Title
Social Context Effects on the Impact of Category Labels

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 32(32)

ISSN
1069-7977

Authors
Stephens, Rachel
Prefors, Amy
Navarro, Daniel

Publication Date
2010

Peer reviewed

Social Context Effects on the Impact of Category Labels

Rachel G. Stephens (rachel.stephens@adelaide.edu.au)
Amy Perfors (amy.perfors@adelaide.edu.au)
Daniel J. Navarro (daniel.navarro@adelaide.edu.au)
School of Psychology, University of Adelaide
Adelaide, SA 5005, Australia

Abstract

We explore whether social context affects how labels (relative to other features) affect category learning. We taught 104 participants four novel categories using a feature inference task. In a between-participants design, we manipulated: 1) the social context of the task (social context vs. on the computer); and 2) which dimension of the category members could be used to perfectly predict the target feature: the category label, a biased feature (which is salient and already associated with the target feature in the correct way) or a non-biased feature (which is less salient and not already associated with the target feature in any way). Learning curves were used to assess whether participants assumed that labels were uniquely helpful compared to other features. The results suggest that the extent to which labels are privileged depends on the context in which the category learning task is presented. When the task is social, people learn quickly regardless of whether a label or another feature is the most informative. When the task is not, both novel labels and biased features are more useful than non-biased features.

Keywords: categorization; feature inference; labels; features; social context.

Introduction

There is a Chinese proverb that says, “The beginning of wisdom is to call things by their right names.” Category labels feel like an important and special part of our conceptual knowledge. We need labels to communicate about classes of objects, and labels (often unlike other features) are a property that all members of a category share. One might expect that labels help learners pick out the category members that have important similarities to each other, and that are different from members of other categories. However, although this has been the topic of study for decades, it remains unclear whether there is a psychological distinction between category labels and other types of features. It is also unknown (especially for adults) whether the effect of a label is affected by the social-referential context in which it is offered.

Work with children broadly supports the notion that labels have a privileged psychological status, although it is still debated what the source of that privilege is. Verbal labels appear to facilitate infant category learning (e.g., Balaban & Waxman, 1997), with shared names highlighting commonalties between objects (Waxman & Braun, 2005). Labels influence the number of categories formed by infants, overriding the categories that are suggested by perceptual similarity (Plunkett, Hu & Cohen, 2008). Additionally, when making decisions about whether a category feature can be generalized to a new object, preschool children rely more on category membership conveyed by a label than they do on perceptual similarity (Gelman & Markman, 1986; 1987). These experiments suggest that labels are special in some way, but these studies do not address whether a single salient feature possessed by all category members might produce the same effects.

The question is further complicated by the fact that, for children at least, the social and linguistic context influences the effects of category labels. Fulkerson and Haaf (2003) found that labels can help infants to form categories that are otherwise *not* formed when only a non-labeling sound or no sound is used in their place. However, for older infants (15 rather than 9 months) the source of the label matters: these infants formed categories when the labels were presented orally, but not when they were presented by a voice recorder. Consistent with this, Campbell and Namy (2003) found that infants learned object names only when the label was presented in a normal social-referential interaction. The names were learned when the label was verbalized by the experimenter and embedded in a familiar naming routine, such as, “Look at what you have! *Tillen*. That’s what we call that one.” Learning was unsuccessful when the label was emitted from a baby monitor and was not timed with the naming routine.

It is unclear whether we should expect similar effects of social context in adults, or whether any differences between adults and children are due to differences in the social context of label presentation. Unlike for children, most adult category learning experiments do not incorporate a social element: category labels are presented in written form on screen or on paper, or (at most) as a recorded sound. Grice’s conversational maxims (Grice, 1975) suggest that labels presented in a social context should especially be presumed to be relevant and informative to the task at hand. Alternatively, since educated adults are well practiced at using labels, social context may not significantly change how labels are treated.

Labels do seem to play an important role in adult categorization. Categorical perception research suggests that learning that stimuli share a label can be sufficient to increase perceptions of similarity of the stimuli (e.g., Goldstone, Lippa & Shiffrin, 2001). In addition, Yamauchi and Markman (1998; 2000a) have begun to directly address the issue of whether category labels have a privileged status over other features for adults. They claimed that people employ different strategies to make feature inferences or classifications (i.e., infer labels), and that learning novel categories through a classification task or a feature inference task will produce different category representations. However, the experimental tasks used were unfair as a test of a general distinction between labels and other features, because the category structures of the classification and inference tasks were not equivalent: labels were a diagnostic feature in the inference task, which seemed to drive the differences between conditions (see also Johansen & Kruschke, 2005). However, the studies could suggest that people expect labels to be useful in a feature inference task (see also Yamauchi, Love & Markman, 2002).

In a further series of experiments Yamauchi and Markman (2000b) showed adults a set of labeled exemplars of two categories and asked them to compare novel stimuli to the exemplars. Classifications of the novel stimuli were generally made according to the total number of features consistent with the appropriate category prototype, but feature inferences for novel stimuli were strongly influenced by the observed category label. As a result, when similarity and category membership were placed in opposition, participants were more likely to base their inferences on the label. This effect was decreased when the labels referred to a feature rather than to category membership, or when the label was replaced by a perceptual feature. These studies suggest that to the extent that labels convey category membership, they are privileged over other features. However, the experimental tasks used in these experiments were fairly unnatural, since participants did not have to learn the categories: they simply compared stimuli on a sheet in front of them.

This work has two aims. First, it contributes one of the first explorations in the adult literature focused on the question of whether social context has an impact on the status of labels. Second, it investigates whether labels have a privileged status over other category features for adults. Are people biased to assume that labels are uniquely helpful compared to other stimulus features when learning about novel categories? If so, they should assume that labels are important to pay attention to and therefore categories should be easier to learn when the labels are useful predictors. However, categories should be more difficult to learn when another feature is the more useful predictor (depending on the type of feature). Are these effects mitigated or amplified depending on the nature of the social context in which the labels are presented? Are labels assumed to be especially important in a social category learning context, involving communication with a knowledgeable human teacher?

Method

Participants learned about four novel categories during a feature inference task. Two between-participants experimental factors were manipulated to form a 3x2 design. The first factor was which aspect or dimension of the category members could be used to perfectly predict the target feature. This diagnostic feature dimension could be the CATEGORY LABEL, a BIASED FEATURE or a NON-BIASED FEATURE. The second factor was whether the category learning task was performed alone on a personal computer (PC), or in a more social context with the experimenter (SOCIAL).

Figure 1: Two sample images used in the category learning task. The image on the right includes the feedback of the hammer.

Participants

106 adults (either undergraduates at the University of Adelaide, or people recruited from the general community; 41 males) took part in the experiment. Ages ranged from 18 to 57 years. They received course credit or AU$10. One participant's data was removed from the SOCIAL, NON-BIASED FEATURE condition because the participant withdrew from the study before training was completed. An outlier was removed from the PC, LABELS condition for taking 26 blocks to complete the training task (this was more than 3 SDs above the mean for that condition). 16 to 18 people remained in each condition.

Materials

The category learning task was designed to be realistic and engaging, in order to encourage ecologically valid responses. Participants learned about four novel "alien people" categories, each of which contained four members. Images for the categories were created using *World of Warcraft*, an online computer game produced by Blizzard. Examples of the images are shown in Figure 1, and the category structure used across all conditions is shown in Table 1.

Participants were asked to predict the nature of a certain target feature: which item each alien wanted to buy (options were a timber axe, a dagger, a hammer or a staff). The four category members varied on five dimensions, which could each take one of four values. The five dimensions and their possible values were:

1) category labels, presented as community names: Goloth, Bragen, Lathor and Durgal
2) clothing: leather warrior-like garb, a robe, tradesperson-like overalls and "lumberjack" attire
3) hair style: long, cropped, bald and ponytail
4) skin color: red, cream, brown and blue-grey
5) facial hair: short square beard, long plaited beard, medium pointed beard and broad beard with upturned moustache.

Table 1: Category structure for the feature inference learning task, for all conditions. The diagnostic feature dimension perfectly predicts the target feature dimension, and demarcates the four categories. (F = feature)

Target features	Diagnostic features	F1	F2	F3	F4
Timber axe	1	1	3	2	1
	1	1	1	3	2
	1	2	1	1	3
	1	3	2	1	1
Dagger	2	2	4	3	2
	2	2	2	4	3
	2	3	2	2	4
	2	4	3	2	2
Hammer	3	3	1	4	3
	3	3	3	1	4
	3	4	3	3	1
	3	1	4	3	3
Staff	4	4	2	1	4
	4	4	4	2	1
	4	1	4	4	2
	4	2	1	4	4

As Table 1 demonstrates, four of the feature dimensions contributed to a family resemblance category structure, but one "best predictor" or diagnostic dimension perfectly predicted the target feature. Thus all conditions had rule-based categories: only a single dimension was needed to solve the categorization problem. One experimental factor was *which* dimension was the best predictor of the target feature. In the LABEL conditions, the community name was the diagnostic dimension. Thus participants could learn to perfectly predict which item an alien wanted to buy using only its community name. Alternatively, in the BIASED FEATURE conditions, the clothing was the diagnostic dimension. In these conditions, the clothing "value" corresponded to the target item one might expect based on prior background knowledge: the aliens wearing the robe wanted the staff, the aliens with the leather garb wanted the dagger, the aliens in the overalls wanted the hammer, and the aliens with the "lumberjack" attire wanted the timber axe. Finally, in the NON-BIASED FEATURE conditions, the facial hair was the diagnostic dimension, arbitrarily matched with the target items.

Procedure

Participants were randomly allocated to one of the six conditions. All participants were asked to imagine they were a space traveler who began working in a general store on another planet and needed to learn about the customers of the store. Participants were told that they needed to learn to predict which item each of 16 customers wanted to buy. They learned by trial and error, and the learning task continued until they made the correct prediction for all 16 customers. This criterion was chosen to encourage optimal performance: participants knew that the task would continue until no errors were made.

On each block the 16 trials were presented in random order. For each trial, participants were presented with an image of a customer with a label. In the SOCIAL conditions, the experimenter displayed a card with the image and verbally presented the label (e.g., "This is a Goloth"). In the PC conditions, the label was written in bold blue capital letters above the image on the screen. Participants were then asked to predict the target feature, either verbally to the experimenter in the SOCIAL conditions, or by clicking the appropriate button on the screen in the PC conditions. They were given immediate corrective feedback after each trial consisting of an image of the customer holding the correct item, with the community name written above the image and the name of the correct target item below the image. Participants also received additional feedback after each block of 16 stimuli about their total number of correct responses for that block.

Design

There were two factors of interest in this experiment. One factor was whether the target feature dimension was best predicted by: 1) the community LABEL; 2) the NON-BIASED feature (facial hair), which participants should not have expected to be useful *a priori*; or 3) the BIASED feature (clothing), which people should have had a prior bias to find useful for predicting what the creatures wanted to buy, since the sets of clothing each corresponded to the appropriate target item. This experimental factor tests whether people are biased to assume that labels are uniquely useful features, or whether they are similar to highly salient or useful features (like the BIASED feature). In each level of the factor, the diagnostic feature dimension plays an identical role in the category structure, allowing a fair test of the relative status of labels and features. If labels are not special, all else being equal, learning performance should be equivalent regardless of whether labels or other features are the diagnostic dimension. However, if labels have a special status, participants should be quicker to learn to predict the target feature when the LABEL is the diagnostic dimension. Learning should be slowest in the NON-BIASED FEATURE conditions, where labels are less useful and an unexpected feature *is* useful. Learning should be more rapid in the BIASED FEATURE conditions, where an unsurprising feature *is* useful, and it is relatively easy to remember which particular feature value (i.e., particular outfit) matches with each target item. Of critical importance, then, is whether learning in the LABEL conditions is closer to learning in the BIASED or in the NON-BIASED FEATURE conditions.

The other experimental factor was whether the labels were presented in a social context or not. In the PC conditions, participants worked on a computer; the labels and images were presented on the screen. In the SOCIAL conditions, participants learned by interacting with the experimenter, who presented the images on cards and verbally presented the category labels. This manipulation tests whether people assume labels to be particularly special when the context is more social and interactive.

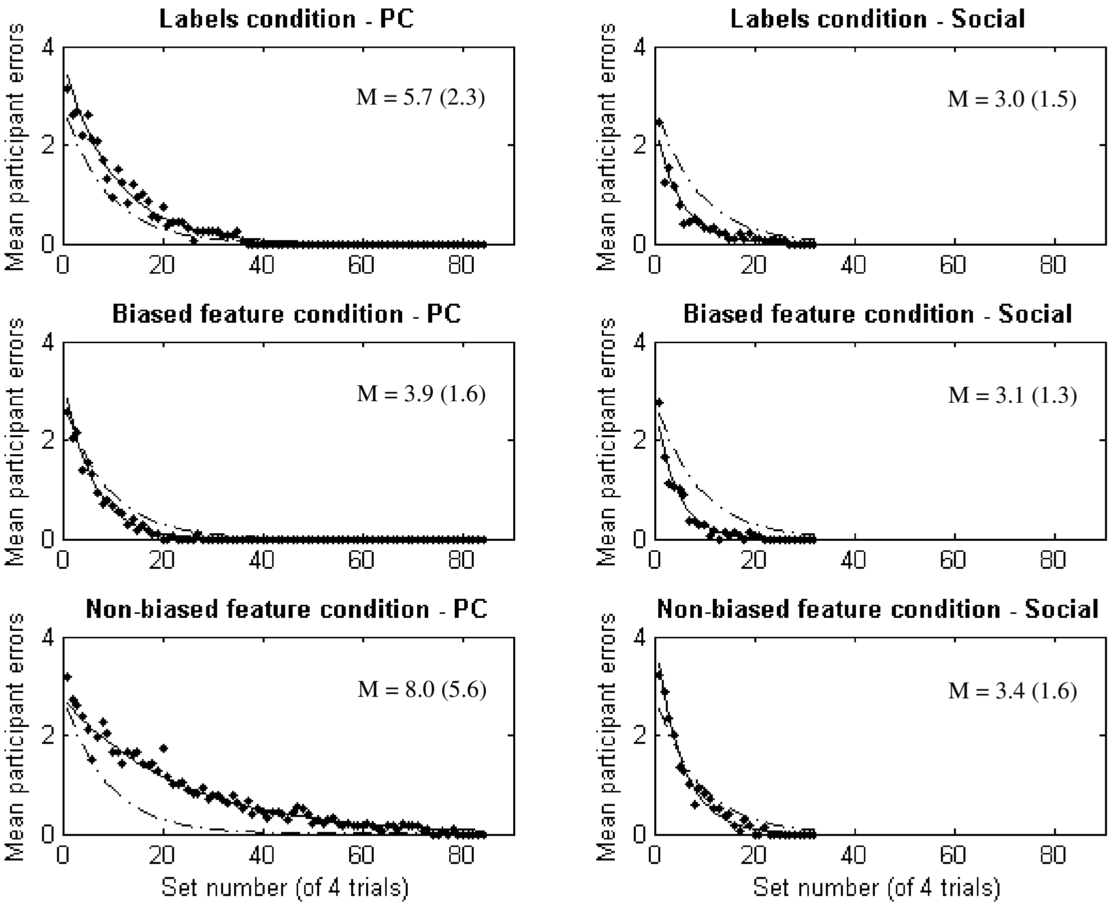

Figure 2: Learning curves averaged across sets of 4 trials, and averaged across participants in each condition. Two models using an exponential function are fit to the error data using BIC: a model with unique parameter values for each condition (solid line) is preferred over a model with the same two parameter values for all conditions (dashed line). M = the mean (SD) number of blocks taken to complete the learning task.

Results

Our results suggest that the extent to which labels are privileged depends on the context in which the task is presented. When the task is social, people learn quickly regardless of the nature of the diagnostic dimension. When it is not, labels are as useful as biased features.

Figure 2 shows the learning curves with error data averaged across sets of four trials, and across all participants in each condition (the black dots). Exponential functions of the form $Y = a \exp(-bX)$ (see Heathcote, Brown & Mewhort, 2000) were fit to the data of each condition. Model parameters were fit using maximum likelihood estimation, and Bayesian Information Criterion (BIC; Schwarz, 1978) was used for model selection. BIC is a measure that considers the data fit, but penalizes models for having excessive parameters (Myung & Pitt, 1997). The model that minimizes BIC should be preferred.

Two (of several[1]) simple models that were fit to the mean error data are also shown in Figure 2. Both models used the exponential function, but one model allowed unique parameter values for each condition, while the other model used the same two parameter values for all conditions. These two models were compared, using BIC to estimate the Bayes Factor (see Myung & Pitt, 1997), which provides the odds in favor of the model with the lower BIC score. The model that allowed unique parameter values was preferred according to this criterion (BIC for unique parameters model = 186.0 vs. BIC for same parameters model = 1499.1; according to the Bayes Factor approximation, a difference between BIC scores of such a large magnitude translates to extremely strong evidence in favor of the full model). This suggests that each condition

[1] Other models that were compared with the "unique parameter values" model to test for interaction effects were also found to be inferior according to BIC. (For instance, a model that allowed the PC, NON-BIASED FEATURE condition to have different parameter values to the other five conditions.)

had different learning curves – that is, that participants did not behave identically in each condition.

As Figure 2 demonstrates, the main effect was that learning was faster overall when the task occurred in a social context than when it was presented on a computer. Presenting the category learning task in a social context led to improved learning performance overall[2]. It seems that participants were much more engaged, and thus solved the task quite quickly across all three of the SOCIAL conditions.

There was a more pronounced difference between the LABEL, BIASED FEATURE and NON-BIASED FEATURE conditions when the stimuli were presented on the PC than when they were presented in a social context. That said, in both the SOCIAL and PC conditions, learning was slowest in the NON-BIASED FEATURE conditions, and fast in the BIASED FEATURE conditions, as expected. Of primary interest is to compare performance in the LABEL conditions with that of the other conditions. For both the SOCIAL and PC conditions, learning speed in the LABEL condition was closer to that of the BIASED FEATURE condition than to that of the NON-BIASED FEATURE condition. However, while on the PC, learning in the LABEL condition was slower than learning in the BIASED FEATURE condition. In contrast, in the SOCIAL conditions, the learning curves of LABEL and BAISED FEATURE conditions were very similar. This suggests a weak effect that labels were more privileged in the social context than on the computer. Nonetheless, in either context, diagnostic labels did not help category learning beyond help that could be given by a diagnostic biased feature.

Why was learning not *fastest* in the LABEL conditions? Let us consider the difference in the learning task between the LABEL and BIASED FEATURE conditions. To successfully complete the category learning task, participants in all conditions needed to: 1) notice that one particular feature dimension was diagnostic (e.g., the labels); and 2) learn the match between each particular diagnostic feature value and a target feature value (e.g., that "Bragens" wanted the dagger, and "Lathors" wanted the hammer). However, the BIASED FEATURE condition was easier: the diagnostic feature dimension (clothing) was not only salient and meaningful (and thus easy to notice), but each outfit also meaningfully corresponded to an item (e.g., the robe outfit matched with the staff). Thus, participants could essentially come to the task already knowing the correct answers. The LABEL condition was actually a more difficult task, because although the diagnostic feature dimension was perceptually salient, people needed to learn an arbitrary match between the novel names and the target features. It is interesting that despite the added difficulty, learning in the LABEL

[2] Exponential functions for LABEL, BIASED FEATURE and NON-BIASED FEATURE conditions were $Y = 3.75\exp(-0.10X)$, $Y = 3.37\exp(-0.17X)$, and $Y = 2.76\exp(-0.04X)$ for the PC conditions, and $Y = 2.53\exp(-0.20X)$, $Y = 2.90\exp(-0.24X)$, and $Y = 4.14\exp(-0.19X)$, for the SOCIAL conditions, respectively. Note that the two parameters vary between conditions; larger a indicates more errors at the beginning of training and larger b indicates faster learning.

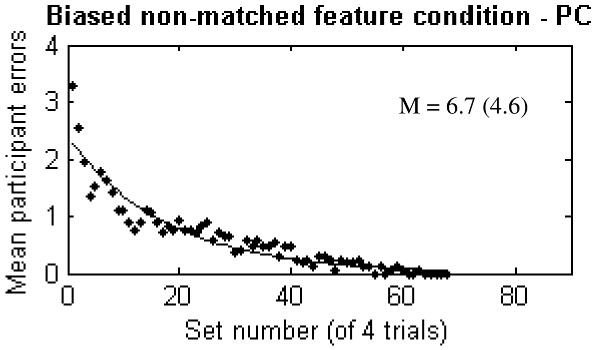

Figure 3: Learning curve averaged across sets of 4 trials, and averaged across participants in this condition. The exponential function is fit to the error data.

conditions was similar to that of the BIASED FEATURE conditions. However, the added difficulty might explain why learning in the LABEL conditions was not fastest.

To determine whether this explains the observed similarity between the LABEL and BIASED FEATURE conditions, we ran another experimental condition on the PC with 17 new participants[3]. The category learning task was identical[4] to that of the PC, BIASED FEATURE condition, with clothing as the diagnostic feature dimension; however, the feature values no longer matched with the expected target feature values. Counterintuitively, the "lumberjack" wanted the dagger, the "warrior" wanted the hammer, the "tradesperson" wanted the staff, and the "wizard/priest" wanted the timber axe. This new condition still had the salient and expectedly meaningful clothes features as the diagnostic dimension, but the match between feature values was arbitrary. As Figure 3 shows, learning was slowed nearly to the same level as in the PC, NON-BIASED FEATURE condition. This suggests that the fast learning in the BIASED FEATURE condition was due to the pre-existing knowledge of the mapping between outfits and target items. Learning could be fast in the LABEL condition because the names were completely novel and "blank"; unlike the clothes in the new condition, the LABEL condition did not require any unlearning of associations between the diagnostic dimension and the target feature dimension.

Discussion

We set out to explore whether social context has an impact on the status of category labels for adults. We investigated whether people would pay special attention to labels presented in a social context, and hence learn quickly when these labels *are* most informative. While there was some suggestion that labels were more privileged in a social context than they were on a computer, the main result was

[3] Participants were undergraduates at the University of Adelaide, or recruited from the general community (9 males). Ages ranged from 17 to 38 years. Participants received AU$5.

[4] Feedback was slightly different in this condition, due to the availability of images: participants did not see an image of the correct target item being held by the alien creature.

that people can solve a category learning task much faster when they are in an engaging, social context, regardless of whether a label or another feature is actually more informative. We suspect that the participants were more motivated to do well in the presence of a human teacher and with a more enjoyable, interactive task. When the task is not social, novel labels are privileged over non-salient and arbitrarily-matched features, but are no more useful than biased features. Nonetheless, it is interesting that learning in the LABEL conditions was fast, despite participants having to learn an arbitrary match between each novel label and the target feature. Presumably, if the labels were *not* novel and were appropriately matched (e.g., "Wizard" and "Warrior"), the task would become trivially easy and participants would learn even faster.

The results of this study support an intermediate view between labels being "just another feature" and having a unique, privileged status. Novel labels can aid category learning better than arbitrary (or biased but arbitrarily matched) features can: labels are salient, and novel labels permit new associations with features to be learned without being hindered by knowledge about existing feature associations. However, context matters: if people are already fully engaged with the category learning task, there is less scope for labels to aid learning beyond other features. One caveat to this finding is that perhaps the influence of labels in the social context was somewhat hidden by a ceiling effect, since our rule-based category structure was a simple one, quickly solved by most participants. More challenging category structures may reveal a larger influence of labels within a social context.

Further work is required to determine whether adults *process* labels differently to other features, or simply weight them more heavily. Gliozzi, Mayor, Hu and Plunkett (2009) contrast an *unsupervised feature-based account* and a *supervised name-based account* of category formation. According to the latter account, objects given the same name belong to the same category, and so labels act as invitations to form categories and highlight commonalities between objects. The unsupervised feature-based account says that labels have the same status as other features. Labels may vary in salience, just like other features, but are handled with the same statistical inference processes as are other features. The model by Gliozzi et al. (2009) suggests that for infants, labels play a mundane but powerful role as simply additional features. Our study, and the experiments by Yamauchi and Markman (e.g., 1998; 2000a) are consistent with this view.

Finally, this experiment has implications for adult category learning studies that are not presented in an engaging, social context. We found that presenting a category learning task in the absence of a social context does not encourage optimal learning behavior. Category learning on a computer may not reflect category learning in the more engaging situations typically encountered in real life, so it is worth understanding category learning in more naturalistic, social contexts.

Acknowledgments

RGS was supported by an Australian Postgraduate Award and DJN was supported by an Australian Research Fellowship (ARC grant DP-0773794). Thank you to the reviewers for their helpful suggestions.

References

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology*, *64*, 3–26.

Campbell, A. L. & Namy, L. L. (2003). The social-referential context in verbal and nonverbal symbol learning. *Child Development*, *74*, 549–563.

Fulkerson, A. L. & Haaf, R. A. (2003). The influence of labels, non-labeling sounds, and source of auditory input on 9- and 15-month-olds' object categorization. *Infancy*, *4*, 349–369.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, *23*, 183–209.

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development*, *58*, 1532–1541.

Gliozzi, V., Mayor, J., Hu, J-F., & Plunkett, K. (2009). Labels as features (not names) for infant categorization: A neurocomputational approach. *Cognitive Science*, *33*, 709–738.

Goldstone, R. L., Lippa, Y., & Shiffrin, R. M. (2001). Altering object representations through category learning. *Cognition, 78*, 27–43.

Grice, H. P. (1975). Logic and conversation. In Cole, P. & Morgan, J. (Eds.) *Syntax and Semantics*, *3*. New York: Academic Press.

Heathcote, A. J., Brown, S. D., & Mewhort, D. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin and Review*, *7*, 185–207.

Johansen, M. K., & Kruschke, J. K. (2005). Category representation for classification and feature inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1433–1458.

Myung, I. J., & Pitt, M. A. (1997). Applying Occam's razor in modeling cognition: A Bayesian approach. *Psychonomic Bulletin and Review, 4*, 79–95.

Plunkett, K., Hu, J-F., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. *Cognition*, *106*, 665–681.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461-464.

Waxman, S. R., & Braun, I. (2005). Consistent (but not variable) names as invitations to form object categories: New evidence from 12-month-old infants. *Cognition*, *95*, B59–B68.

Yamauchi, T., Love, B. C., & Markman, A. B. (2002). Learning nonlinearly separable categories by inference and classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 585–593.

Yamauchi, T., & Markman, A. B. (1998). Category learning by inference and classification. *Journal of Memory and Language*, *39*, 124–148.

Yamauchi, T., & Markman, A. B. (2000a). Learning categories composed of varying instances: The effect of classification, inference and structural alignment. *Memory & Cognition*, *28*, 64–78.

Yamauchi, T., & Markman, A. B. (2000b). Inference using categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 776–795.