Developmental Science 9:6 (2006), pp 597–603

REPORT

Children use categories to maximize accuracy in estimation

Sean Duffy,[1] Janellen Huttenlocher[2] and L. Elizabeth Crawford[3]

1. *Department of Psychology, Rutgers University, USA*
2. *Department of Psychology, The University of Chicago, USA*
3. *Department of Psychology, The University of Richmond, USA*

Abstract

The present study tests a model of category effects upon stimulus estimation in children. Prior work with adults suggests that people inductively generalize distributional information about a category of stimuli and use this information to adjust their estimates of individual stimuli in a way that maximizes average accuracy in estimation (see Huttenlocher, Hedges & Vevea, 2000). However, little is known about the developmental origin of this cognitive process. In the present study, 5- and 7-year-old children viewed stimuli that varied in size and reproduced each from memory. Consistent with the predictions of a Bayesian model of category effects on estimation, responses were adjusted toward the central value of the stimulus distribution. Additionally, the dispersion of the stimulus distribution affected the pattern of bias and variability of responses in a way that is predicted by the model. The results suggest that, like adults, children use categories for increasing average accuracy in estimating inexact stimuli.

Introduction

The present study tests a model of category effects on stimulus judgment in children. It is well known that the ability to form categories from experience with objects is essential for adaptive functioning in a variety of cognitive domains. A well-known finding in the adult memory literature is the *schema effect* in which a stimulus is remembered as being more typical of the category of which it is a member than it actually was (Bartlett, 1932; Brewer & Nakamura, 1984; Poulton, 1989). While this bias has been described as memory distortion, Huttenlocher, Hedges and Vevea (2000) have proposed that this bias results from an adaptive process that improves accuracy in estimating inexactly remembered stimuli.

Huttenlocher *et al.* propose a category adjustment model in which people code stimuli hierarchically, at fine-grain and category levels, and combine information from both levels when reconstructing a stimulus from memory. This combination of information across levels moves estimates of stimuli towards the central value of the category and increases average accuracy by decreasing the random error surrounding the true value for a stimulus. The model and its predictions are described in greater detail below.

A variety of studies show that the ability to induce statistical properties of categories arises early in development (Strauss, 1979; Bomba & Siqueland, 1983; Eimas & Quinn, 1994; Saffran, Aslin & Newport, 1996; Kirkham, Slemmer & Johnson, 2002). However, whether children use categories for improving the accuracy of their estimates of stimuli from memory is unknown. There appear to be two potential explanations for the origin of such category effects. One possibility is that the use of categories in estimation arises gradually through experiences in which adjusting stimuli towards the center of a category increases accuracy. Children may then begin relying upon this strategy automatically whenever estimating stimuli. If so, category effects should be relatively small or non-existent in younger children and emerge slowly over the course of childhood. Alternatively, category effects might arise early in development from innately available, hardwired cognitive mechanisms that are available from birth. They might have arisen over the course of evolution in a variety of species because high accuracy is adaptively important in a broad range of ecological situations. If so, the use of categories should be seen early in development. To evaluate different possible explanations for the origins of category effects, the first step is to determine if they are found in young children at all, and if so, whether they are similar to the

Address for correspondence: Sean Duffy, Department of Psychology, Rutgers University, Camden, NJ, 311 North Fifth Street, Camden NJ, 08102, USA; e-mail: seduffy@camden.rutgers.edu

effects observed in adults. The present study explores how categories influence estimation in 5- and 7-year-old children by exploring memory biases that arise in a sequential reproduction task that are predicted by Huttenlocher *et al.*'s (2000) category adjustment model.

The category adjustment model

According to the category adjustment model (Huttenlocher *et al.*, 2000), an estimate of a stimulus is reconstructed from information about the stimulus and the category of which it is a member. The model defines a category as a cognitive structure consisting of a bounded range of stimulus values along a set of relevant dimensions. The model posits that summary information such as the central value, dispersion of members, and boundaries is available at the time of estimation. A stimulus is defined as a value along a relevant stimulus dimension such as a line of a particular length. A stimulus is considered to be a member of a category if it falls within the bounded range for that category; however, there may be some uncertainty about category membership if the stimulus is near a boundary region of the category.

The category adjustment model (Huttenlocher, Hedges & Duncan, 1991; Huttenlocher *et al.*, 2000) posits that estimation relies on a process that approximates a Bayesian inference. Bayes theorem provides a method for combining a prior distribution with a sampling distribution to produce estimates that minimize average error. The category adjustment model proposes that when a stimulus is encountered, it is coded at two levels of detail: a categorical level and a particular level. Category-level knowledge operates as a prior distribution and the particular-level knowledge as a sampling distribution, and these sources of information are weighted and combined during stimulus estimation. Combining category-level and particular-level information in this way is a rational strategy because while it introduces bias in estimates, it also reduces their variability, producing a lower mean squared error of responses than if the category information had not been used. The extent of adjustment towards the central value depends upon three factors: the inexactness of the fine-grain memory, the dispersion of the category, and the degree of certainty about the category membership of stimuli. We examine these in greater detail below.

Effects of the exactness of fine-grain memory

Categories improve accuracy only when memory for a stimulus is inexact. If stimuli are remembered precisely, there is no inexactness in memory and responses will always equal the true stimulus values. For instance, if you just measured your height with a ruler, your estimate of your height will equal your actual height, and the use of category information is unnecessary for increasing accuracy. Alternatively, if nothing is known or remembered about the particular stimulus except for its category membership, the central value for the category will always be given, and there will be no variability in responses but bias will be maximal. For instance, if you have never met the authors' friend Kim but know she is a woman, your best estimate for her height is the average height for women. For intermediate levels of inexactness, stimuli will be adjusted toward the category center, with the extent of the adjustment varying as a function of the exactness of the fine-grain memory.

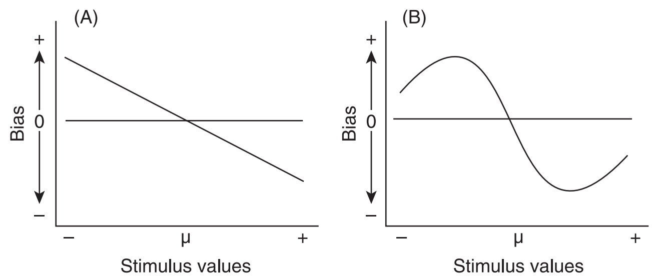

Figure 1 *Hypothetic bias curves for categories exhibiting a uniform distribution (A) and a peaked distribution (B) as predicted by the category adjustment model.*

Bias is defined as the difference between a participant's response and the true stimulus value. Positive bias indicates overestimation; negative bias indicates underestimation. The schematic graph in Figure 1A depicts bias on the vertical axis and actual stimulus values on the horizontal axis. If category information is not used in estimation, responses are not adjusted towards the central value and the bias curve has no slope. In such conditions, responses will on average equal the actual stimulus value across the entire stimulus space. Alternatively, if responses are adjusted towards the category center, smaller values are overestimated and larger values are underestimated and the bias curve has a negative slope. With greater inexactness in the fine-grain memories, stimuli will be adjusted to a greater extent towards the category center, resulting in a steeper negative slope for the bias curve.

Effects of category dispersion on bias strength and variability of responses

The concentration of instances within a category affects the variability of responses. Within a category, stimuli might be distributed uniformly, with instances spread evenly across the entire category, or peaked, with instances

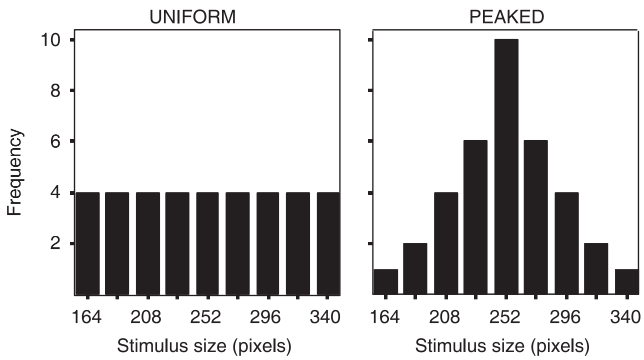

Figure 2 *Frequency distributions used in the present study.*

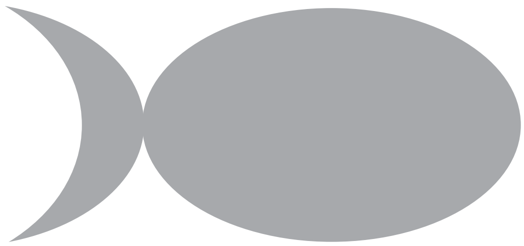

Figure 3 *Schematic fish used in the present study.*

concentrated near the center (see Figure 2). When the dispersion of values is nearly infinite, the central value provides very little information about any given category member. In this situation, category-level information is used to a minimal degree for adjustment and thus the variability of responses to any particular stimulus will be fairly high. At the other extreme, if the category has only one value, all responses will be of that value and there will be no variability in responses. Between these two extremes, the variability of estimates for a particular stimulus should vary as monotonic function of the dispersion of instances in the category. Thus, estimates for a stimulus value will be less variable when that value is embedded in a distribution of lower variability, such as a peaked distribution, than when it is embedded in a distribution of higher variability, such as a uniform distribution.

Effect of category dispersion on the shape of the bias curve

The dispersion of instances will also influence the shape of the bias curve due to the degree of certainty about category membership for extreme category values. A stimulus will be adjusted only if it is judged to be a category member. There is uncertainty about category membership for values near boundaries. That uncertainty is greater for a peaked distribution in which most values are concentrated near the center of the category and only a few values are near the boundaries, than for a uniform distribution, in which stimuli are evenly distributed across the entire stimulus space and there are many stimuli near the boundaries. When the category membership of a stimulus is uncertain, the extent of the adjustment should be proportional to the degree of uncertainty about its membership, so that bias should be weaker near the category boundaries. Thus the shape of the bias pattern across a category should differ for uniform versus peaked distributions: the curve for a uniform distribution should be quite linear whereas the curve for a peaked distribution should be curvilinear (schematically depicted in Figures 1A and 1B).

Category adjustment in children

This study examines whether young children's stimulus estimates exhibit a pattern of bias that is predicted by the category adjustment model. We present 5- and 7-year-old children with a set of schematic fish stimuli on a computer, as in Figure 3. The fish vary in fatness, with nine distinct stimulus values. Half the children estimate stimuli from a uniformly distributed category: there are an equal number of fish from each of the nine stimulus values. The other half estimate stimuli from a peaked category having lower dispersion; the majority of fish are from the central five stimulus values. Figure 2 presents these two distributions. In the task, children view a target fish for 2 seconds, it disappears for 2 seconds, then a response fish appears that they must adjust to be equal in size to the target fish. Both distributions consist of 36 fish that children sequentially reproduce in random order.

If children use information in categories for adjusting responses, their estimates of stimuli from those categories should exhibit a central tendency bias. In addition, if younger children's memory for individual stimuli is less exact than older children's, the model predicts that bias should be stronger for younger children than for older children. Additionally, if children are sensitive to the dispersion of instances across the category and use this information to optimize accuracy, the slope of the bias curve for the central values of the category should be steeper in the peaked than in the uniform condition. This is because for those items that are clearly within the category, the central value in the peaked condition is more informative than in the uniform condition and thus is given more weight in estimation. This prediction pertains to only central values of the category because, as noted above, the category membership of stimuli near the edges is more uncertain. Thus, the shape of the bias curve should be less linear for the peaked than for the uniform distribution because stimuli near the edges will not be adjusted as much in the peaked condition. Finally, the standard deviation of responses to individual stimuli should be less for the peaked than for the uniform distribution due to the fact that responses are adjusted less in the uniform condition.

Method

Participants

The participants were 60 kindergarten (25 boys and 35 girls) and 40 second-grade students (22 boys and 18 girls) at the University of Chicago Laboratory School. The mean age of the 5-year-olds was 5;2 (years;months) (range: 4;8–5;6), and the 7-year-olds' mean age was 7;3 (range: 6;10–7;8).

Materials

The presentation of stimuli and the collection of responses were controlled by a Macintosh PowerBook laptop with a 15.2-inch screen. The stimuli were blue fish consisting of an elliptical body and a fan-shaped tail. They varied only in fatness (the vertical dimension); their length was always 400 pixels (1 pixel = $\frac{1}{30}$th cm). Each stimulus fish was presented near the center of the monitor.

The children were randomly assigned to the uniform condition or the peaked condition. In both there were nine distinct stimulus values, varying in fatness from 164 pixels to 340 pixels in 22-pixel increments. In the uniform condition, each stimulus fish was presented four times in random order. In the peaked distribution, fish in the middle category were presented more often than the fish at the tail ends of the distribution. The frequencies of these distributions are presented in Figure 2.

Procedure

The procedure was a memory test in which a fish appeared for 2 seconds and disappeared. After a delay of 2 seconds, an 80-pixel thin response fish appeared. The child used the keyboard to adjust the response fish to match the size of the target fish they had just observed. The keyboard had stickers on the 'S' and 'K' keys depicting a thin or a fat fish, respectively. Pressing the fat fish key made the response fish expand in fatness while pressing the thin fish key made the fish contract and become thinner.

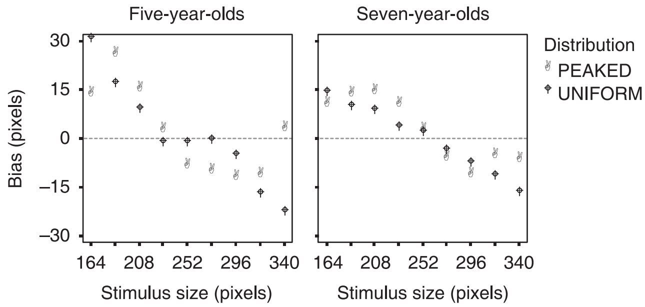

Figure 4 *Bias curves (in pixels) for 5- and 7-year-olds.*

Children were told that they were going to play a game where they would be making 'fish friends' and that it is important to make the second fish be the same size as the first fish on each trial. They were given two practice trials where they were taught how to use the keyboard to adjust the response fish. The fish stimuli used during practice were randomly sampled from within the category for each child. After the child adjusted the response fish and indicated satisfaction with its size, the experimenter pressed the return key to move to the next trial. Following the two practice trials the actual experimental session of 36 trials began. The experiment lasted about 12 minutes.

Results

For each response, bias was computed as the difference between the participant's estimate and the actual stimulus size. Responses greater than 3 standard deviations from the mean for a particular stimulus value were dropped, eliminating less than 1% of the data. These points represented trials in which the participant may have been distracted and failed to attend to the target presentation.

Do children use categories in estimation?

Figure 4 presents the bias curves for both age groups and conditions. If children use categories in estimation, the bias of responses by actual stimulus magnitude should have a significant negative slope. To test this, linear regressions were performed for each of the four groups on average bias for each stimulus value against actual stimulus size. For both 5- and 7-year-olds in both the uniform and peaked condition, the slopes were negative and their coefficients were significant, indicating that children adjusted their estimates towards the central values of the set of instances. The regression estimates are presented in Table 1.

Table 1 *Regression coefficients for all nine stimuli for 5- and 7-year-olds*

	Intercept				Slope			
Condition	Beta	SE	*t*	*p*	Beta	SE	*t*	*p*
5 Peaked	51.68	7.69	6.72	.001	−0.221	0.03	−7.32	.001
5 Uniform	64.77	6.66	9.71	.001	−0.264	0.026	−10.18	.001
7 Peaked	47.30	4.428	10.68	.001	−0.182	0.017	−10.46	.001
7 Uniform	40.83	4.21	9.69	.001	−0.168	0.016	−10.24	.001

We next tested whether the slopes of the bias curves differed between 5 and 7 years of age. If younger children's memory for particular stimuli is less exact than older children, they should introduce greater bias into their estimates. To test this prediction, we collapsed the data by condition and performed a multiple regression analysis comparing bias slope by using dummy variables. This analysis showed that the 5-year-olds had significantly steeper slopes than the 7-year-olds ($t(17) = 3.508$, $p < .001$), indicating that the 5-year-olds introduced greater bias into their estimates than the 7-year-olds.

An additional question concerns whether children adjust their responses towards the central value of the category or simply towards the value of the immediately preceding stimulus. With aggregated data, the two alternatives lead to the same pattern of bias towards the central region of the category, although by very different mechanisms. To test which of these possibilities best explain the observed results, we first eliminated the participant's first response because there is no prior stimulus and analyzed the data for each subsequent response by the order within the series that the stimulus appeared. We performed multiple linear regressions on the bias of each response with two predictors: the running average magnitude of all stimuli observed, and the magnitude of the immediately preceding stimulus. The analysis showed that for each condition, the magnitude of the prior stimulus did not affect bias; however, the running average of all stimuli observed was a significant predictor in each condition. This result suggests that children are adjusting responses towards the central region of the category rather than towards the immediately preceding stimulus.

Does the shape of the bias curve vary by stimulus distribution and age?

The model predicts that the shape of the bias curve should differ depending on whether the distribution is uniform or peaked. The bias curve in the peaked distribution should depart from linearity to a greater extent than the bias curve for the uniform distribution. To test this, the mean response for each of the nine stimuli for each child was entered as a within-subjects factor in a repeated measures ANOVA. The polynomial contrasts for the linear and non-linear components of the within-subjects factor were used to calculate η^2 – the linear component divided by the sum of the linear and non-linear components. When $\eta^2 = 1$ the bias curve is a straight line, and lower values indicate increasing departure from linearity. To test whether values of η^2 differed significantly, we conducted a nested *F* test following the procedure outlined in Huttenlocher *et al.* (2000).

The result of this analysis indicated that for both 5- and 7-year-olds, the slope of the bias curve was more linear for the uniform than for the peaked distribution. For the 5-year-old uniform condition, $\eta^2 = .990$, while for the peaked condition, $\eta^2 = .810$ ($F(2, 310) = 14.25$, $p < .001$). For the 7-year-old uniform condition, $\eta^2 = .999$, in the peaked condition, $\eta^2 = .951$ ($F(2, 188) = 4.29$, $p < .01$).

Another model prediction is that the bias for the central category members should be stronger for the peaked than for the uniform distribution. To test this we performed a regression on only the central five stimulus values using dummy variables to compare whether the slope of the bias curve differs significantly between the uniform and peaked condition. For 5-year-olds, the coefficient of the dummy variable for the slope was marginal, $t(9) = 2.192$, $p = .071$, but for 7-year-olds the coefficient is significant, $t(9) = 4.637$, $p < .005$. The result of this analysis and that of the linearity of the bias curve provide additional evidence that children are sensitive to the dispersion of instances across the category.

Does response variability vary by distribution and age?

Finally, the model predicts that responses in the peaked condition will exhibit lower variability than in the uniform condition. To test this, standard deviations were calculated for each stimulus value in both conditions for both age groups. An ANOVA was conducted on log-transformed standard deviations for each stimulus value with age and distribution as between-participant factors. Note that stimuli near the edges of a category are treated differently than more central values. As noted in the introduction, category membership of stimuli near the boundaries is more uncertain in peaked than in uniform distributions. The differences we predict in variability between distributions should be clearest among stimuli that are clearly treated as category members. Therefore we excluded responses to the two smallest and largest stimuli and ran the analysis using only the central five stimuli.

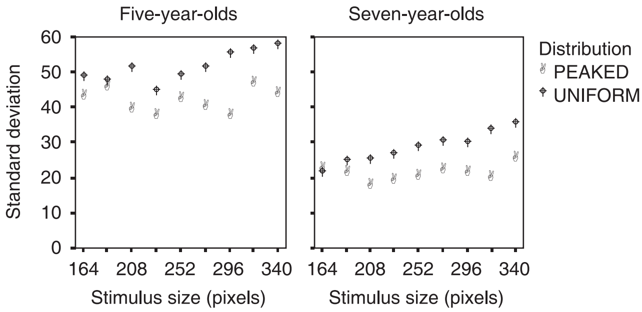

Figure 5 *Standard deviations (in pixels) for 5- and 7-year-olds.*

This analysis yielded a significant effect of distribution, $F(1, 16) = 65.54$, MSE = 0.078, $p < .001$, effect size = .80, and age, $F(1, 16) = 331.9$, MSE = .397, $p < .001$, effect size = .95, consistent with the prediction that children are sensitive to and use frequency information in estimation. There was no interaction between these factors, $F(1, 16) = 1.86$, *ns*. For the 5-year-old group, the mean standard deviation of responses in the peaked condition was 38.57 pixels and in the uniform condition 49.10 pixels. For the 7-year-olds, average standard deviation in the peaked condition was 19.25 pixels and in the uniform condition 26.94 pixels. The standard deviations are shown by stimulus value in Figure 5.

Discussion

This study examined whether young children form an inductive category from a set of stimuli varying along a continuous dimension and use information about the distribution of stimuli for reconstructing estimates of individual stimuli. We found that both 5- and 7-year-old children exhibited a pattern of responses indicating that they formed and used categories when estimating stimuli. Furthermore, their responses were consistent with the predictions of the category adjustment model (Huttenlocher *et al.*, 2000).

Children biased estimates of stimuli towards the central value of the set of stimuli they reproduced. The model posits that adjustment towards the central region of categories arises when people combine inexact information about stimuli with category-level information about the prior set of stimuli they reproduced. The category adjustment model posits that such a combination approximates a Bayesian procedure of adjusting a present distribution of inexactness surrounding the true value for a stimulus towards a prior distribution of the category of stimuli, introducing bias into responses yet decreasing their overall variability.

Furthermore, the pattern of bias and response variability suggests that the frequency distribution of the stimuli influenced children's estimates in a manner predicted by the model. First, for both 5- and 7-year-olds, the bias curve in the uniform condition exhibited greater linearity than in the peaked condition, suggesting uncertainty about category membership for extreme values in the peaked condition. Second, responses in the central region of the stimulus distribution exhibited stronger bias in the peaked condition relative to the uniform condition, suggesting that the central value in the peaked condition exhibited greater weight in estimation because it provided more information than the central value in the uniform condition. Third, for both age groups, the variability of estimates for individual stimulus values was less for children in the peaked than in the uniform distribution condition due to the greater adjustment in this condition. Each of these findings follows the predictions of the category adjustment model, suggesting that children, like adults, use inductive categories to adjust estimates of inexactly encoded stimuli, and use this information to maximize accuracy.

The finding that children are sensitive to differences in frequency distributions is consistent with the perspective that frequency is encoded automatically. For example, studies using word lists by Hasher and Zacks (1979) show that people remember the number of times they see particular words, even if they do not do so intentionally (see also Zajonc, 1968). Using stimuli varying on a continuous dimension, we also find that children automatically encode frequency information about a set of stimuli.

Overall, 5-year-olds exhibited greater variability in their responses than did 7-year-olds, suggesting greater memory inexactness for each stimuli. A variety of age-related cognitive changes may explain this difference, such as the development of attention and working memory capacities. However, 5-year-olds also biased their responses to a greater extent than 7-year-olds. Recall that the model predicts that bias increases as a function of decreasing exactness in the fine-grain memories. The pattern of bias we have observed across the two age groups suggests that 5-year-olds are adjusting their responses to a greater extent towards the prototype in order to compensate for the inexactness in their fine-grain memories for stimuli. Seven-year-olds introduce less bias into their responses because their memory for individual stimuli is more exact.

At first glance, the finding that younger children rely on categories in estimation appears to contradict the findings of a recent study by Sloutsky and Fisher (2004) demonstrating that young children perform similarity- rather than category-based inductive reasoning. However, note that the design and purpose of the Sloutsky and

Fisher study and the present investigation differ in a number of ways. Sloutsky and Fisher (2004) explored recognition memory for individual members of natural categories containing salient individuals (individual animals with distinctive features). The present study explored a more basic process: *estimation* of stimuli from within a single category of stimuli varying along a continuous dimension. For such categories, similarity, rather than an abstract conceptual rule, defines the category, and thus our results are consistent with those reported by Sloutsky and Fisher. Moreover, it is plausible that the use of category information in estimation and induction differs, and more work is needed to explore the relationship between these two kinds of tasks.

That children as young as 5 adjust use categories in estimation much like adults suggests that these processes must be available early in development. To determine exactly when these procedures arise, it would be useful to explore category effects with younger children and even infants. Unfortunately, methodological challenges in testing younger participants in serial reproduction tasks limit our ability to determine whether younger children or even infants use categories in reconstructive memory. However, if category adjustment procedures are available from the start of life, an important question is how elaborate these initial mechanisms might be. For example, it may be the case that category adjustment procedures begin as a more simple process of moving estimates towards the central region of a category, and over time and experience, sensitivity to the dispersion of category members become incorporated into memory processes.

While the findings reported here are consistent with the predictions of the category adjustment model, one direction for future research may be to compare the predictions of the category adjustment model against other categorization models. For instance, Nosofsky's Generalized Context Model (Nosofsky, 1986) proposes that category effects in judgment arise at the point of encoding rather than reproducing stimuli. To rule out alternative models, a future direction may be to study category effects that arise in estimating stimuli from two adjacent categories, or manipulate the order of stimuli to determine whether the category adjustment model or alternative models better explain the observed phenomena. However, the fact that young children exhibit category effects in stimulus estimation provides a first step towards better understanding the role of categories in the development of memory.

Acknowledgement

The research reported here was supported in part by a grant from the National Science Foundation REC0337360.

References

Bartlett, F.C. (1932). *Remembering: A study in experimental psychology.* Cambridge: Cambridge University Press.

Bomba, P.C., & Siqueland, E.R. (1983). The nature and structure of infant form categories. *Journal of Experimental Child Psychology*, **35**, 609–636.

Brewer, W.F., & Nakamura, G.V. (1984). The nature and function of schemas. In R.S. Wyer & T.S. Srull (Eds.), *Handbook of social cognition* (Vol. 1, pp. 119–160). Hillsdale, NJ: Erlbaum.

Eimas, P.D., & Quinn, P.C. (1994). Studies on the formation of perceptually based basic-level categories in young infants. *Child Development*, **65**, 903–917.

Hasher, L., & Zacks, R.T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, **108**, 356–388.

Huttenlocher, J., Hedges, L., & Duncan, S. (1991). Categories and particulars: prototype effects in estimating spatial location. *Psychological Review*, **98** (3), 352–376.

Huttenlocher, J., Hedges, L., & Vevea, J. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology: General*, **129** (2), 220–241.

Kirkham, N.Z., Slemmer, J.A., & Johnson, S.P. (2002). Visual statistical learning in infancy: evidence for a domain-general learning mechanism. *Cognition*, **83**, B35–B42.

Nosofsky, R.M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, **115**, 39–57.

Poulton, E.C. (1989). *Bias in quantifying judgements.* Hillsdale, NJ: Erlbaum.

Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning by 8-month old infants. *Science*, **274**, 1926–1928.

Sloutsky, V.M., & Fisher, A.V. (2004). When learning and development decrease memory: evidence against category-based induction. *Psychological Science*, **15**, 553–556.

Strauss, M.S. (1979). Abstraction of prototypical information in adults and 10-month-old infants. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **5**, 618–632.

Zajonc, R.B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology Monograph Supplement*, **9**, 1–27.

Received: 1 September 2004
Accepted: 23 February 2006