Li, P. & Gleitman, L. (2002). Turning the tables: language and spatial reasoning. *Cognition*, 83, 265-294.
- Lupyan, G. (2006). Labels Facilitate Learning of Novel Categories. In *The Sixth International Conference on the Evolution of Language*. Leipzig, Germany.
- Miller, N. E. & Dollard, J. (1941). *Social Learning and Imitation*. New Haven: Yale University Press.
- Pica, P., Lemer, C., Izard, W., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, 306, 499-503.
- Ratterman, M. J. & Gentner, D. (1998). The effect of language on similarity: The use of relational symbols improves young children's performance on a mapping task. In K.D.Holyoak, D. Gentner, & B. Kokinov (Eds.), *Advances in analogy research: Integration of theory and data from the cognitive, computational and neural sciences* (Sophia: New Bulgarian University.
- Roberson, D. & Davidoff, J. (2000). The categorical perception of colors and facial expressions: The effect of verbal interference. *Memory & Cognition*, 28, 977-986.

- Roberson, D., Davidoff, J., & Braisby, N. (1999). Similarity and categorisation: neuropsychological evidence for a dissociation in explicit categorisation tasks. *Cognition*, 71, 1-42.
- Robinson, J. S. (1955). The Effect of Learning Verbal Labels for Stimuli on Their Later Discrimination. *Journal of Experimental Psychology*, 49, 112-114.
- Rossman, I. L. & Goss, A. E. (1951). The Acquired Distinctiveness of Cues - the Role of Discriminative Verbal Responses in Facilitating the Acquisition of Discriminative Motor Responses. *Journal of Experimental Psychology*, 42, 173-182.
- Roy, D. (2005). Grounding words in perception and action: computational insights. *Trends in Cognitive Sciences*, 9, 389-396.
- Rumelhart, D. E., Smolensky, D. E., McClelland, J. L., & Hinton, G. E. (1986). Parallel Distributed Processing Models of Schemata and Sequential Thought Processes. In J.L.McClelland & D. E. Rumelhart (Eds.), *Parallel Distributed Processing Vol II* (pp. 7-57). Cambridge, MA: MIT Press.
- Vygotsky, L. S. (1962). *Thought and Language*. Cambridge, MA: MIT Press.
- Waxman, S. R. & Hall, D. G. (1993). The Development of A Linkage Between Count Nouns and Object Categories - Evidence from 15-Month-Old to 21-Month-Old Infants. *Child Development*, 64, 1224-1241.
- Waxman, S. R. & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology*, 29, 257-302.
- Whorf, B. L. (1956). *Language, Thought, and Reality*. Cambridge, MA: MIT Press.
- Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition*, 85, 223-250.

Yoshida, H. & Smith, L. B. (2005). Linguistic cues enhance the learning of perceptual cues.
Psychological Science, 16, 90-95.

Notes:

1. We define a category as a group of stimuli that evoke a common response. The stimuli typically bear some perceptual or functional relationship to each other. That response may be verbal (calling some colors “blue” despite substantial variation in their hue), or nonverbal, e.g., performing the same action with different-looking objects, or clustering all the blue items together during a sort.
2. A few participants seemed to create more elaborate schemas, e.g., imagining that the ridge on the head created “eyes” on the sides of the ridge, making the category distinction a featural one: items with eyes versus items without eyes. There was no obvious relationship between the use of such strategies and classification performance.
3. Despite the adjective surface property of the labels, the instructions encouraged participants to think of them as referring to kinds rather than properties. To confirm this intuition, we had 13 naïve subjects read the instructions to the *label* condition and answer the question: “Do you think the label “grecious” [“leebish”] refers to a kind of alien or to a property possessed by an alien.” Only 2/13 selected “property” for leebish (significant by two-tailed binomial test, $p < .05$). Five out of 13 ($p = .29$) selected “property” for “grecious” (perhaps due to its unintentional similarity to “gracious”). There was no evidence of a difference in the facilitation effect between the two labels.
4. We thank a reviewer for bringing up the possibility that some combinations of the aliens and explorer motions may have led participants to form incorrect hypotheses such as “the

alien attacks the explorer.” We tested this possibility by performing an ANOVA with alien-motion and direction of approach/escape as within-subject factors. There were no significant main effects, and importantly, the interaction was not significant, $F(3,48)=1.18, p>.3$. Interestingly, we did find a main effect of category-type (approach versus escape). In both experiments, and in all conditions, participants performed better on aliens that were to be approached (Experiment 1: $F(1,42)=4.96, p_{rep}=.94$; Experiment 2: $F(1,71)=41.55, p_{rep}>.99$). The same stimuli were classified more accurately when they were in the *approach* category than when they were in the *escape* category. There are several possible explanations: a simple bias to respond with approach, a deeper encoding of to-be-approached aliens due to the explorer moving towards (thus attracting attention towards them) than when moving away. Most interesting is a possibility that if aliens to be approached were “positive” (many participants described them as “nice” and “friendly”) and aliens to be escaped were “negative” (often described as “unfriendly” and “evil”), then classifying such positive examples may be easier than classifying negative ones.

Figure Captions

Figure 1. The two categories learned by the participants. The stimuli on the left have flatter bases and a subtle ridge on their “head.” The stimuli on the right have more rounded bases, and a smoother head.

Figure 2. Experiment 1 mean classification accuracy for training (left) and testing (right). Bars show one SE of the mean.

Figure 3. Mean classification accuracy for training for the four conditions of Experiment 2.

Figure 1

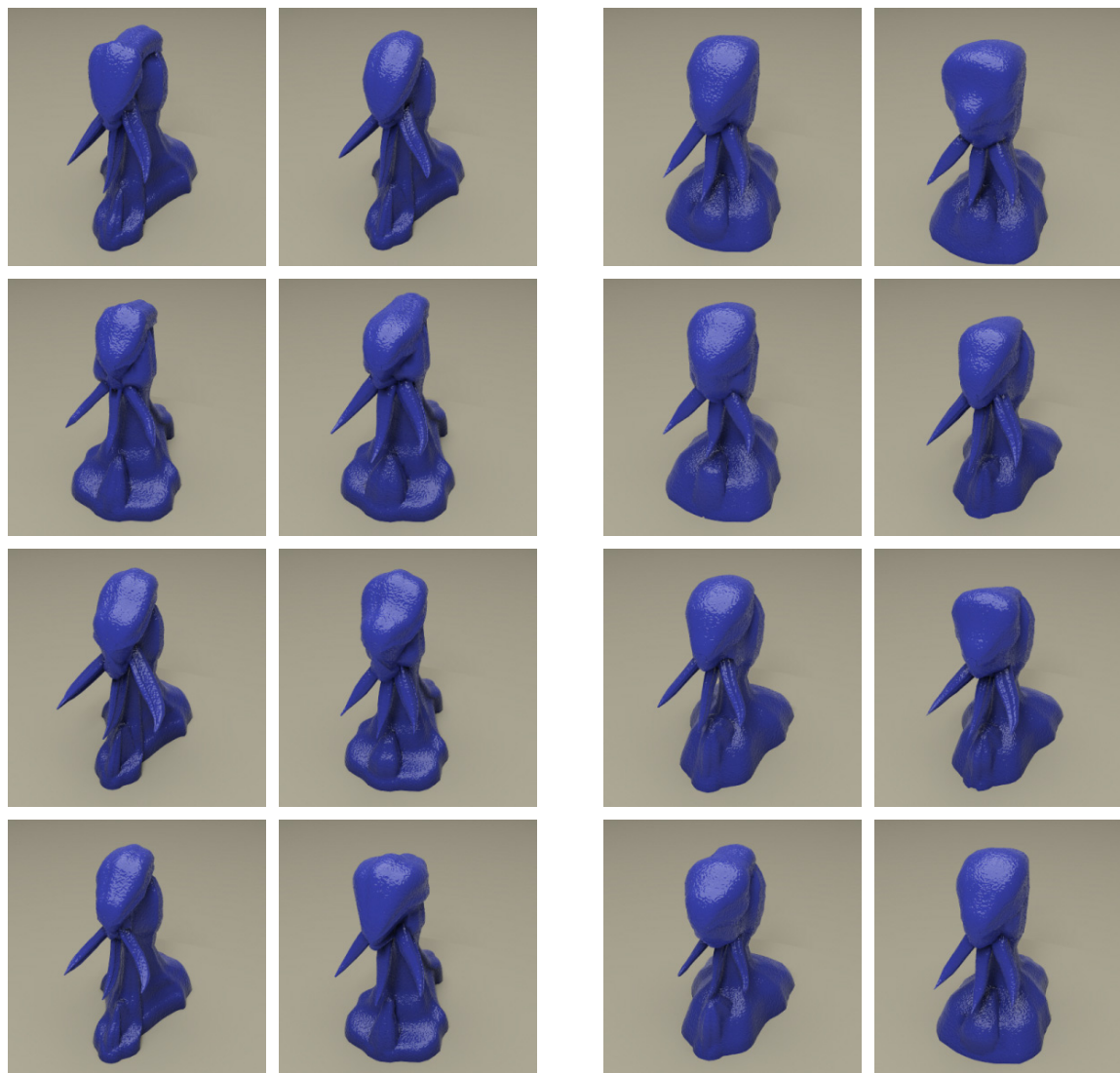


Figure 2

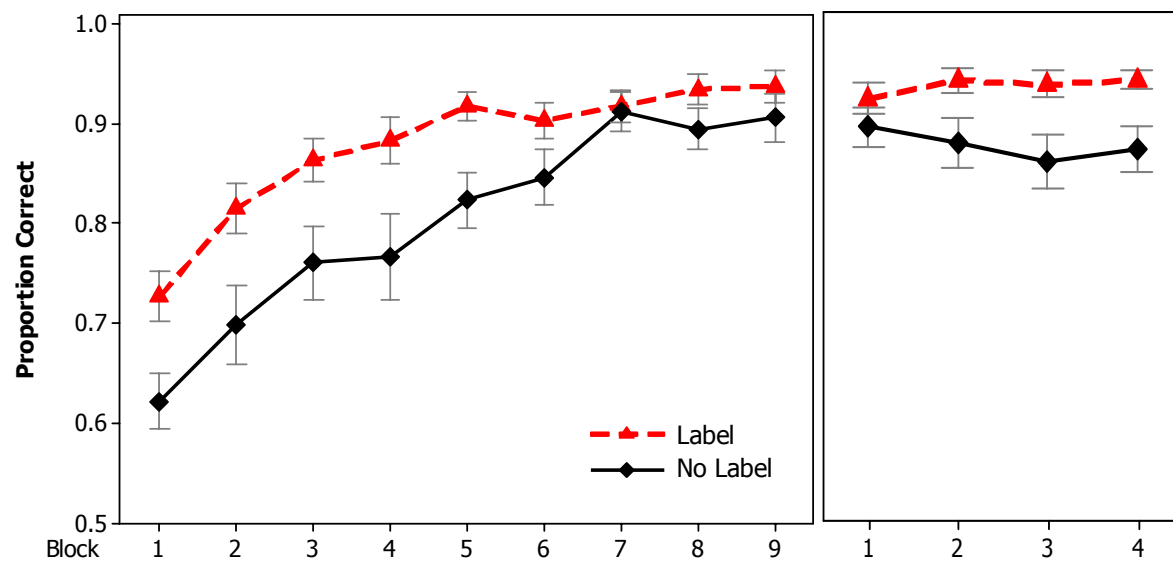


Figure 3

