Research Article

Language Is Not Just for Talking

Redundant Labels Facilitate Learning of Novel Categories

Gary Lupyan,[1,2] David H. Rakison,[1] and James L. McClelland[3]

[1]Department of Psychology, Carnegie Mellon University; [2]Center for the Neural Basis of Cognition, Pittsburgh, Pennsylvania; and [3]Department of Psychology, Stanford University

ABSTRACT—*In addition to having communicative functions, verbal labels may play a role in shaping concepts. Two experiments assessed whether the presence of labels affected category formation. Subjects learned to categorize "aliens" as those to be approached or those to be avoided. After accuracy feedback on each response was provided, a nonsense label was either presented or not. Providing nonsense category labels facilitated category learning even though the labels were redundant and all subjects had equivalent experience with supervised categorization of the stimuli. A follow-up study investigated differences between learning verbal and nonverbal associations and showed that learning a nonverbal association did not facilitate categorization. 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 allows for linguistic communication, and it has been argued that words stabilize abstract ideas in working memory and make them available for inspection (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? Although nonhuman animals categorize nonlinguistically, humans have the potential benefit of labels—category names—as they decide to which category an object belongs. The crucial question, then, is how the presence of labels affects human categorization.

Investigations of the role of names in categorization are related to investigations of how categorization affects 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 of a bias in similarity ratings due to knowing that two objects have different labels and thus belong to different categories (Goldstone, Lippa, & Shiffrin, 2001). Though this line of research is relevant to the present study, the question addressed by the present study is whether category labels facilitate category formation when categorization experience is controlled.

The specific question of how labels affect category learning has been most thoroughly investigated in two contexts. First, work on children's language acquisition has revealed that infants as young as 9 months of age more readily individuate labeled than unlabeled objects (Xu, 2002), and that by 12 months of age, infants have an expectation that words refer to object categories (Waxman & Hall, 1993; Waxman & Markow, 1995). Waxman and Markow argued that words serve as "invitations to form categories" and that superordinate labels, such as "vehicle," lead children to form the appropriate category. Labels that are correlated with regularities in the world (e.g., shape is predictive of solid, but not nonsolid, categories) have been shown to improve the learning of these regularities (Yoshida & Smith, 2005).

Address correspondence to Gary Lupyan, 342C Baker Hall, Carnegie Mellon University, Pittsburgh, PA 15213, e-mail: glupyan@cnbc.cmu.edu.

[1]We define a category as a group of stimuli that evoke a common response. That response may be verbal (e.g., 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). The stimuli within a category typically bear some perceptual or functional relationship to each other.

In experiments such as those conducted on infants by Waxman and her colleagues, the labels are semantically empty. A different sort of labeling influence can be seen in older children. For instance, learning to associate small, medium, and large groupings with the labels "baby," "mommy," and "daddy," respectively, facilitates relational judgments, enabling children to transfer the size relation to novel stimuli; in this case, the effect of the labels is tied to their semantics (Kotovsky & Gentner, 1996; Ratterman & Gentner, 1998).

A second line of research on how labels affect category learning builds on James's (1890) *law of dissociation by varying concomitants*—the idea that associating A with B on one occasion and with C on another leads to A becoming dissociated from both B and C, and thereby becoming a more abstract object. Extending James's reasoning, Miller and Dollard (1941) hypothesized that associating different responses with otherwise similar stimuli should increase the perceived difference between the stimuli. Miller and Dollard saw object names as a kind of motor response, and so hypothesized that associating undifferentiated stimuli with different names increases the differences between the stimuli and facilitates placing them into separate categories. This 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 for us to be clear about what we mean by the term *label*. We use the term to refer to anything that is (a) consistently correlated with a category and (b) used to refer to a category. The category can comprise objects, sounds, actions, spatial relations, and so on. 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, whereas the subjects in the study reported here, 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 people 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 rather whether labeled categories are easier to acquire than unlabeled categories *because* they have a name—even when categorization can be performed without relying on labels. Thus, the two experiments we report in this article are the first to directly test the idea that labels make category differences "more concreted" (James, 1890, p. 333). Our experiments address two main questions: Are labeled categories easier to learn than unlabeled categories even when the labels are entirely redundant, contributing no additional information? How does associating stimuli with verbal labels compare with learning a nonlinguistic category association?

EXPERIMENT 1

In the first experiment, we measured performance of subjects learning to associate novel objects with behavioral responses. Some subjects performed this task while learning names for the stimulus categories, and others did not learn names. We expected that if it is easier to learn named than unnamed categories, performance would be superior in the former condition.

Method

Subjects

Forty-eight Carnegie Mellon University undergraduates (ages 18–24) participated in the experiment for course credit. The subjects were randomly assigned to *label* and *no-label* groups. Data from 4 subjects were excluded because they did not follow instructions. Data for the test phase of the experiment were not available for 2 subjects because of experimenter error.

Materials

The stimuli were a subset of the YUFO stimulus set (Gauthier, James, Curby, & Tarr, 2003). Items in one category (shown on the left in Fig. 1) had flatter bases and a subtle ridge on their "heads." Items in the other category (shown on the right in Fig. 1) had more rounded bases and smoother heads. Subjects' responses on a questionnaire following the category training

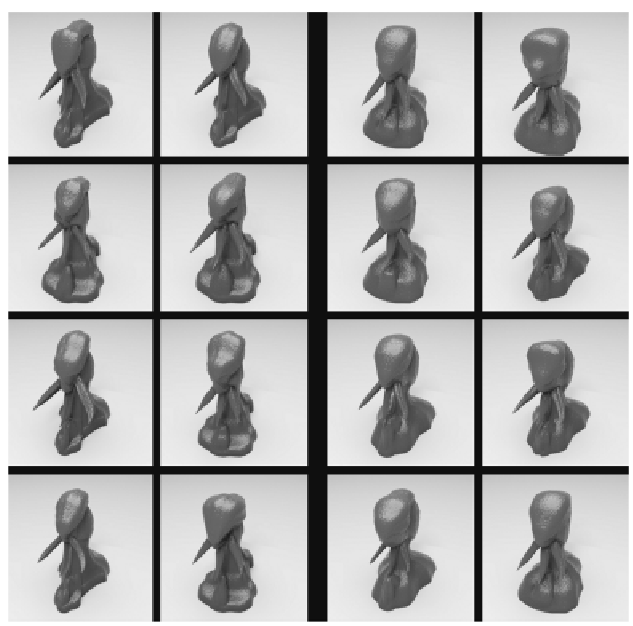

Fig. 1. The two categories learned by the subjects. The stimuli on the left have flatter bases and a subtle ridge on their "head"; the stimuli on the right have more rounded bases and a smoother head.

indicated that they were overwhelmingly more attuned to the distinction in the heads than to the distinction in the bases; subjects used adjectives like "pointy" versus "fat" and "bumpy" versus "smooth" when asked to verbalize the difference between the two categories.

The stimuli were presented on a black background on a 17-in. computer 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.[2]

Training Procedure

Subjects were told to imagine that they were explorers on another planet and were 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, 1 of the 16 aliens appeared in the center of the screen. After 500 ms, an outline of a character in a space suit (the "explorer") appeared in one of four positions—to the left of, to the right of, above, or below the alien. Subjects were instructed to respond with the appropriate direction key depending on the category of the alien. For instance, if the explorer appeared above the alien, they needed to press the "down" key to move toward the alien or the "up" key to move away; after the key press, the explorer moved toward or away from the alien, as indicated. Auditory feedback—a buzz for an incorrect response and a bell for a correct response—sounded 200 ms after the explorer stopped moving. In the label condition, a printed label ("leebish" or "grecious") appeared to the right of the alien 300 ms after the feedback. After another 1,500 ms, the alien (and label, in the label condition) disappeared from the screen, and a fixation cross marked the start of the next trial. The total trial duration and exposure to the stimulus were equal for the two conditions.

The pairing of the labels with the categories (move away vs. move toward) and with the perceptual stimuli (left vs. right side of Fig. 1) was counterbalanced across subjects. Subjects in the label condition were told that previous visitors to the planet had found it useful to name the two kinds of aliens, and that they should pay careful attention to the labels. All subjects received the same number of categorization trials (nine blocks of 16 trials each) and had equal exposure to the stimuli. The only difference between the two conditions was whether or not a verbal label appeared after each response.

[2]Although the labels may appear to have surface properties typical of adjectives, the instructions encouraged subjects to think of the labels as referring to kinds, rather than properties. To confirm the intuition that the context of the instructions encouraged participants to treat the labels as referring to kinds, we had 13 naive subjects read the instructions for the label condition and answer the following question: "Do you think the label 'grecious' ['leebish'] refers to a kind of alien or to a property possessed by an alien?" Only 2 of the 13 subjects answered that "leebish" referred to a property (significant by two-tailed binomial test, $p < .05$). Five of the 13 subjects ($p = .29$) indicated that "grecious" referred to a property (perhaps because of this label's unintentional similarity to "gracious"). There was no evidence that the two labels differed in their ability to facilitate categorization performance.

Testing Procedure

Following the training trials, subjects completed a test phase. On each test trial, one of the aliens appeared in the center of the screen, and the task was to classify it as a kind to be approached or escaped from. No feedback was given, and the names learned in the label condition were not presented. So that we could determine whether subjects had formed category representations or just memorized specific examples, the test stimuli included not only previously categorized stimuli, but also previously unseen stimuli from the learned categories. There were four blocks of 24 trials each (16 old stimuli plus 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 condition (label vs. no-label) as a between-subjects factor and block as a within-subjects factor. Initial analyses revealed that there were no differences between different stimulus-label, label-response, and stimulus-response pairings, all Fs < 1, so these factors were collapsed for subsequent analyses. Performance improved over time, $F(8, 42) = 33.12$, $p < .001$, $\eta_p^2 = .45$, with final performance in both conditions reaching asymptote near the 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 × Block interaction, $F(8, 336) = 2.59$, $p < .05$ (Fig. 2). For instance, subjects in the no-label group reached the 80%-correct level of performance after approximately 72 trials; subjects in the label group reached the same level of performance after approximately 30 trials.

Reaction times did not differ between the conditions (label: $M = 1{,}229$ ms, $SD = 254$ ms; no-label: $M = 1{,}178$ ms, $SD =$

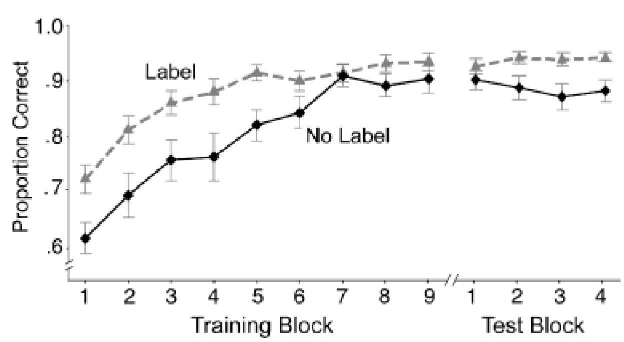

Fig. 2. Mean classification accuracy in the training and test phases of Experiment 1. Error bars indicate standard errors of the means.

270 ms), $F(1, 42) < 1$; reaction times also did not show a significant Condition × Block interaction, $F < 1$.

Test Phase

In the test phase immediately following training, no feedback or labels were provided. Overall, generalization performance was far above chance ($M = .86$, $SD = .11$), though below the level of performance for the previously seen items ($M = .92$, $SD = .08$), $F(1, 40) = 15.41$, $p_{rep} > .99$. For the last three blocks of training, performance in the label condition and performance in the no-label condition were statistically indistinguishable. When response feedback was removed in the test phase, performance became significantly higher in the label condition ($M = .93$, $SD = .04$) than in the no-label condition ($M = .87$, $SD = .10$), $F(1, 40) = 7.01$, $p_{rep} = .96$ (Fig. 2). Reaction times did not differ between conditions (label: $M = 904$ ms, $SD = 332$ ms; no-label: $M = 875$ ms, $SD = 253$ ms), $F(1, 40) < 1$.

An important goal of the test phase was to see whether representations formed in the presence of labels were more robust than representations formed without labels. Analyses of the responses to the novel stimuli indicated that they were. Generalization performance improved with time among subjects in the label condition, $F(3, 60) = 3.79$, $p < .05$, but not among subjects who were not exposed to labels, $F(3, 60) < 1$. Thus, subjects who previously learned labels were apparently better able to generalize the learned categories to unseen items even without feedback or labels to guide them. There is a suggestion that generalization performance of the label group was superior to that of the no-label group, though the Condition × Block interaction did not reach significance, $F(3, 120) = 1.63$.

In Experiment 1, associating initially meaningless labels with novel categories facilitated categorization. This was true even though the names provided no additional category information. Subjects who learned the categories in the context of labels not only learned to categorize the two families of aliens more quickly than subjects who were not given labels, 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 attracted subjects' attention to parts of the stimuli that were particularly informative about the category distinction. In Experiment 2, we addressed this possibility by introducing an *auditory-label* condition that equated the visual information presented in the label and no-label conditions. Second, it was unclear whether the facilitation observed in Experiment 1 was specific to labels or whether the same type of facilitation could be produced by learning any additional association. 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 would be expected to 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. In Experiment 2, to test whether the effect was specific to labels, we included a *location* condition in which the stimuli were associated with locations, rather than verbal labels.

Method

Subjects

Seventy-five Carnegie Mellon University undergraduates (ages 18–24) participated in the experiment for course credit. The subjects were randomly assigned to four conditions: no-label ($n = 22$), written-label ($n = 18$), auditory-label ($n = 18$), and 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 recorded sound clips of a female saying "leebish" and "grecious." In the location condition, subjects were told that some aliens lived on one side of the planet, and others lived on the other side. On each trial, after the subject responded (approach/escape) and auditory feedback was given, the alien moved up or down to signal where it "lived." The motion started 300 ms after response feedback and lasted approximately 400 ms. The trial ended 1,300 ms after the alien stopped moving. Thus, the alien was visible for a longer total time in the location condition compared with the label conditions. Finally, the test phase was omitted from this experiment.

To measure the degree to which subjects learned the association between stimuli and labels or locations, we included verification trials as part of the training procedure. Verification trials were presented after a random 10% of training trials. On each verification trial in the label conditions, one of the aliens appeared with a query asking: "Is this one leebish [grecious]? yes/no" (the label was randomly selected). On the verification trials in the location condition, the alien moved up or down, and subjects responded to the query, "Is this correct? yes/no"; subjects 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 condition as a between-subjects factor and block as a within-subjects factor showed a significant main effect of condition, $F(3, 71) = 3.18$, $p < .03$, $\eta_p^2 = .08$, and a

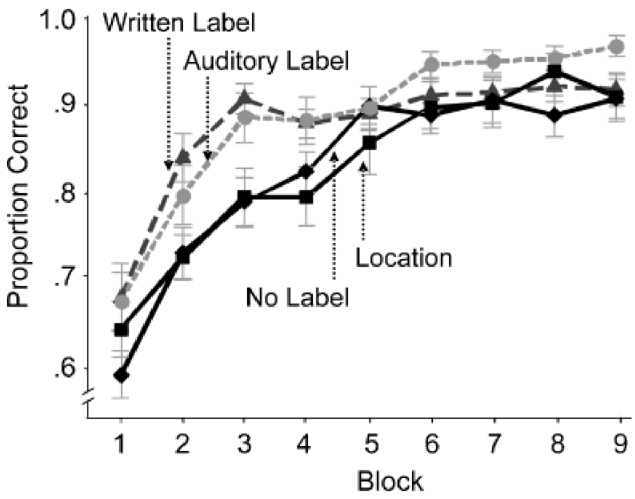

Fig. 3. Mean classification accuracy during training in the four conditions of Experiment 2. Error bars indicate standard errors of the means.

Block × Condition interaction, $F(24, 568) = 1.66$, $p < .01$, $\eta_p^2 = .07$ (Fig. 3). Planned comparisons (simultaneous Tukey tests with 95% confidence intervals) showed that the two label conditions did not differ from each other, $F(1, 34) = 0.26$, and that the no-label condition did not differ from the location condition, $F(1, 37) = 0.0$. An ANOVA comparing the pooled data from the two label conditions and the pooled data from the no-label and location conditions revealed a significant main effect of condition, $F(1, 73) = 11.10, p_{rep} > .99, \eta_p^2 = .15$, and a Condition × Block interaction, $F(8, 584) = 2.99$, $p < .001$, $\eta_p^2 = .04$. A direct comparison of the written-label condition with the location condition revealed a significant main effect of condition, $F(1, 33) = 4.11$, $p_{rep} = .92$, $\eta_p^2 = .08$, and a significant Condition × Block interaction, $F(8, 264) = 3.58$, $p < .001$, $\eta_p^2 = .09$. As in Experiment 1, labels improved categorization. This improvement was seen for both written and spoken labels, relative to both the no-label condition and the location condition, in which subjects learned to associate alien categories with locations rather than nonsense words.[3]

[3]We thank a reviewer for bringing up the possibility that some combinations of the alien's motion and the explorer's motion may have led subjects to form incorrect hypotheses such as "the alien attacks the explorer." We tested this possibility by performing an ANOVA with alien motion and category (approach vs. escape) as within-subjects factors. There were no significant main effects, and most important, the interaction was not significant, $F(3, 48) = 1.18, p > .3$. Interestingly, we did find a main effect of category type. In both experiments, and in all conditions, subjects performed better on aliens that were to be approached than on aliens that were to be avoided, $F(1, 42) = 4.96, p_{rep} = .94$, in Experiment 1 and $F(1, 71) = 41.55$, $p_{rep} > .99$, in Experiment 2. There are several possible explanations for why the same stimuli were classified more accurately when they were in the "approach" category than when they were in the "escape" category. There may have been a simple bias to respond with "approach," or encoding may have been deeper when the explorer was moving toward the aliens (thus attracting attention toward them) than when the explorer moved away. Given that aliens to be approached were often described in positive terms (many subjects described them as "nice" and "friendly") and aliens to be escaped were described in negative terms (often described as "unfriendly" and "evil"), these results suggest that classifying positive examples may be easier than classifying negative ones.

Overall, individual participants' verification accuracy was correlated with training accuracy, $r(50) = .34$, $p = .02$. The three-way comparison of the location, written-label, and auditory-label conditions, with verification accuracy as the dependent variable, was significant, $F(2, 50) = 5.28$, $p < .01$, $\eta_p^2 = .17$. Planned comparisons revealed that accuracy was significantly greater in the auditory-label condition ($M = .92$, $SD = .07$) than in the location and written-label conditions, but the latter two conditions did not differ (both Ms = .82, SDs = .13). Although spoken labels were learned better than written labels, there was no difference in categorization performance between the written- and auditory-label conditions. At the same time, subjects learned where the aliens lived just as well as they learned the written labels, yet categorization was facilitated by learning the labels, but not the locations.

GENERAL DISCUSSION

At the outset of this study, we posed two main questions: Are labeled categories easier to learn than unlabeled categories? How does associating stimuli with verbal labels compare with learning a nonlinguistic category association? Given equal experience categorizing novel stimuli, subjects 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 more than did learning nonverbal associations. There is also some evidence that categories learned with labels were 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 study (Lupyan, 2006), 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. These results further suggest that learning labels results in more robust category representations.

These findings provide empirical support for the idea that words do more than communicate information between individuals (Clark, 1997; James, 1890; Whorf, 1956) and bear directly on the language-and-thought debate (e.g., Gentner & Goldin-Meadow, 2003; Gleitman & Papafragou, 2005). Our results show 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 the plants 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 a classification advantage.

The categories used in the current experiments were by no means incommensurate with linguistic coding of English-speaking adults. Indeed, many subjects in the no-label condition reported generating their own labels during the learning task. It remains to be seen whether labeling can produce a still greater facilitation in category learning when 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 subjects to learn a subtle perceptual distinction on the basis of experience with individual category exemplars. In this context, the labels can be thought of as providing perceptually simple correlates to an otherwise perceptually complex distinction. That is, the labels may have allowed subjects to more easily represent the somewhat fuzzy perceptual distinction between categories ("more rounded and smooth" vs. "less rounded, with ridges") in terms of a simpler verbal distinction ("leebish" vs. "grecious"). This general account that naming a category causes items within that category to cohere because the name serves as a reliable cue 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. According to this account, as a label is paired with individual exemplars, it becomes associated with features most reliably associated with the category. When activated, it then dynamically creates a more robust category attractor. It is unclear whether once formed, category attractors continue to depend, in part, on labels. Subjects in the label condition of Experiment 1 continued to show facilitation when labels were removed, but this may have been due to continued implicit labeling.

If the category representations responsible for superior performance when categories are named depend on on-line activation of verbal representations, then somehow inactivating the labels may cause categorization impairments. This possibility is supported by two sources of evidence. First, linguistic impairments have long been implicated in categorization impairments (Goldstein, 1948). Some types of aphasia are reliably associated with nonlinguistic categorization impairments, and there is a correlation between naming and categorization deficits, even in the absence of comprehension (i.e., semantic) impairments (De Renzi, Faglioni, Savoiardo, & Vignolo, 1966; Hjelmquist, 1989; Roberson, Davidoff, & Braisby, 1999). Second, in normal adults, verbal, but not visual, interference eliminates the cross-category advantage in the domains of colors and facial expressions (e.g., Gilbert, Regier, Kay, & Ivry, 2006; Roberson & Davidoff, 2000).

A remaining question is why learning the associations between stimuli and nonverbal cues in Experiment 2 ("lives above" vs. "lives below") failed to facilitate performance. According to the simple account that labels augment otherwise fuzzy categories with discrete, perceptually simple features, one might have expected that associating the stimulus categories with these arguably equally discrete and simple nonverbal cues would have similarly facilitated category learning. One explanation for not finding a corresponding facilitation from learning the nonverbal association may be that adults have an expectation that words, but not facts (e.g., "lives above" and "lives below"), refer to shape-based object categories (Colunga & Smith, 2005). It remains to be seen whether nonword labels used in an explicitly referential context can facilitate categorization. It is also possible that the advantage for labels may have been due to differences in inner rehearsal; that is, participants may have rehearsed labels more effectively than the up/down distinction.

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

Acknowledgments—This work was supported in part by a National Science Foundation Graduate Fellowship to the first author and by National Institute of Mental Health Program Project Grant MH64445. We would like to thank Brian Mathias and Ashleigh Molz for their help with data collection and Robert Goldstone, Eliana Colunga, and an anonymous reviewer for insightful commentary and suggestions.

REFERENCES

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*, 273–275.

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 W.D. Gray & C.D. Schunn (Eds.), *Proceedings of the 24th Annual Meeting of the Cognitive Science Society* (pp. 136–141). Mahwah, NJ: Erlbaum.

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.

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., Savoiardo, 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, USA*, *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, England: Cambridge University Press.

Goldstein, K. (1948). *Language and language disturbances*. New York: Grune & Stratton.

Goldstone, R.L. (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 A. Cangelosi, A.D.M. Smith, & K. Smith (Eds.), *The Sixth International Conference on the Evolution of Language* (pp. 190–197). Singapore: World Scientific.

Miller, N.E., & Dollard, J. (1941). *Social learning and imitation*. New Haven, CT: 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* (pp. 274–282). Sofia, Bulgaria: 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.

(Received 1/4/07; Revision accepted 3/26/07;
Final materials received 3/28/07)