
Biederman, Irving, Yue, Xiaomin and Davidoff, Jules B.

Representation of Shape in Individuals From a Culture With Minimal Exposure to Regular, Simple Artifacts: Sensitivity to Nonaccidental Versus Metric Properties

You may cite this version as: Biederman, Irving, Yue, Xiaomin and Davidoff, Jules B.. 2009. Representation of Shape in Individuals From a Culture With Minimal Exposure to Regular, Simple Artifacts: Sensitivity to Nonaccidental Versus Metric Properties. Psychological Science, 20(12), pp. 1437-1442. ISSN 09567976 [Article] : Goldsmiths Research Online.

Available at: http://eprints.gold.ac.uk/4930/

PSYCHOLOGICAL SCIENCE

Research Report

Representation of Shape in Individuals From a Culture With Minimal Exposure to Regular, Simple Artifacts

Sensitivity to Nonaccidental Versus Metric Properties

Irving Biederman,[1] Xiaomin Yue,[2] and Jules Davidoff[3]

[1]Department of Psychology/Neuroscience, University of Southern California; [2]Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School; and [3]Centre for Cognition, Computation and Culture, Department of Psychology, Goldsmiths, University of London

ABSTRACT—*Many of the phenomena underlying shape recognition can be derived from the greater sensitivity to nonaccidental properties of an image (e.g., whether a contour is straight or curved), which are invariant to orientation in depth, than to the metric properties of an image (e.g., a contour's degree of curvature), which can vary with orientation. What enables this sensitivity? One explanation is it derives from people's immersion in a manufactured world in which simple, regular shapes distinguished by nonaccidental properties abound (e.g., a can, a brick), and toddlers are encouraged to play with toy shape sorters. This report provides evidence against this explanation. The Himba, a seminomadic people living in a remote region of northwestern Namibia where there is little exposure to regular, simple artifacts, were virtually identical to Western observers in their greater sensitivity to nonaccidental properties than to metric properties of simple shapes.*

Do individuals from a culture with only limited exposure to developed-world artifacts have the same kinds of shape representations as those evidenced by the typical artifact-immersed laboratory subject? To answer this question, we turned to the Himba, a seminomadic people who live in a remote region of northwestern Namibia and have little contact with the manufactured products that are so prevalent in daily life in developed societies. Figures 1a through 1d show scenes that are typical of the Himba environment.

We compared representations of shape between the Himba and individuals immersed in the artifacts of the developed world. Specifically, we assessed sensitivity to differences in nonaccidental properties (NAPs) and metric properties (MPs) of simple shapes (geons). A NAP is a viewpoint-invariant characteristic of an image that provides strong evidence that the property is true of the object projecting that image. Unlike MPs, such as degree of curvature, which can vary continuously with rotation in depth, NAPs tend not to vary under such rotations—for example, straight contours in the image remain straight, and curved contours remain curved, at almost all orientations of the object. For both types of stimulus variation, we used a model of V1-like Gabor filters to scale the physical similarity of the to-be-discriminated shapes (see Scaling Shape Similarity). This model predicts psychophysical shape similarity almost perfectly as long as there are no nonaccidental or part differences between the shapes (Lades et al., 1993; Yue, Tjan, & Biederman, 2006).

THEORETICAL BACKGROUND

Why study the Himba? The Himba provide an opportunity to assess the effects of lack of exposure to artifacts on the repre-

---

Address correspondence to Irving Biederman, University of Southern California, Department of Psychology/Neuroscience, 3641 Watt Way, Hedco Neuroscience Building, Room 316, Los Angeles, CA 90089-2520, e-mail: bieder@usc.edu.

PSCI 2465

P S C I	2 4 6 5	B	Dispatch: 6.10.09	Journal: PSCI	CE: Blackwell
Journal Name	Manuscript No.		Author Received:	No. of pages: 6	PE: TMS

Q1

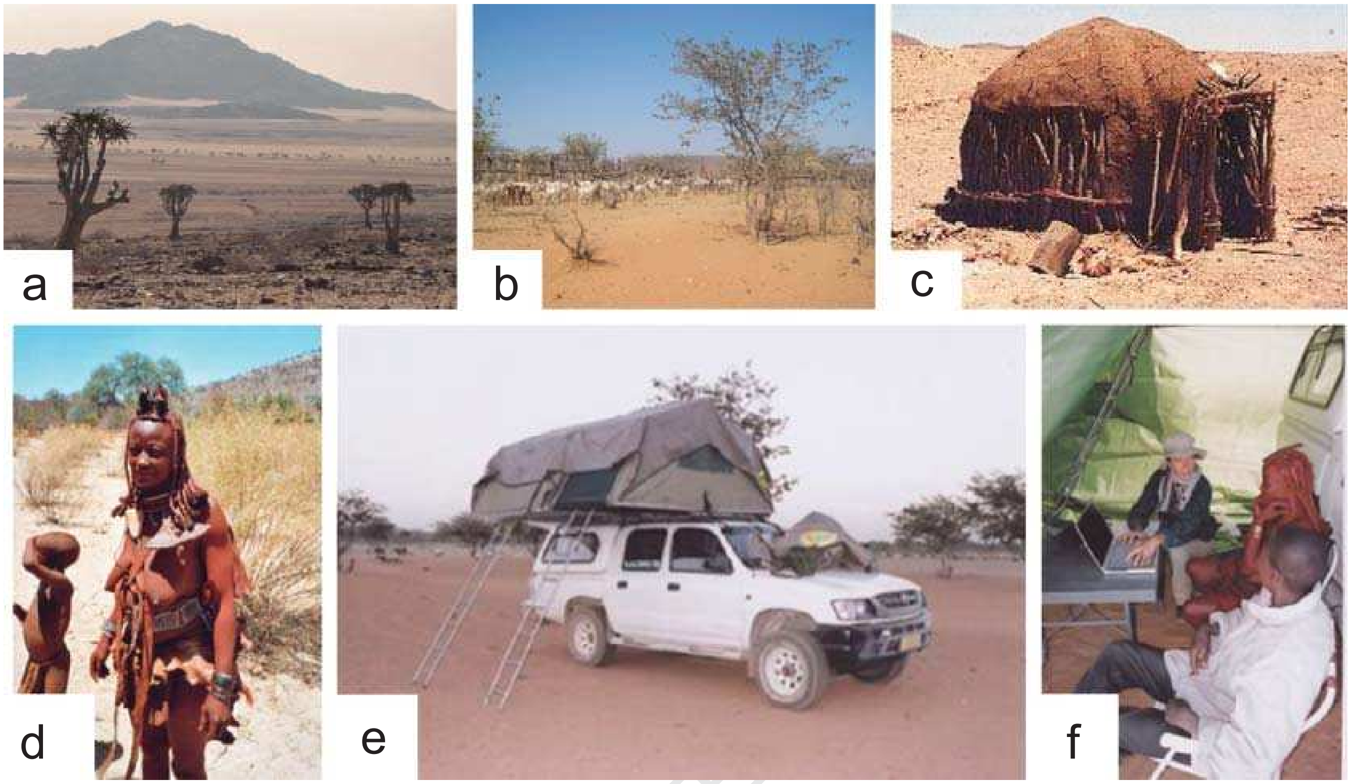

Fig. 1. Typical scenes in the world of the Himba and the experimental setup for the Himba. The typical scenes show (a) a landscape in northwestern Namibia, the Himba homeland; (b) a Himba encampment (the stick fence in the background serves as a corral for the livestock); (c) a dung-and-stick hut typical of Himba dwellings; and (d) a mother and child. The photos in (e) and (f) show the Old World Image Understanding Laboratory. A tarpaulin was draped over the ladders on the vehicle to create an enclosure (f) where images on the computer screen would not be washed out by sunlight. The photo in (f) shows the guide, a subject, and the experimenter, from front to back. As noted in the text, the Himba were reluctant to touch the computer, so the experimenter keyed in their responses while maintaining her gaze on the subject (rather than the computer screen).

sentation of shape. Specifically, the availability of simple manufactured objects (e.g., cans, bricks) that are distinguished solely by NAPs might allow facile learning of such differences. In the extreme, toddlers in developed-world environments are encouraged to play with toy shape sorters (see Fig. 2) that allow direct comparisons between contrasting NAPs. The issue under examination is not whether straight versus curved, or parallel versus nonparallel, contours are present at different frequencies in Himba versus developed-world environments. We assume that there are no differences in such frequencies. Rather, the issue is whether the opportunity for direct contrasts afforded by simple objects affects sensitivity to those differences. That the low-level image statistics of scale and orientation do not differ between natural and artifactual scenes is supported by studies of Switkes, Mayer, and Sloan (1978) and Tadmor and Tolhurst (1994).

Herero, the language of the Himba, includes few of the terms for simple shapes (e.g., "square," "circle," and "triangle") or for shape characteristics (e.g., "parallel") that are common in languages of developed societies (Roberson, Davidoff, & Shapiro, 2002). In a particularly revealing demonstration by Roberson et al. that preceded their main experiment, 2 Himba participants, separated by a panel so they could not see each other, had identical sheets of paper with 21 shapes—seven variants each of a circle, a square, and a triangle. The variants were created, for example, by adding a gap in a contour, changing a contour from straight to curved (or vice versa in the case of the circle), lengthening a portion of a figure, and drawing a somewhat irregular version of the basic shape by hand. The task was for 1 participant to communicate to the other which one of the shapes was currently being designated. This task would be trivially easy

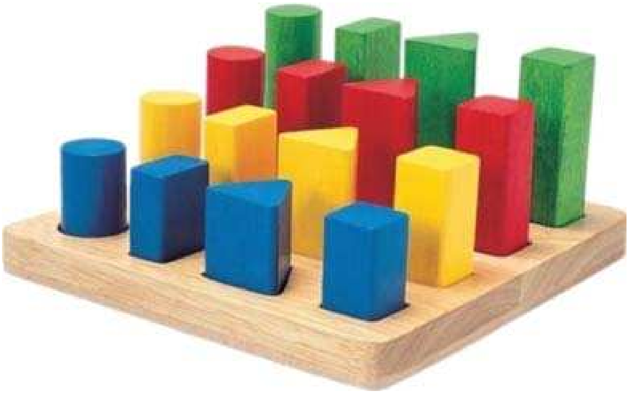

Fig. 2. An example of a toy shape sorter (Plan Toy Geometric Peg Board; PlanToys, Inc., Palo Alto, CA).

for Westerners, who could make statements such as "the circle with a gap" or "the square with wobbly lines." The Himba were at chance in determining the general shape categories (circle, square, or triangle) of the targets. Spontaneous groupings and paired-associate learning by the Himba also failed to reflect what appear to be obvious shape categories (to Westerners).

Although there was some uncertainty as to how the Himba interpreted the tasks, Roberson et al. (2002) concluded that grouping according to Western prototypical shape categories is the product of convention and language. Following Willats (1992), who on the basis of children's drawings argued that the most basic shape categories are "lumps" and "sticks," Roberson et al. suggested that the Himba might have regarded all the prototypes as lumps. Given that the prototypes differed so markedly in NAPs, this would suggest that the Himba were less sensitive to NAPs than would be expected on the basis of the performance of Western participants. However, Roberson et al. did allow the possibility that their task might have reflected language-influenced cognitive groupings rather than basic perceptual processes.

Humans (in developed-world environments) and laboratory animals show greater sensitivity to NAPs than to MPs (e.g., Biederman, 1995; Biederman & Bar, 1999; Lazareva, Wasserman, & Biederman, 2008; Logothetis, Pauls, Bülthoff, & Poggio, 1994). This greater sensitivity to NAPs has also been found in studies of inferotemporal cell tuning of macaques (Kayaert, Biederman, & Vogels, 2003, 2005; Vogels, Biederman, Bar, & Lorincz, 2001). However, the inference that such results provide strong evidence for a culture-free development of shape coding can be challenged. Laboratory animals, if anything, live in a more "geonic" environment than do humans living in a developed world.

The possibility that the perception of shape by the Himba may be similar to the perception of shape by people from the developed world was suggested by Lescroart, Biederman, Yue, and Davidoff (in press), who investigated whether the Himba would spontaneously use the same generalized cone dimensions as subjects from developed countries. That study involved a texture-segregation task in which subjects viewed a series of 5 × 5 arrays of curved cylinders (resembling macaronis). The cylinders differed, metrically, in the curvature of the axis (slightly curved vs. highly curved) and in aspect ratio (thin vs. thick). Each subject, a Himba or a University of Southern California (USC) student, had to indicate whether the border between two texture fields in each array was vertical or horizontal. The border, which was always between the second and third or third and fourth columns or rows (i.e., one field had 10 macaronis and the other 25), could be defined by one dimension (e.g., thin on one side and thick on the other, with curvature varying randomly within the field) or by a conjunction of both dimensions (e.g., thin macaronis with high curvature and thick macaronis with low curvature in one field and thick macaronis with high curvature and thin macaronis with low curvature in the other). If the dimensions were psychologically accessible, then when the border was defined by only one dimension, subjects could selectively attend to that dimension and ignore the variation in the other; in this case, there would be a performance advantage for the single-dimension task relative to the conjunction task. If the dimensions were not psychologically accessible, subjects would not be able to selectively attend to only one of the two dimensions in the single-dimension task. Instead they would have to attend to each macaroni individually, and, because there were 25 identical macaronis in all displays, there would be no difference between the conjunction and single-dimension tasks. The Himba showed the same advantage as the USC students for the single-dimension over the conjunction task, which suggests that both groups exploited the dimensional structure in the single-dimension displays to achieve texture segregation.

SCALING SHAPE SIMILARITY

Principled comparisons of the sensitivity to different variations of shapes would be impossible without a scaling of the physical differences between stimuli. Until recently, this problem appeared to be an unsolvable apples-and-oranges dilemma. How large of a difference in an MP (e.g., degree of curvature) is equivalent to a given NAP difference (e.g., straight vs. curved)? Our solution to this problem built on what is known about the ventral pathway for shape representation.

The tuning of V1 simple cells can be well modeled as a Gabor filter. A *Gabor jet* (Lades et al., 1993) is a column of multiscale, multioriented Gabor filters, with the receptive fields of the filters within the jet centered on a common point in the visual field. A jet thus models a complex cell hypercolumn. In the model of Lades et al., a lattice of Gabor jets covers the visual field. A measure of early-stage physical similarity between two stimuli can be computed by correlating the activation values that a pair of stimuli produce in the Gabor kernels of these jets. (See Yue et al., 2006, or Fiser, Biederman, & Cooper, 1996, for a discussion of how Gabor similarity is computed.) A higher correlation indicates greater similarity. This measure of image similarity predicts discriminability of metrically varying complex shapes almost perfectly. In a match-to-sample task with highly similar faces and blobs resembling teeth, correlations between error rates and the calculated similarity between the distractor and match (which was identical to the sample) were in the mid .90s (Yue, Subramaniam, & Biederman, 2007).

Our rationale for using the Gabor measure was that we assumed that the cortical processes that might differentiate sensitivity to NAPs and sensitivity to MPs are an outgrowth of cell tuning in stages after V1 (and, likely, after V4 as well). The previously cited studies (Kayaert et al., 2003; Vogels et al., 2001) documenting greater sensitivity in macaque inferotemporal cortex to NAPs than to MPs equated the two kinds of stimulus differences according to a V1 measure of similarity (the Gabor-jet model or its near equivalent). By equating stimulus variation according to a V1 measure of similarity, we had some confidence that any observed differences in the sensitivity to different kinds of stimulus variations (i.e., NAPs vs. MPs) would be generated by later stages of visual processing beyond V1.

METHOD

Logistics

Data from the Himba participants were collected on 6-day excursions in a four-wheel-drive vehicle (Fig. 1e). These excursions started from (and returned to, for refueling, charging batteries, and provisioning) Opuwo, a township at the edge of Himba territory (and the home of the guide). Because of increasing contact of the Himba with developed-world institutions, it was necessary to go to remote regions, 1- or 2-day drives from Opuwo, to search for tribal encampments. The Himba are seminomadic, so searching was indeed necessary. When an encampment was encountered, the guide would approach the village chief (or stand-in, if the chief was away) and ask for permission to camp on the outskirts of the Himba compound and have members of the tribe participate in the experiment. A gift of 0.5 kg of sugar was made to the headman to show respect and to ask for his permission to engage the tribespeople in the experiment. Participants were compensated with 0.5 kg of maize (corn meal) regardless of whether they finished the experiment.

So that the images on the computer screen would not be washed out by sunlight, a tarpaulin was draped over the ladders of the vehicle (Fig. 1e) to create an enclosure, as shown in Figure 1f. (The enclosure, unfortunately, often served as an attractant to shade-seeking goats.)

The Task

The experiment employed a match-to-sample task in which the sample was a simple, regular geon that appeared to be three-dimensional. The sample was to be matched against one of two other geons, which differed either in an MP (see Fig. 3, left panel) or in a NAP (see Fig. 3, right panel). The matching shape was, in all cases, identical to the sample. Figure 4 shows how the same sample shape could have either an MP or a NAP distractor. The stimuli were the same shapes used in Kayaert et al. (2003). On each trial, the difference between the nonmatching (distractor) shape and the matching shape defined the degree of Gabor similarity and whether the difference involved a NAP or an MP.

Subjects and Procedure

A total of 15 Himba and 8 USC students served as subjects for data analysis. Each Himba (who completed the experiment) performed 366 trials following 32 practice trials. In addition, 10 Himba were excluded for failure to reach a criterion of 70% accuracy or for reluctance to complete the task. (However, inclusion of data from the excluded subjects would have slightly increased the NAP advantage.) The USC subjects performed twice as many (732) trials. The stimuli for each subject were presented in randomly appearing sequences of NAP and MP trials, sampled from a large set of possible shapes.

Stimuli were presented and responses recorded on a Macintosh G3 laptop with a 15-in. screen. Because the Himba were reluctant to touch the computer, the experimenter pressed the response keys while maintaining her gaze on the subjects (rather than the computer screen; see Fig. 1f); the subjects pointed to the location (left or right) of the matching stimulus. For the Himba, displays were terminated by the key press. Pilot testing indicated that with unlimited display times, the USC subjects' performance would be at ceiling. Consequently, displays were presented for 300 ms for the students. Error feedback was conveyed by a beep. The USC subjects responded by pressing the left and right arrow keys on the computer keyboard.

RESULTS

The data were grouped, over the Gabor similarity values, into six bins, each holding an equal number of trials. Figure 5 shows error rates as a function of subject group (Himba or USC), physical (Gabor-jet) similarity of the distractor to the matching stimulus, and the nature of the difference between the distractor and the matching stimulus (NAP or MP). Only five of the six levels of similarity are shown, because there were no errors for

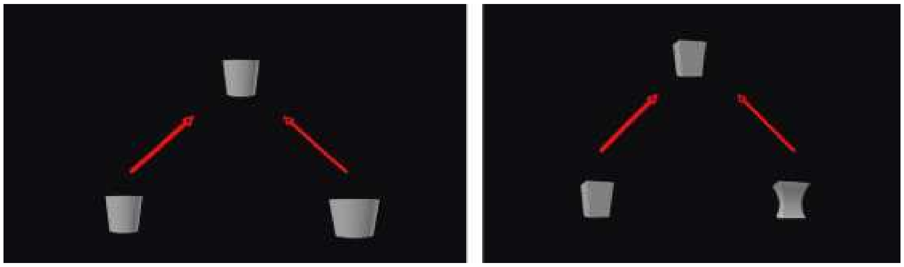

Fig. 3. Two examples of the match-to-sample trials. In each illustration, the sample is at the top, and the distractor and matching stimulus are in the bottom row. In the trial illustrated on the left, the distractor (on the right) differs from the matching stimulus in a metric property. In the trial illustrated on the right, the distractor (on the right) differs in from the matching stimulus in a nonaccidental property.

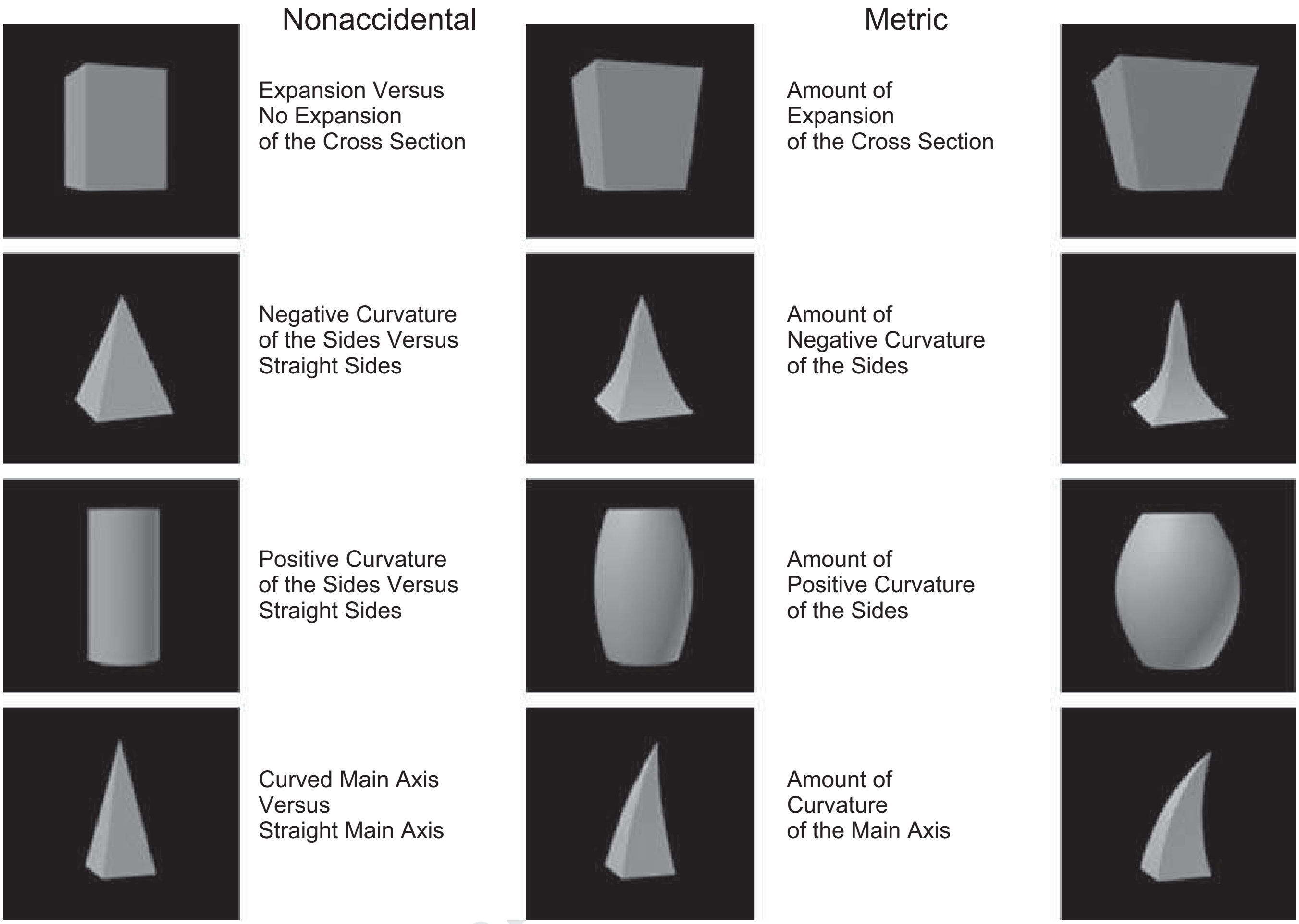

Fig. 4. Examples of stimulus differences in nonaccidental properties (NAPs) and metric properties (MPs). Within each row, the stimulus in the center column is a basic shape that served as a sample in the match-to-sample task. In each row, the shape change between the center and left column is a NAP difference that creates a different geon. The shape change between the center and right column is an MP difference that does not change the geon. Within each row, the physical image differences between the left and center columns and between the right and center columns are equivalent, as measured by the Gabor-jet model.

stimuli at the highest level of dissimilarity. Himba and USC subjects showed almost identical advantages (lower error rates) for NAP compared with MP differences: 19% for the Himba, $t(14) = 10.08$, $p_{\mathrm{rep}} > .99$, $d = 2.60$; 18% for the USC students, $t(5) = 6.06$, $p_{\mathrm{rep}} = .89$, $d = 1.24$. The effect of similarity was significant both for the Himba, $F(4, 56) = 22.80$, $p_{\mathrm{rep}} > .99$, $\eta_p^2 = .614$, and the USC students, $F(4, 20) = 16.72$, $p_{\mathrm{rep}} > .99$, $\eta_p^2 = .77$; more similar distractors were associated with higher error rates. The advantage for NAPs over MPs held over the full range of similarity. The magnitudes of the effects of property difference (NAP or MP) and physical similarity were equivalent for the two subject groups, leading to a nonsignificant three-way interaction of these variables, $F(4, 76) = 0.62$, $p_{\mathrm{rep}} < .50$, $\eta_p^2 = .032$. The interactions of property difference and group, $F(1, 19) = 3.12$, $p_{\mathrm{rep}} = .82$, $\eta_p^2 = .141$, and physical similarity and group, $F(3, 57) = 1.23$, $p_{\mathrm{rep}} = .63$, $\eta_p^2 = .061$, were also not significant. Part of the ranges of similarity values for the Himba and USC groups were nonoverlapping, so including the full ranges would have generated a spurious interaction (i.e., one group would have performed the task on stimuli with higher similarity values than the other group). To produce more equivalent similarity ranges in the test for the Similarity × Group interaction, we used only four of the five similarity levels for that test.

Caution is required in interpreting interactions (or noninteractions) when the dependent variable (e.g., error rate, in the present case) may not be on an interval scale, as interactions could be eliminated (or created) by a monotonic transformation. However, the overall performance levels of the two groups were essentially equivalent, so we have some confidence that any interactions (or noninteractions) in the data of one group would likely be replicated in the data of the other. We draw no conclusions from the fact that the Himba and USC groups had equivalent overall levels of performance, as the testing conditions and familiarity with similar tasks (e.g., experience playing video games) were so markedly different between the groups.

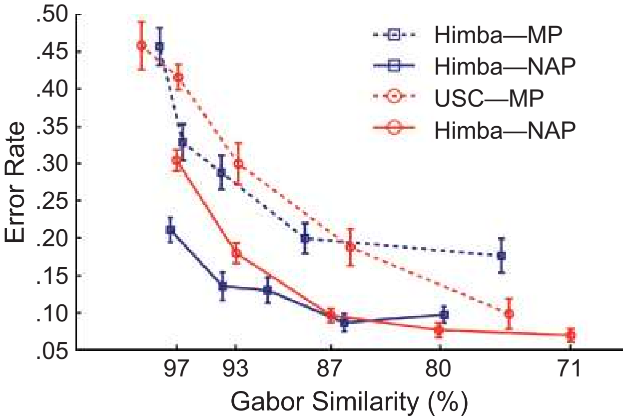

Fig. 5. Percentage error as a function of Gabor-jet similarity of the distractor and the matching stimulus, the nature of the difference between the distractor and the matching stimulus (nonaccidental property, NAP; metric property, MP), and subject group (Himba; University of Southern California, USC, students). The maximum possible Gabor-jet similarity is 100% (which would be the similarity of a stimulus with itself). For each subject group, trials were grouped into five similarity bins, with an equal number of trials in each bin. The error bars represent the standard errors of the means computed over the subjects in each group at each data point.

DISCUSSION AND CONCLUSION

This experiment offers, to our knowledge, the most rigorous assessment of the effects of exposure to modern artifacts on the representation of shape. The bottom line is that the Himba did not differ from individuals living in what is, arguably, the most artifactual of environments (Los Angeles). Specifically, the reduced exposure of the Himba to regular artifacts and the absence of simple shape terms in their language (a) did not result in reduced sensitivity (compared with USC students) to NAPs and (b) did not produce a difference in sensitivity to physical variation, as assessed by the Gabor-jet measure.

Why would there not be an effect of frequent exposure to simple, manufactured artifacts on sensitivity to NAPs relative to MPs? A genetic predisposition for coding shape by NAPs is one possibility. In addition, infants' attention to moving objects might provide support for a particular architecture tuned to NAPs. Whereas rotation in depth might lengthen or shorten the contour of an object or vary the degree of curvature of a contour with nonzero curvature, the fact that the contour had nonzero (or zero) curvature would not vary. Thus, during attention engaged by rotation of an object, the effective input to an infant's self-organizing network would tend to be tuned to NAPs of contours defining orientation and depth discontinuities (Biederman, 1987). It is this behavior, which can facilitate the development of the neural connectivity tuned to NAPs, that is likely universal and that could likewise render the representation of shape universal as well.

Acknowledgments—This research was supported by National Science Foundation Grants 04207994, 0531177, and 0617699 to I.B. and by European Union Project SEDSU (Stages in the Evolution and Development of Sign Use; 012-984) to J.D. We thank Greet Kayaert for constructing the stimuli and Marissa Nederhouser for running the subjects.

REFERENCES

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–147.

Biederman, I. (1995). Visual object recognition. In S.M. Kosslyn & D.N. Osherson (Eds.), *An invitation to cognitive science: Vol. 2. Visual cognition* (2nd ed., pp. 121–165). Cambridge, MA: MIT Press.

Biederman, I., & Bar, M. (1999). One-shot viewpoint invariance in matching novel objects. *Vision Research*, *39*, 2885–2899.

Fiser, J., Biederman, I., & Cooper, E.E. (1996). To what extent can matching algorithms based on direct outputs of spatial filters account for human object recognition? *Spatial Vision*, *10*, 237–271.

Kayaert, G., Biederman, I., & Vogels, R. (2003). Shape tuning in macaque inferior temporal cortex. *Journal of Neuroscience*, *23*, 3016–3027.

Kayaert, G., Biederman, I., & Vogels, R. (2005). Representation of regular and irregular shapes in macaque inferotemporal cortex. *Cerebral Cortex*, *15*, 1308–1321.

Lades, M., Vorbruggen, J.C., Buhmann, J., Lange, J., von der Malsburg, C., Wurtz, R.P., et al. (1993). Distortion invariant object recognition in the dynamic link architecture. *IEEE Transactions on Computers*, *42*, 300–311.

Lazareva, O.F., Wasserman, E.A., & Biederman, I. (2008). Pigeons and people are more sensitive to nonaccidental than to metric changes in visual objects. *Behavioral Processes*, *77*, 199–209.

Lescroart, M.D., Biederman, I., Yue, X., & Davidoff, J. (in press). A cross-cultural study of the representation of shape: Sensitivity to underlying generalized-cone dimensions. *Visual Cognition*.

Logothetis, N.K., Pauls, J., Bülthoff, H.H., & Poggio, T. (1994). View-dependent object recognition by monkeys. *Current Biology*, *4*, 401–414.

Roberson, D., Davidoff, J., & Shapiro, L. (2002). Squaring the circle: The cultural relativity of good shape. *Journal of Cognition and Culture*, *2*, 29–53.

Switkes, E., Mayer, M.J., & Sloan, J.A. (1978). Spatial frequency analysis of the visual environment: Anisotropy and the carpentered environment hypothesis. *Vision Research*, *18*, 1393–1399.

Tadmor, Y., & Tolhurst, D.J. (1994). Discrimination of changes in the second-order statistics of natural and synthetic images. *Vision Research*, *34*, 541–554.

Vogels, R., Biederman, I., Bar, M., & Lorincz, A. (2001). Inferior temporal neurons show greater sensitivity to nonaccidental than to metric shape differences. *Journal of Cognitive Neuroscience*, *13*, 444–453.

Willats, J. (1992). Seeing lumps, sticks and slabs in silhouettes. *Perception*, *21*, 481–496.

Yue, X., Subramaniam, S., & Biederman, I. (2007, November). *Predicting the psychophysical discriminability of faces and other complex stimuli based on a measure of image similarity.* Paper presented at the annual meeting of the Society for Neuroscience, San Diego, CA.

Yue, X., Tjan, B.S., & Biederman, I. (2006). What makes faces special? *Vision Research*, *46*, 3802–3811.

(Received 1/3/09; Revision accepted 4/7/09)