Perception, 2012, volume 41, pages 556–568

doi:10.1068/p7096

Experience produces the atypicality bias in object perception

Justin Kantner§, James W Tanaka
Department of Psychology, University of Victoria, PO Box 1700, Victoria, BC V8W 3P5, Canada;
e-mails: kantner@psych.ucsb.edu; jtanaka@uvic.ca
Received 9 August 2011, in revised form 26 March 2012

Abstract. When a morph face is produced with equal physical contributions from a typical parent face and an atypical parent face, the morph is judged to be more similar to the atypical parent. This discontinuity between physical and perceptual distance relationships, called the "atypicality bias" (Tanaka et al 1998, *Cognition* **68** 199–220), has also been demonstrated with non-face objects (birds and cars; Tanaka and Corneille 2007 *Perception & Psychophysics* **69** 619–627). We tested whether the atypicality bias can be induced for a novel set of artificial objects. Two categories of "blob" stimuli were generated, each composed of typical and atypical members. Morphs averaged from typical and atypical parent exemplars were used to test the presence of an atypicality bias before and after participants were familiarized with blob items. In experiment 1, participants were trained to discriminate between the two blob categories. An atypicality bias was evident after, but not prior to, category training. In experiment 2, participants rated the pleasantness of the blobs instead of learning to categorize them; an atypicality bias was present only after the ratings task. This finding suggests that relatively passive exposure to exemplars is sufficient to influence perceptions of similarity, and that the atypicality bias is a manifestation of this influence.

Keywords: object perception, perceptual learning, similarity, atypicality, bias

1 Introduction

A useful method for representing the perceptual similarity relationships among members of an object class is to conceive of them as occupying points in a multidimensional space (Nosofsky 1986; Shepard 1962). Just as the proximity of cities on a map can be determined by comparing the values of each city along the dimensions that define their locations (latitude and longitude), the similarity of exemplars in psychological space is a function of their values along the *n* dimensions of the space, where *n* is the number of features used in evaluating their similarity. Among its numerous applications, this framework provides a means of operationalizing the concepts of stimulus typicality and atypicality (or distinctiveness): exemplars are typical to the extent that they possess values on each dimension that are close to the values of most of the other exemplars in the space, and atypical to the extent that they possess extreme values on one or more dimensions. Typical exemplars occupy densely populated regions of psychological space, a reflection of their similarity to many other exemplars, while atypical exemplars lie in sparsely populated regions of the space. A prominent theoretical application of similarity space is Valentine's (1991) "face–space" model, which accounts for typicality effects in face recognition and categorization based on the locations of face representations in a multidimensional space (see, eg, Lee et al 2000; Leopold et al 2001; Valentine 2001). The multidimensional space framework has also been used to describe object similarities in non-face domains (eg Cutzu and Edelman 1996) and used in theories of speech perception (Iverson and Kuhl 1995, 2000).

§ Current address: Department of Psychological and Brain Sciences, University of California, Santa Barbara, Santa Barbara, CA 93106, USA.

The distribution of items in dense and sparse regions of similarity space is important because it may affect how the items themselves are perceived (eg Krumhansl 1978). Tanaka et al (1998) investigated the relationship between typicality and object perception by creating morphs comprised of equal physical contributions from a typical parent and an atypical parent. Based upon a purely structural metric of similarity, these morphs were equidistant to their typical and atypical parents. However, participants tended to perceive the morph as bearing a stronger resemblance to the atypical parent. Tanaka et al (1998) demonstrated this "atypicality bias" using a preference task in which the two parent faces were presented either immediately before (experiments 1, 2, and 4) or alongside (experiment 3) their morph; participants were asked to choose to which parent the morph was more similar. Across experiments, the atypical parent was chosen on 56% to 63% of trials ($M = 60\%$). In a follow-up study, Tanaka and Corneille (2007, experiment 3) extended the atypicality bias to bird and car stimuli, demonstrating that the bias is not a face-specific effect. Tanaka et al (2011) found that children as young as 3–4 years show an atypicality bias with faces and birds, indicating that knowledge of category structures can develop with relatively little perceptual experience.

Tanaka and colleagues accounted for their results using an attractor field model of face and object perception (Tanaka and Corneille 2007; Tanaka et al 1998; see also Humphreys and Johnson 2007). According to attractor field theory, every exemplar represented in memory is associated with a range of stimulus inputs that will activate that representation, as is the case when a familiar face is recognized despite being viewed from an unusual angle. This "attractor field" corresponds to an area of similarity space surrounding each exemplar; stimuli falling within this area are sufficiently similar to the exemplar to activate its stored representation when presented. Tanaka and colleagues proposed that items in densely populated regions of similarity space possess relatively small attractor fields because the attractor fields of many neighboring representations restrict their range. The attractor fields of items in sparsely populated regions, by contrast, are much larger owing to the lack of competition from nearby representations. Atypical exemplars, then, should generally possess larger attractor fields than typical exemplars. This theory provides a ready account of the atypicality bias: a morph which lies at the midpoint of a typical and atypical parent in similarity space is more likely to be drawn to the larger attractor field of the atypical parent, manifest perceptually as greater similarity to that parent.

Although the attractor field model provides an intuitive account of the atypicality bias and was demonstrated to be computationally feasible by neural network simulations (Tanaka et al 1998), the cognitive mechanisms underlying the phenomenon remain unknown. A principal question concerns the means by which atypical items come to be perceptually "attractive." The physical structure of the stimuli alone might not be sufficient to produce an atypicality bias. That is, because one cannot know which members of an object class are (a)typical until one has accrued experience with members of that class, the atypicality bias might emerge only after such learning has occurred. In this scenario, a given set of stimuli might shift from eliciting no bias to eliciting an atypicality bias once knowledge of the associated category structure has been acquired.

This proposal is supported by past research indicating that category learning can alter the perception of the categorized objects themselves. For example, Schyns and colleagues (Schyns and Rodet 1997; Schyns et al 1998) have demonstrated that feature learning in novel stimuli can be guided by category diagnosticity, and that the choice of "functional" (category-relevant) features early in learning can affect the perception of subsequent exemplars. By this view, the features that drive the perception of a given stimulus are not fixed structural characteristics of the stimulus but are at least in part a product of the demands of prior learning within the domain (Schyns and Rodet 1997). Identical stimuli, then, might be perceived differently before and after such learning has occurred.

Goldstone and colleagues have investigated how changes in perception with category learning might reflect changes in the dimensions of similarity space. Goldstone (1994) found that category training heightened perceptual sensitivity to category-relevant stimulus dimensions (and, in some cases, diminished sensitivity to irrelevant dimensions). Goldstone and Steyvers (2001) examined two means by which this sensitization can occur: the differentiation of relevant and irrelevant dimensions, and selective attention to relevant dimensions. Goldstone et al (2001) provided evidence that knowledge of category membership alone can influence ratings of similarity, such that within-category members are judged to be more similar to one another than between-category members on the basis of their shared category affiliation (see also Corneille et al 2006). However, these researchers also demonstrated that category training can create perceptual changes at the level of the mental representation of an object: within-category exemplars were not only perceived to be more similar to one another following training, they were perceived to be more similar to a neutral stimulus with no categorical association. All of these findings suggest that learning a category structure produces changes in similarity space, an assumption built into some models of categorization (eg Nosofsky 1986).

One factor driving such changes may be the formation of dense and sparse regions of similarity space as the central tendency of a category is learned. According to Krumhansl's (1978) distance–density model, the similarity of two exemplars is jointly determined by their distance in the space and the density of the regions in which they lie. The distance–density model gives the formula for calculating inter-item similarities as follows:

$$\mathrm{sim}(e_1, e_2) = d(e_1, e_2) + \alpha\delta(e_1) + \beta\delta(e_2)\,,$$

where e_1 and e_2 are exemplars, $d(e_1, e_2)$ is their inter-point distance in similarity space, $\delta(e_1)$ and $\delta(e_2)$ are the spatial densities of exemplars in the regions surrounding the two exemplars, and α and β are weights applied to those densities. The central prediction of the model is that two exemplars of equal physical distance will be less similar to each other if located in a dense region of the space than if located in a sparse region of the space. Since, as noted above, dense and sparse regions of space are established through experience with category members, Krumhansl's (1978) model is consistent with the possibility that category learning results in an expansion of dense regions of the space, contraction of sparse regions of the space, or both. These dynamics of similarity space may underlie the atypicality bias if the dimensions of the space are altered in such a way that a morph otherwise perceptually equidistant from its typical and atypical parent is drawn toward its atypical parent (or, alternatively, drawn away from its typical parent).

The present experiments were designed to determine whether the atypicality bias emerges as a function of perceptual experience with an object class. This possibility cannot be tested with stimuli from familiar categories such as faces, birds, and cars because the structure of those categories is well known to participants pre-experimentally. Hence, the magnitude of the atypicality bias before and after category learning cannot be established. Nor is it clear what stimulus properties (eg color, shape, size, surface properties) are driving the judgments of typicality that distinguish typical and atypical faces, birds, and cars. To address the shortcomings of real-world objects, we created shape stimuli called "blobs" (see, eg, Curran et al 2002) with which participants had no pre-experimental familiarity. Sample blobs appear in figure 1. Each category of blobs was based on a prototype whose structure varied randomly within user-defined parameters (for details of blob generation, see section 2.1). Typical category members were created by making small variations on this prototype, rendering them highly similar to the prototype and each other. Atypical members were allowed a greater magnitude of variance from the prototype, and were thus highly distinct from one another and the typical exemplars.

Importantly, each category contained an equal number of typical and atypical exemplars. Note that within natural categories such as those used in previous studies of the atypicality bias, structural typicality and frequency of occurrence are generally confounded, as atypical exemplars tend to be fewer in number and encountered less frequently than typical exemplars (Barsalou 1985). Thus, although a similarity space-based account of the atypicality bias implies that the effect is driven by the structural uniqueness of atypical exemplars, the potential contribution of the relative rarity of such exemplars complicates this interpretation. In the current study, we used artificial stimuli to disentangle these two elements of distinctiveness, holding frequency constant and manipulating only the structural typicality of blob exemplars.

Figure 1. Sample typical and atypical parent blobs and their morph.

Within each blob category, morphs of every pairwise combination of typical and atypical parents were created. Pilot testing using the same preference task as Tanaka et al (1998) established two categories for which no inherent bias was present. In experiment 1, participants completed the preference task before and after learning to discriminate exemplars of these two categories through a training task. We hypothesized that if knowledge of a category structure yields an atypicality bias, these blobs should elicit no bias before training but an atypicality bias after training. In experiment 2, the category learning task from Experiment 1 was replaced with pleasantness ratings of the same blob exemplars to determine whether explicit category training is necessary to observe an atypicality bias.

2 Experiment 1

2.1 *Method*

2.1.1 *Participants.* Twenty-five University of Victoria undergraduates participated for bonus credit in a psychology course.

2.1.2 *Materials.* Typical blobs, atypical blobs, and morphs were created with Blob Maker, a Matlab-based program. Blob Maker produces categories of blobs by creating a category prototype and varying the prototype to generate category exemplars. Exemplars were created for pilot testing in two rounds; the parameters used differed slightly between the two, and these differences are noted below.

Prototypes were created as follows: first, a circle was divided evenly by either 15 or 17 rays originating at its center. Second, a point was placed at a randomly determined position along each ray. Third, adjacent points were connected with a straight line, creating a jagged outline. The outline was filled in with a blue color to complete the prototype. Exemplars were variations on the prototype in which the positions of the points on each ray were allowed to shift relative to their locations on the prototype. Typical exemplars were allowed up to 5% or 3%–7% deviation from the prototype, while atypical exemplars were allowed up to 15% or 8%–12% deviation. Thus, within-category blobs reflected a common underlying structure (that of the prototype), with typical members more homogeneous than atypical members by virtue of closer adherence to that structure.

Category exemplars were selected for use in the pilot studies according to the judgment of the experimenters, a procedure meant to ensure that the typical blobs used in the experiment were adequately discriminable from one another, and that the atypical blobs bore sufficient resemblance to the category prototype to be identified with that category. Each typical exemplar was then morphed with each atypical exemplar. Morphs were created by averaging the lengths from the origin of each corresponding point of a given typical–atypical pair.

24 categories were pilot tested via a preference task also used in experiments 1 and 2 and described in greater detail below. Participants saw two blobs on opposing sides of the screen, one a typical parent and the other an atypical parent. The morph of the two parents then appeared in a central position, and participants indicated whether the central blob was more similar to the blob on the left or to the blob on the right. 16 parent–morph–parent triads from each pilot category were presented in randomly intermixed order.

Of the 24 categories tested, two were selected for the experiment based on the criteria that they (a) were free of bias in the preference task (ie the atypical and typical parents were chosen approximately equally often); (b) were sufficiently dissimilar to one another that successful category training was possible; and (c) were sufficiently similar to one another that category learning would require careful attention to fine-grained structural characteristics of the blobs. The two categories are displayed in figure 2. The atypical parent was selected as bearing greater similarity to the morph than the typical parent on 47.0% and 52.2% of trials in categories 1 and 2, respectively. Neither of these values differed significantly from the 50% (unbiased) level (both *p*s > 0.20). Each blob category was assigned to "family A" for half of participants and to "family B" for the other half. The experiment was conducted with E-Prime software (http://www.pstnet.com).

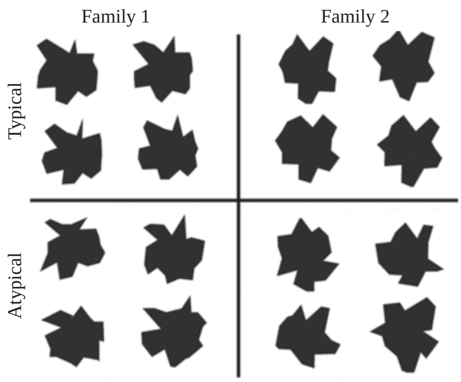

Figure 2. Typical and atypical members of the two blob families used in experiments 1 and 2.

2.1.3 *Procedure.* The experiment consisted of three phases: an initial preference task, category training to criterion, and a second, post-training preference task. Participants were tested individually. Instructions for the pre-training preference task informed participants that they would be seeing two images on opposing sides of the screen followed by a third image positioned centrally between the first two. Their task was to indicate whether the central image was more similar to the one on the left or to the one on the right. Participants were told to make quick "gut" responses and that a maximum of 3 s was allowed for each response. Each trial began with a 500 ms fixation cross positioned in the center of the screen, followed by the presentation of two parent blobs, one typical and one atypical, approximately 5 cm to the left and right of center. Typical and atypical parents were counterbalanced with respect to screen position. After 2500 ms the morph of the two parents appeared in the center of the screen along with the words "Left or Right?" below. Participants pressed the "1" key to indicate that the left blob was more similar to the center blob or the "0" key to choose the right blob. All three items remained on the screen until a response was made or the 3 s time limit elapsed, at which time the following trial began. The preference task contained 32 trials, one for each blob triad from the two categories.

The category training task followed. Participants were informed that they would be viewing a series of images, each of which belonged either to "family A" or to "family B", and that their task was to indicate to which family each image belonged. Participants were told that they would initially be guessing, but that they would learn the correct responses by paying attention to feedback on each trial. Category training was divided into blocks of 16 trials. In each block, the 8 "A" and 8 "B" blobs were presented for category judgments in a random order. On each trial, a blob was presented in the center of the screen with the words "Family A or Family B" below. Responses were made by pressing the "A" key to signify a "family A" judgment or the "L" key to signify a "family B" judgment. Blobs remained on the screen until a response was made. Responses initiated immediate feedback of "CORRECT!" in blue type or "Incorrect" in red type that remained on the screen for 1000 ms. The following trial began immediately thereafter. Categorization trials continued until three consecutive blocks were completed with no more than one error per block. Upon reaching this criterion, participants graduated to the post-training preference task.

In the post-training preference task, the same blob triads as in the pre-training task were presented in a new random order and with the positions of the typical and atypical parents reversed relative to their positions in the pre-training phase (ie a triad displayed with the arrangement atypical-morph-typical pre-training was displayed typical-morph–atypical post-training). Thus, the post-training phase consisted of 32 preference judgments.

2.2 *Results and discussion*

All twenty-five participants successfully reached the performance criterion in the category training phase and advanced to the post-training preference task. Preference task trials in which no response was given within the time limit (2.1% of trials pre-training, 0.9% of trials post-training) did not figure in the analyses of preference data.

Participants required an average of 9.4 (SD = 5.8) training blocks to reach criterion in the category training phase. Categorization accuracy was reliably higher for typical blobs than for atypical blobs ($M_{\text{typical}} = 90.6\%$, $M_{\text{atypical}} = 80.7\%$, $t_{24} = 8.508$, $p < 0.001$). Categorization RTs were reliably lower for typical blobs than for atypical blobs ($M_{\text{typical}} = 921$ ms, $M_{\text{atypical}} = 1070$ ms, $t_{24} = 4.785$, $p < 0.001$).

The mean pre- and post-training atypicality bias is displayed for each family in figure 3. In the pre-training preference task, the atypical parent was chosen as more similar to the morph on 52.3% of trials; after category training, the atypical parent was chosen on 58.3% of trials. A 2 (preference phase: pre-training or post-training) × 2 (family: 1 or 2) repeated-measures ANOVA indicated a significant increase in the atypicality bias from pre- to post-training ($F_{1,\,24} = 8.980$, $p < 0.01$, $\eta_{\text{p}}^2 = 0.272$). Neither the main effect of family nor the preference phase × family interaction were significant (both ps < 0.37). t-tests confirmed that the pre-training bias was not significantly greater than the 50% level in either family ($M_{\text{family1}} = 50.3\%$, $M_{\text{family2}} = 54.5\%$), though it approached significance in the latter case ($p = 0.08$). The level of bias following training was significantly greater than 50% in both families ($M_{\text{Family1}} = 57.9\%$, $t_{24} = 3.542$, $p < 0.01$; $M_{\text{family2}} = 58.7\%$, $t_{24} = 2.972$, $p < 0.01$). The increase in atypicality bias from pre-training to post-training was also significantly greater than 50% when measured by item ($t_{31} = 2.232$, $p < 0.05$), indicating that the bias was not driven by a minority of parent–morph–parent triads but was consistently observed across the stimulus set. Overall, nineteen out of twenty-five participants showed a directional increase in atypicality bias after category training.

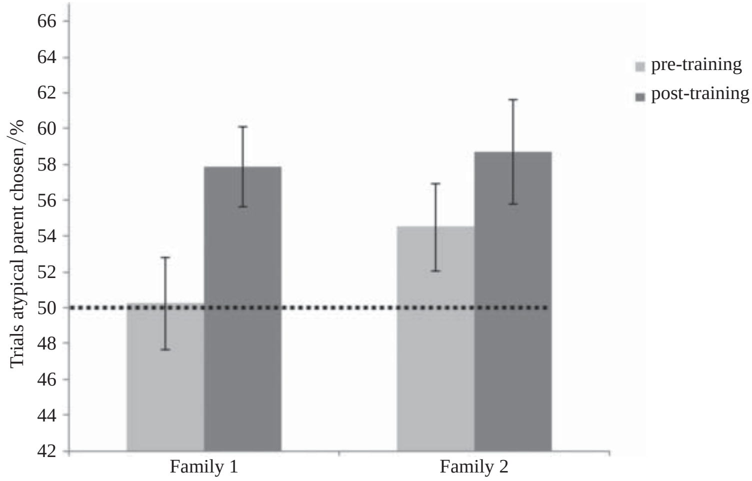

Figure 3. Percentage of trials on which the morph was called more similar to the atypical parent before and after category training in experiment 1. Error bars represent one standard error of the mean. The dashed line represents the 50% (unbiased) level.

These results support the hypothesis that the atypicality bias is observed following, but not prior to, category learning within a novel stimulus domain. They do not, however, imply that the explicit category training undertaken in experiment 1 is necessary to develop an atypicality bias. Prototype extraction tasks have demonstrated that information about family resemblance within a stimulus domain can be gleaned incidentally from exposure to exemplars in non-categorization tasks (eg Knowlton and Squire 1993; Smith 1998). If such information drives the atypicality bias, it might be the case that any task involving visual examination of the stimuli would foster incidental feature learning sufficient to produce the bias. We tested this possibility in experiment 2 by replacing the category training participants received in experiment 1 with a blob pleasantness rating task. If the visual analysis required to judge the pleasantness of blobs confers enough information about their normative structural attributes to allow learning of the associated category structures, participants' preference task responses should evidence an atypicality bias following, but not prior to, the pleasantness rating phase. This result would suggest that the atypicality bias arises under general experience-based learning conditions and is not limited to circumstances in which item categories are learned explicitly.

3 Experiment 2

3.1 *Method*

3.1.1 *Participants.* Thirty University of Victoria undergraduates participated for bonus credit in a psychology course.

3.1.2 *Materials.* The materials were identical to those used in experiment 1.

3.1.3 *Procedure.* The procedure was identical to that of experiment 1, except that a pleasantness rating task was administered between the two preference-judgment tasks instead of category training. Prior to making the pleasantness ratings, participants were informed that they would be presented with a series of images and that they would be asked to rate the pleasantness of each image on a scale of 1 ("not pleasant") to 7 ("very pleasant"). Participants were encouraged to use the entire scale in making their judgments and were told to make quick responses based on their first instinct. As in experiment 1, they were informed that they would be seeing the same group of 16 images over repeated blocks.

Blobs were presented in the center of the screen and remained until the participant made a response via key press. A blank 750 ms interval separated each trial. As in experiment 1, trials were organized in blocks consisting of a random ordering of the 8 exemplars from each family. As there was no basis for establishing a criterial performance level akin to that of the category training in experiment 1, each participant completed 10 blocks, a number chosen to approximate the average amount of exemplar exposure participants in experiment 1 received during category training ($M = 9.4$ blocks). Therefore, judgments were made to each of the 16 blobs 10 times for a total of 160 pleasantness rating trials.

3.2 *Results and discussion*

The data of one participant repeatedly failing to respond within the time limit during the preference task were removed prior to analysis. Among the remaining participants, responses not made before the deadline comprised 1.9% and 1.5% of trials in the first and second preference tasks, respectively, and were not part of subsequent analyses.

Pleasantness ratings for typical and atypical blobs did not differ significantly ($M_{\text{typical}} = 3.73$, $M_{\text{atypical}} = 3.53$, $t_{28} = 1.701$, $p = 0.10$). RTs during the pleasantness task were similar for typical and atypical blobs ($M_{\text{typical}} = 1590$ ms, $M_{\text{atypical}} = 1650$ ms, $t_{28} = 1.264$, $p = 0.22$).

Preference task results were very similar to those observed in experiment 1 (see figure 4). The atypical parent was chosen as more similar to the morph on 53.1% of trials in the initial preference task and on 59.6% of trials following the pleasantness ratings task. A preference phase × family ANOVA revealed that this increase in the atypicality bias from the pre- to post-pleasantness task was significant ($F_{1,\,28} = 4.680$, $p < 0.05$, $\eta_{\text{p}}^2 = 0.143$). The magnitude of the increase in bias was moderately greater for family 2 than for family 1, but the associated preference phase × family interaction was not significant ($p = 0.13$). The main effect of family was non-significant ($p = 0.52$). *t*-tests again confirmed that the atypicality bias did not significantly exceed the 50% level for either family in the first preference task ($M_{\text{family1}} = 53.8\%$, $M_{\text{family2}} = 52.7\%$; both *p*s > 0.22), but was significantly greater than 50% for both families in the second ($M_{\text{family1}} = 57.0\%$, $t_{28} = 3.061$, $p < 0.01$; $M_{\text{family2}} = 62.1\%$, $t_{28} = 3.765$, $p < 0.001$). As in experiment 1, the increase in atypicality bias following the pleasantness task was also significantly greater than 50% when measured by item ($t_{31} = 2.730$, $p < 0.05$). Overall, twenty out of twenty-nine participants showed a directional increase in atypicality bias after the pleasantness rating task.

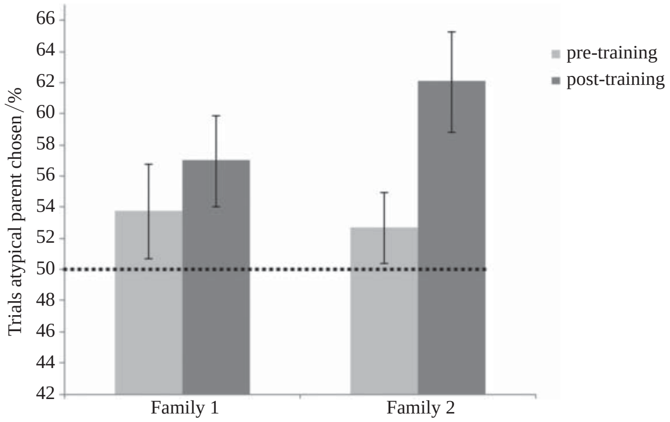

Figure 4. Percentage of trials on which the morph was called more similar to the atypical parent before and after the pleasantness rating task in experiment 2. Error bars represent one standard error of the mean. The dashed line represents the 50% (unbiased) level.

The results of experiment 2, then, were strikingly similar to those of experiment 1. The central finding, a significant increase in the atypicality bias following the pleasantness rating task, suggests that the visual inspection of the blob stimuli necessary to produce ratings of pleasantness was sufficient to make salient the distinction between typical and atypical exemplars, and that this indirect form of training changed the perception of those exemplars. Indeed, the 6.5% increase in the atypicality bias observed in experiment 2 mirrored the 6.1% increase following explicit category training in experiment 1. The similar magnitudes of these increases suggest that category training and pleasantness ratings exerted a comparable influence on the perception of the blob exemplars. We discuss the implications of this result below.

4 General discussion

These results support the hypothesis that the atypicality bias represents a change in object perception that develops as a function of perceptual experience with a given class of stimuli. The blob morphs used in the present experiments were called more similar to their typical and atypical parents with approximately equal frequency when first encountered, as measured by pilot testing and the pre-training preference tasks. After repeated exposure to the parent exemplars via category training (experiment 1) or a pleasantness rating task (experiment 2), however, participants selected the atypical parent as more similar to the morph on a significantly greater proportion of trials. This finding suggests that previous demonstrations of the atypicality bias (Tanaka and Corneille 2007; Tanaka et al 1998) did not require a learning phase because they used stimuli from commonly encountered object classes (faces, birds, and cars) for which the analogous experience had already accrued through everyday exposure. The present experiments may be viewed as having captured the emergence of the atypicality bias during a single laboratory session.

What aspects of category training lead to the development of the atypicality bias? In the process of discerning the structural features of the stimuli (or informative fragments thereof; eg Hedge et al 2008) that are diagnostic of category membership, individuals become aware of the normative appearance of those features as well as deviations from the norm, thus establishing a sense of typical and atypical category exemplars, respectively. Evidence that this distinction took place in the current experiments comes from the result that atypical exemplars were categorized slower and less accurately than typical exemplars in experiment 1, a common finding in research on natural categories (McCloskey and Glucksberg 1979; Murphy and Brownell 1985). We propose that learning what constitutes category typicality and atypicality alters the mental representation of the exemplars within that category, and that this alteration results in the observed atypicality bias.

The relative difficulty of classifying atypical exemplars in experiment 1 mirrors results from research on natural categories, and thus represents a desirable quality of the present blob stimuli. However, it also constitutes a confounding of structural typicality and ease of classification. Consequently, it is possible that the increased selection of atypical exemplars in the subsequent preference task was influenced by the fact that these exemplars posed a particular challenge in the category training phase. Indeed, as Nosofsky (1991) has discussed in detail, stimulus-driven biases can be difficult to disentangle from response biases in judgments of similarity. In the present case, evidence suggests that the atypicality bias was not driven by the difficulty of categorizing atypicals per se. In experiment 2, a significant atypicality bias followed a pleasantness rating task in which explicit categorization judgments were not made and neither mean ratings nor RT differed as a function of typicality. Moreover, in a previous experiment (not reported here) we tested a different set of blobs in which the typical items from one category were very similar to the typical items from the second category. As a result of this similarity, participants were faster and more accurate in categorizing atypical exemplars, the opposite of the pattern found in the present experiments. Despite the atypical advantage in categorization, an atypicality bias of roughly the same magnitude as those reported here emerged following category training. Thus, the learning of category structure, not categorization performance on typical versus atypical items, appears to be the critical element in producing the atypicality bias.

If category learning is the mechanism underlying the atypicality bias, why was the pleasantness rating task in experiment 2 associated with as great an increase in the bias as the category training task in experiment 1? A partial explanation might be a ceiling on the level of atypicality bias measurable via the preference task. Previous experiments conducted in our lab have consistently found atypicality bias levels at roughly 60%, as has past work using stimuli from natural categories (Tanaka and Corneille 2007; Tanaka et al 1998), though bird stimuli may an exception (Tanaka and Corneille 2007; Tanaka et al 2011). Perhaps category training is more directly associated with the development of the atypicality bias, but both the category training and pleasantness rating manipulations in the present experiments were sufficient to raise the atypicality bias to maximal levels.

Whether or not such a limit characterizes the measurement of the atypicality bias, the significant increase in the bias following pleasantness ratings suggests that feedback-driven category training is not necessary for its emergence. Rather, the atypicality bias appears to arise via the ability of the perceptual system to discern rapidly the structural regularities of a category of stimuli even when the task at hand does not necessitate category learning. By this reasoning, any form of perceptual experience with the stimuli that makes evident the distinction between typical and atypical category members might be expected to produce the effect. This account of the present results dovetails with the Tanaka et al (2011) finding of an atypicality bias in 3–4-year-old children, who have amassed far less perceptual categorization experience than adults. Given the broad boundary conditions that characterize the effect, the atypicality bias would appear to be a pervasive perceptual phenomenon.

We argued above that the emergence of the atypicality bias with experience can be couched in terms of theories that express exemplar similarities as a function of their distances in a multidimensional psychological space (eg Goldstone 1994; Krumhansl 1978; Tanaka et al 1998). We believe that the present data are consistent with the view that: (a) the accrual of experience with a category of stimuli populates the central and peripheral regions of the associated similarity space with typical and atypical category members, respectively; (b) the dimensions of the space are altered in the process, becoming relatively more compact in the sparsely populated regions of the similarity space than in the densely packed regions and producing altered similarity relationships between exemplars; and (c) these relationships conform to the predictions of Krumhansl's (1978) distance–density model, which holds that items in sparsely populated regions of the space will be more similar to each other than items of equivalent physical distance in densely populated regions. We depict these general dynamics of similarity space in figure 5.

The hypothesized changes in similarity space with category learning depicted in figure 5 and the resulting emergence of the atypicality bias are also broadly consistent with Love and colleagues' supervised and unsupervised stratified adaptive incremental network (SUSTAIN) model of category learning (Love et al 2004). SUSTAIN represents similar category exemplars (such as the typical blobs in the present experiments) as clusters embedded in a multidimensional similarity space. The model forms new clusters in remote areas of similarity space when distinctive items are encountered, a mechanism that could allow it to model the effects of learning on the atypicality bias reported here. SUSTAIN also allows for both supervised and unsupervised learning of category structure, consistent with the present

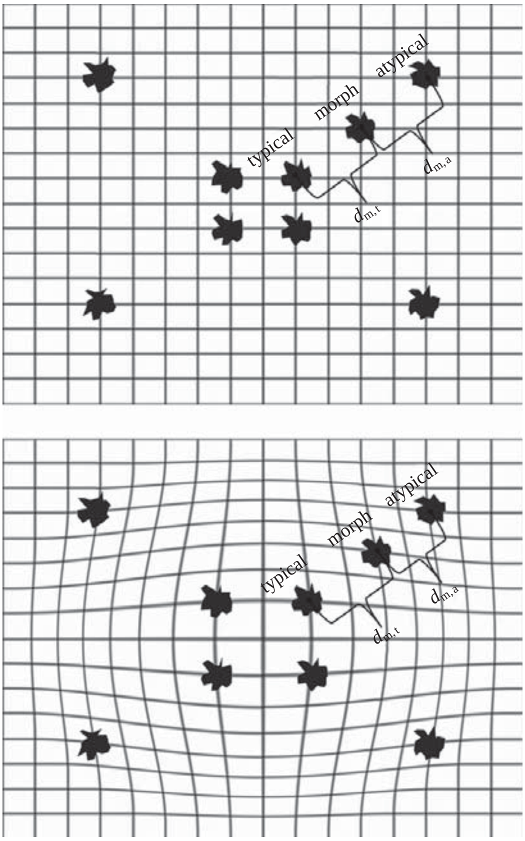

Figure 5. A depiction of hypothetical changes in similarity space with stimulus experience theorized to underlie the development of the atypicality bias. Blobs displayed constitute family 1 from the current experiments. Atypical exemplars are positioned at the periphery of the space; typical exemplars occupy the central portions of the space. Before experience with category exemplars is accrued (top panel), psychological distances between exemplars are characterized by a Euclidean metric across the space, and a 50–50 morph of a typical and an atypical exemplar is equidistant from both parents ($d_{\mathrm{m,\,t}} = d_{\mathrm{m,\,a}}$). After perceptual experience (bottom panel), the dimensions of the space are stretched as a decreasing function of distance from the center of the space. The resulting shift in the relative positions of the exemplars produces the atypicality bias ($d_{\mathrm{m,\,t}} > d_{\mathrm{m,\,a}}$).

finding of an atypicality bias following both feedback-driven (experiment 1) and relatively implicit (experiment 2) category learning.

The result of such learning is that a morph positioned at the midpoint of the physical distance between two parents will appear more similar to the parent residing in the more compact, sparsely populated region of the space, producing the atypicality bias reported here. The present experiments serve as further reminders that our perceptions of similarity are not fixed, but are constantly being formed and reformed by category experience.

Acknowledgments. This research was supported by an Erich and Shelley Mohr Fellowship to Justin Kantner and grants from the James S McDonnell Foundation, the National Science Foundation (#SBE-0542013), and the Natural Sciences and Engineering Research Council of Canada to Jim Tanaka. We thank Tamara Meixner, Laura Dixon, Brie MacDonald, Chelsea Durber, and Blaire Webster for their assistance.

References

Barsalou L W, 1985 "Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories" *Journal of Experimental Psychology: Learning, Memory, and Cognition* **11** 629–654

Corneille O, Goldstone R L, Queller S, Potter T, 2006 "Asymmetries in categorization, perceptual discrimination, and visual search for reference and nonreference exemplars" *Memory & Cognition* **34** 556–567

Curran T, Tanaka J W, Weiskopf D M, 2002 "An electrophysiological comparison of visual categorization and recognition memory" *Cognitive, Affective, & Behavioral Neuroscience* **2** 1–18

Cutzu F, Edelman S, 1996 "Faithful representation of similarities among three-dimensional shapes in human vision" *Proceedings of the National Academy of Sciences of the USA* **93** 12046–12050

Goldstone R L, 1994 "Influences of categorization on perceptual discrimination" *Journal of Experimental Psychology: General* **123** 178–200

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

Goldstone R L, Steyvers M, 2001 "The sensitization and differentiation of dimensions during category learning" *Journal of Experimental Psychology: General* **130** 116–139

Hedge J, Bart E, Kersten D, 2008 "Fragment-based learning of visual object categories" *Current Biology* **18** 597–601

Humphreys K, Johnson M H, 2007 "The development of 'face-space' in infancy" *Visual Cognition* **15** 578–598

Iverson P, Kuhl P K, 1995 "Mapping the Perceptual Magnet Effect for speech using signal detection theory and multidimensional scaling" *Journal of the Acoustical Society of America* **97** 553–562

Iverson P, Kuhl P K, 2000 "Perceptual magnet and phoneme boundary effects in speech perception: Do they arise from a common mechanism?" *Perception & Psychophysics* **62** 874–886

Knowlton B J, Squire L R, 1993 "The learning of categories: Parallel brain systems for item memory and category knowledge" *Science* **262** 1747–1749

Krumhansl C L, 1978 "Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density" *Psychological Review* **85** 445–463

Lee K, Byatt G, Rhodes G, 2000 "Caricature effects, distinctiveness, and identification: Testing the face-space framework" *Psychological Science* **11** 379–385

Leopold D A, O'Toole A J, Vetter T, Blanz V, 2001 "Prototype-referenced shape encoding revealed by high-level aftereffects" *Nature Neuroscience* **4** 89–94

Love B C, Medin D L, Gureckis T M, 2004 "SUSTAIN: A network model of category learning" *Psychological Review* **111** 309–332

McCloskey M, Glucksberg S, 1979 "Decision processes in verifying category membership statements: Implications for models of semantic memory" *Cognitive Psychology* **11** 1–37

Murphy G L, Brownell H H, 1985 "Category differentiation in object recognition: typicality constraints on the basic category advantage" *Journal of Experimental Psychology: Learning, Memory, and Cognition* **11** 70–84

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

Nosofsky R M, 1991 "Stimulus bias, asymmetric similarity, and classification" *Cognitive Psychology* **23** 94–140

Schyns P G, Goldstone R L, Thibaut J, 1998 "The development of features in object concepts" *Behavioral and Brain Sciences* **21** 1–54

Schyns P G, Rodet L, 1997 "Categorization creates functional features" *Journal of Experimental Psychology: Learning, Memory, and Cognition* **23** 681–696

Shepard R N, 1962 "The analysis of proximities: Multidimensional scaling with an unknown distance function. Part I" *Psychometrika* **27** 125–140

Smith E E, 1998 "The case for implicit category learning" *Cognitive, Affective, & Behavioral Neuroscience* **8** 3–16

Tanaka J W, Corneille O, 2007 "Typicality effects in face and object perception: Further evidence for the attractor field model" *Perception & Psychophysics* **69** 619–627

Tanaka J, Giles M, Kremen S, Simon V, 1998 “Mapping attractor fields in face space: The atypicality of bias in face recognition” *Cognition* **68** 199–220

Tanaka J W, Meixner T L, Kantner J, 2011 “Exploring the perceptual spaces of faces, cars and birds in children and adults” *Developmental Science* **14** 762–768

Valentine T, 1991 “A unified account of the effects of distinctiveness, inversion, and race in face recognition” *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology* **43** 161–204

Valentine T, 2001 “Face–space models of face recognition”, in *Computational, Geometric, and Process Perspectives on Facial Cognition: Contexts and Challenges* Eds M J Wenger, J T Townsend (Mahwah, NJ: Lawrence Erlbaum Associates) pp 83–113