Journal of Experimental Psychology:
Human Perception and Performance
2005, Vol. 31, No. 4, 668–684

0096-1523/05/$12.00 DOI: 10.1037/0096-1523.31.4.668

The Structure of Three-Dimensional Object Representations in Human Vision: Evidence From Whole–Part Matching

E. Charles Leek
University of Wales, Bangor

Irene Reppa
University of Wales, Swansea

Martin Arguin
Université de Montréal and Institut Universitaire de Gériatrie de Montréal

This article examines how the human visual system represents the shapes of 3-dimensional (3D) objects. One long-standing hypothesis is that object shapes are represented in terms of volumetric component parts and their spatial configuration. This hypothesis is examined in 3 experiments using a whole–part matching paradigm in which participants match object parts to whole novel 3D object shapes. Experiments 1 and 2, consistent with volumetric image segmentation, show that whole–part matching is faster for volumetric component parts than for either open or closed nonvolumetric regions of edge contour. However, the results of Experiment 3 show that an equivalent advantage is found for bounded regions of edge contour that correspond to object surfaces. The results are interpreted in terms of a surface-based model of 3D shape representation, which proposes edge-bounded 2-dimensional polygons as basic primitives of surface shape.

Keywords: shape representation, vision, primitives, image segmentation, parts

One fundamental question for research on human vision concerns the way in which the visual system encodes and represents the shapes of three-dimensional (3D) objects (e.g., Feldman, 2003). There are several classes of theory about shape representation in human and machine vision including approaches based on structural decomposition (e.g., Biederman, 1985, 1987; Brooks, 1981; Marr, 1982; Marr & Nishihara, 1978), geometric models (e.g., Lowe, 1987), and multidimensional feature spaces (see Edelman, 1997, 1999). Of these classes of theory, the structural decomposition approach has been particularly influential in studies of human vision (e.g., Biederman, 1987; Marr, 1982; Marr & Nishihara, 1978) and in particular the hypothesis that mental representations of 3D objects consist of volumetric component parts and their spatial configuration. Recognition-by-Components (RBC; Biederman, 1985, 1987; Hummel, 1997, 2000, 2001; Hummel & Biederman, 1992; Hummel & Stankiewicz, 1996) is one much-cited example of this idea.[1] In RBC, shape representation is initially based on the detection of edge boundaries and vertices (i.e., intersections between two or more edges). Volumetric primitives are then used to approximate object shape based on nonaccidental properties of edges (e.g., collinearity, symmetry, parallelism, curvilinearity, and cotermination) and parsing of volumetric components at regions of concavity.[2] A structural description specifying the volumetric components and their spatial configuration is then matched to a similar representation held in long-term memory.

Despite the influence of hypotheses like RBC and of other volumetric models of shape representation (e.g., Brooks, 1981; Guzman, 1968; Marr & Nishihara, 1978), there is surprisingly little empirical evidence supporting a role for volumetric primitives in human vision. Much of the empirical debate has focused on issues concerning the viewpoint dependence of object representations (e.g., Arguin & Leek, 2003; Biederman & Gerhardstein, 1993; Edelman, 1999; Leek, 1998a, 1998b; Tarr & Bulthoff, 1998). In contrast, the nature of the shape primitives mediating high-level object representations remains poorly understood.

E. Charles Leek, Centre for Cognitive Neuroscience, School of Psychology, University of Wales, Bangor, United Kingdom; Irene Reppa, Department of Psychology, University of Wales, Swansea, United Kingdom; Martin Arguin, Départment de psychologie, Université de Montréal, Montréal, Ontario, Canada, and Centre de recherche, Institut Universitaire de Gériatrie de Montréal, Canada.

Some of the data reported in this article were presented at the 41st Annual Meeting of the Psychonomics Society, New Orleans, Louisiana, November 2000. This research was supported by Economic and Social Science Research Council United Kingdom Grant R000239512 awarded to E. Charles Leek and by an operating grant from the Natural Sciences and Engineering Research Council of Canada to Martin Arguin. We thank Daniel Saumier, Claire Hunter, and Maggie Lo for their help with the data collection. We thank Marie-Josèphe Tainturier, Mary Kaiser, and Donald Hoffman for their useful comments and suggestions.

Correspondence concerning this article should be addressed to E. Charles Leek, Centre for Cognitive Neuroscience, School of Psychology, University of Wales, Bangor LL57 2AS, United Kingdom. E-mail: e.c.leek@bangor.ac.uk

[1] The feasibility of RBC also continues to attract debate in computer vision (e.g., Barrow & Tenenbaum, 1993; Dickinson et al., 1997; Edelman, 1999).

[2] Because this article examines the use of volumetric component parts as a general hypothesis about shape representation (rather than any specific implementation of this hypothesis), we use the term *volumetric component* to refer to volumetric primitives rather than, for example, geons—the specific variant of volumetric primitive described in RBC.

Some supporting evidence for RBC is discussed by Biederman (1987). In one study it was found that accuracy for naming line drawings of familiar common objects decreases as the number of volumetric components in the stimulus image is reduced. Biederman (1987) has also shown that consistent with the predictions of RBC, the deletion of edge contour that disrupts the detection of concave discontinuities (required for parsing) and invariant features such as collinearity and curvilinearity (required for edge completion) results in performance costs in recognition, relative to the deletion of equivalent amounts of edge contour at noncritical regions (e.g., midsegments)—at least for short stimulus exposure durations (Biederman, 1987). Some further evidence for volumetric components has been presented by Biederman and Cooper (1991). In one experiment, participants named line drawings of objects across two blocks of trials separated by a 7-min interval. In the priming block, every other edge and vertex was deleted from each stimulus such that 50% of the contour from each volumetric component was preserved in the image. The second block contained either identical images, complementary images containing only the contour deleted from the stimuli in the priming block, or feature-deleted different exemplars sharing the same basic-level name. Priming effects on naming latencies were found in the second block for all conditions. The priming effects were also larger for identical and complementary image pairs than for same name–different exemplar pairs—suggesting that at least some component of the priming effects was purely visual in nature, rather than solely conceptual or phonological (Biederman & Cooper, 1991). Critically, the magnitude of priming was equivalent for identical and complementary feature-deleted image pairs. According to Biederman and Cooper (1991), this finding showed that visual priming was not mediated solely by representations of edges or vertices actually present in the stimuli but rather by representations specifying the volumetric component part structure. In Experiment 2, another condition was used involving the deletion of 50% of volumetric components from each stimulus. Prime-target pairs contained either identical (with the same components deleted) or complementary volumetric components. Here there was more priming between identical prime-target pairs than between complementary component deleted pairs. This finding was taken as evidence that visual priming is mediated by volumetric components present in the priming stimulus; that is, by shape representations that make explicit the volumetric component parts of the object. However, although the evidence from these studies appears to support volumetric models of shape representation, it is not unequivocal (e.g., Edelman, 1999).

First, Edelman (1999) has argued that the pattern of results in Biederman and Cooper (1991, Experiment 1) can also be accounted for by edge completion (e.g., Grossberg & Mingolla, 1985). That is, the identical and complementary conditions might have shown equivalent priming effects because of an edge completion mechanism acting on the complementary feature-deleted primes. Edge completion is also likely to be disrupted more by the deletion of specific types of edge features, such as vertices, which provide clues to collinearity and curvilinearity of deleted edge segments (Grossberg & Mingolla, 1985).

Second, another issue is whether time costs associated with the deletion of volumetric components (Biederman, 1987; Biederman & Cooper, 1991, Experiment 2) necessarily reflect an effect of volumetric parts per se or the deletion of some covariant shape property. One such covariate is surface shape, which might also be an important factor for object recognition (e.g., Barrow & Tenenbaum, 1981; Lee & Park, 2002; Leek & Arguin, 2000; Leek, Reppa, & Arguin, 2002, 2003; Marr, 1982; Marr & Nishihara, 1978; Nakayama, He, & Shimojo, 1995; Nakayama & Shimojo, 1992; Pentland, 1989). In the feature-deletion paradigm, effects of volumetric component deletion are confounded with the deletion of object surfaces that form constituent parts of those components. It could be the case that surfaces, or at least bounded edge regions that potentially correspond to object surfaces (e.g., Camps, Huang, & Kanungo, 1998; Lee & Park, 2002; Leek, Reppa, & Arguin, 2002, 2003), contribute to the apparent volumetric part advantage found in previous studies.

The aim of this study is to examine the basis of volumetric part effects in shape recognition and to elucidate the structure of the shape primitives that mediate 3D object shape representation. The experiments are based on a whole–part matching paradigm (e.g., Ankrum & Palmer, 1991; Palmer, 1977) in which participants match subsets of edge features from 3D objects to whole object shapes. The rationale of the studies is that whole–part matching should be faster (and more accurate) when part stimuli contain image features or configurations of image features that match the primitives that are encoded in mental shape representations of the objects during perception. Thus, performance in matching part stimuli containing different types of image features may be used to make inferences about the content and structure of mental representations of object shapes. For example, if volumetric component parts have a special status in shape representations, then whole–part matching should be more efficient for part stimuli containing volumetric components than for part stimuli containing other nonvolumetric configurations of edge features (e.g., bounded regions corresponding to object surfaces or other nonvolumetric regions of object shape). In Experiments 1 and 2, we contrasted performance in matching whole object shapes to subsets of volumetric components versus equivalent amounts of open or closed nonvolumetric regions of edge contour. In Experiment 3, we contrasted matching of whole objects to subsets of volumetric components versus bounded regions of edge contour corresponding to object surfaces. The empirical question is whether performance in whole–part matching depends on how edge and vertex information is grouped in the part stimulus displays. Observed effects of feature grouping may be taken to reflect the functional significance or special status of that part type in mental representations of 3D object shapes.

Experiment 1

The primary aim of Experiment 1 was to contrast performance in matching part stimuli containing subsets of edge contour that are either grouped into volumetric or nonvolumetric regions in whole object shapes. Models of shape representation assuming the existence of volumetric shape primitives, such as RBC, predict an advantage for matching volumetric parts over those consisting of nonvolumetric configurations of edge contour because, by hypothesis, the perception of whole objects involves image segmentation into constituent volumetric parts.

Method

Participants. Eighteen undergraduate students from the University of Wales, Bangor, United Kingdom, participated in the experiment for course credit. All participants had normal or corrected-to-normal visual acuity.

Apparatus. The experiment was run on a Macintosh G3 computer with a 17-in. RGB monitor using PsycLab software (Gum, 1995).

Stimuli. The stimuli are shown in Figure 1. They consisted of opaque perspective line drawings of 12 3D novel objects. Each object was scaled to fit within a 6 × 6-cm frame that subtended 6.86° of visual angle from a viewing distance of 50 cm. The object set contained both geometrically regular and irregular shaped components. Each object consisted of two volumetric components—a larger principal component and a smaller component. For each object, two types of comparison (part) stimuli were created: volumetric components and open contour. The volumetric component part stimuli ($n = 24$) each consisted of a single volumetric component. There were 12 large volumetric component stimuli containing a mean of 70.21% ($SD = 3.85\%$) of total edge contour and 12 small volumetric component stimuli containing a mean of 40.19% ($SD = 5.02\%$) of total edge contour (see Table 1).

The open contour stimuli ($n = 24$) were made by deleting a proportion of noncontiguous edge contour from both volumetric components of each stimulus. Deleted contour included both vertices and midsegments of contour (Biederman, 1987). There were two subsets of open contour stimuli: large ($n = 12$) and small ($n = 12$). The large open contour set contained a mean of 57.66% ($SD = 7.13\%$) of total edge contour (where total edge contour is the total length of contour in the whole object stimulus). The small open contour set contained a mean of 42.51% ($SD = 7.09\%$) of total edge contour. Large volumetric component stimuli contained a significantly higher proportion of total edge contour than open contour stimuli, $t(22) = 2.14$, $p < .04$. The proportion of total edge contour in the small volumetric component and small open contour part stimuli was not significantly different, $t(22) = 0.51$, *ns*.

Previous studies have also shown that edge vertices carry important information about object shape and its spatial configuration (e.g., Biederman, 1987; Enns & Rensink, 1991; Guzman, 1968). Vertices were defined as the intersection or junction of two or more edges. The total number of Y, T, and L vertices per stimulus was equated between the large open contour and volumetric component conditions and between the small open contour and volumetric component conditions (see Table 1). Whole object displays were also enlarged to 150% of their original size to prevent contour overlap with the part stimuli. This measure served to prevent participants from completing the task using a direct template matching strategy.

Design. The experiment was based on a 2 (part type: volumetric component vs. open contour) × 2 (part size: large vs. small) × 2 (match vs. mismatch) repeated measures design. Each trial consisted of the presentation of a whole novel object, followed by a part stimulus (see Figure 2). For match trials, the part stimuli consisted of either open contour or a volumetric component from the object that preceded it. For mismatch trials, the part stimuli consisted of either open contour or a volumetric component from another visually similar object.

The experiment consisted of 96 experimental (48 match and 48 mismatch) and 12 (6 match and 6 mismatch) practice trials. For both match and mismatch, there were 12 trials in each of the four conditions (large vs. small open contour, large vs. small volumetric component). In total, each whole object stimulus was presented eight times, and each part stimulus twice (once in a match trial and once in a mismatch trial) during the experiment. Trials were presented in two blocks of 48 trials. Half of the participants made match responses with their dominant hand and mismatch responses with their nondominant hand. For the other half, these assignments were reversed. Trial order was randomized, within blocks, for each participant.

Procedure. The procedure is shown schematically in Figure 2. Participants were seated approximately 50 cm from the monitor. Each trial began with the central presentation of a visual prompt *Ready.* The prompt remained on the screen until the subject initiated the trial sequence by pressing on the space bar. Following a blank interstimulus interval of 750 ms, one of the whole object stimuli appeared in the center of the screen for 1,200 ms. After a further blank interstimulus interval of 750 ms, a part

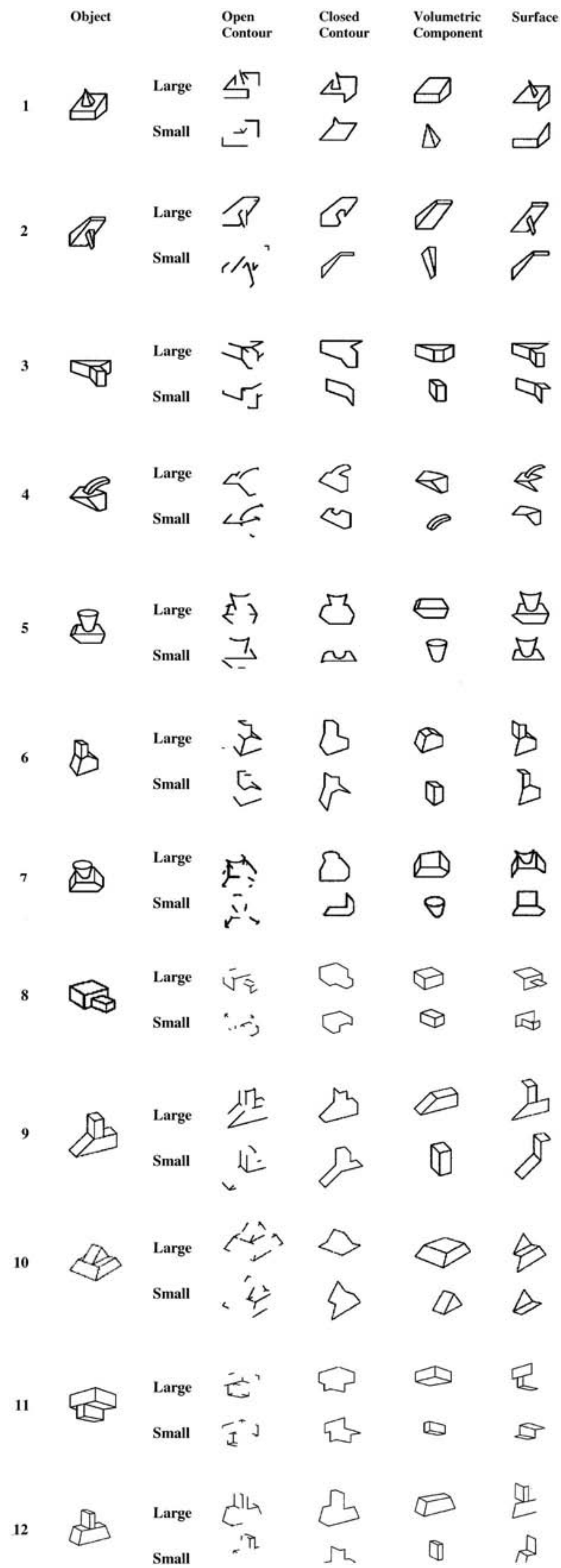

Figure 1. The 12 novel object shapes used in Experiments 1, 2, and 3.

Table 1
Properties of the Open Contour and Volumetric Part Stimuli Used in Experiment 1

Part type	Edge contour (cm)		% total edge contour		*N* edge vertices	
	M	*SD*	*M*	*SD*	*M*	*SD*
Open contour						
Large	15.61	3.70	57.66	7.13	7.50	0.80
Small	11.42	2.70	42.51	7.09	5.83	1.34
Volumetric						
Large	19.03	4.10	70.21	3.85	7.50	0.80
Small	10.86	2.60	40.19	5.02	5.83	1.34

stimulus was presented in the center of the screen until the participant made a response. The task was to decide as quickly and as accurately as possible whether the part stimulus came from the whole object that preceded it. There was a response deadline of 3,000 ms. When a response was incorrect or timed out, the participants received feedback in the form of a short error tone and a visual prompt (*Incorrect*). The participants made their responses by pressing one of two keys labeled *match* or *mismatch* on a standard keyboard. The experiment lasted approximately 25 min.

Results

Analyses of reaction times (RTs). RT analyses were conducted on mean RTs per object for correct responses across conditions. RTs were trimmed to within $\pm$ 2 *SD*s from the mean per condition.

A 2 (match vs. mismatch) $\times$ 2 (part type: volumetric components vs. open contour) $\times$ 2 (part size: large vs. small) repeated measures analysis of variance (ANOVA) showed a significant main effect of match–mismatch, $F(1, 11) = 11.25$, $p < .0006$. There was also a marginally significant two-way interaction of Part Type $\times$ Part Size, $F(1, 11) = 4.07$, $p < .06$. Figures 3A and 3B show mean RTs for the small and large part stimulus conditions by part type and match–mismatch.

Separate 2 (match vs. mismatch) $\times$ 2 (part type) repeated measures ANOVAs were also carried out on mean RTs for the small and large part conditions. For the small parts, there was a significant main effect of match–mismatch, $F(1, 11) = 7.97$, $p < .01$, and a significant interaction between the two factors, $F(1, 11) = 5.32$, $p < .04$. Planned comparisons between critical conditions showed that there was a significant difference between small volumetric components and small open contour parts for match trials, $t(11) = 2.66$, $p < .02$, but not for mismatch trials, $t(11) = 0.68$, *ns*. For the large part conditions, there were no significant main effects or interactions.

Analyses of error rates. The mean percentage error rate per condition was 10.98% of trials ($SD = 3.81\%$): Mean error rates per condition are shown in Figure 4. An analysis, across conditions, using the Friedman nonparametric test for multiple-dependent groups by ranks was not significant, $\chi^2(7, N = 12) = 11.25$, *ns*. Error rates were also examined using pairwise Wilcoxon signed-ranks test. The only significant difference was between the large open contour and large volumetric component conditions in the mismatch trials, $Z(12) = 2.29$, $p < .02$. There was a marginally significant positive correlation between RTs and error rates ($r^2 = .034$), $F(1, 94) = 3.31$, $p = .07$. There was no indication of a speed–accuracy trade-off.

Discussion

The main findings of Experiment 1 were as follows: First, an advantage was found for matching part stimuli consisting of volumetric components over nonvolumetric configurations of open edge contour. Second, this difference was only found for part stimuli containing a relatively small proportion of total edge contour. No differences were observed between conditions for the large part stimuli. Third, RTs across conditions were equivalent in the mismatch trials for both large and small parts.

These results suggest that the way in which edge information is grouped in part stimulus displays can affect performance in whole–part matching—at least under conditions where part stimuli contain relatively small amounts of total edge contour. The advantage in matching partial edge contour that is grouped into volumetric components is consistent with the predictions of volumetric component-based hypotheses of 3D shape representation like RBC. On these accounts, this advantage may be taken to reflect the greater efficiency of matching image features that correspond to volumetric primitives in the shape representations encoded for the whole object shapes. The greater difficulty in matching the open edge contour stimuli may be presumed to reflect an increase in processing time required to map nonvolumetric groups of edge contour onto volumetric components in whole object representations.

The effect of part size is likely to reflect the fact that increasing the proportion of edge contour in the part displays also increases the probability of recovery of critical shape information using contour completion (e.g., Biederman, 1987; Biederman & Cooper, 1991). When there are insufficient image features, contour completion processes may be unable to resolve ambiguities in the part stimulus displays. It is relevant to note that the apparent advantage for matching volumetric components in the small part displays cannot be explained by differences between part–match conditions in terms of the amount of visible edge contour or vertices present in the stimuli. In the small part displays, the volumetric component and open contour part stimuli were matched on both of these factors.

Arguably, one could also account for the volumetric component advantage in terms of a potential difference in the ease of discrimination of volumetric parts and open contours between match and mismatch trials. That is, volumetric components might show an advantage because in the match and mismatch trials, these stimuli are more visually distinct from each other than those of the open contour stimuli. However, if this were the case, then an advantage for volumetric components should also have been found in the mismatch trials. Thus, the absence of differences between conditions in the mismatch trials suggests that the advantage for volumetric components cannot be due solely to greater ease of discrimination.

There is also another possible interpretation of these results. The advantage for matching volumetric parts may not actually reflect a genuine effect of volumetric structure per se but rather an advantage for encoding and matching perceptually closed forms (e.g., Ankrum & Palmer, 1991; Tversky, Geisler, & Perry, 2003). The volumetric part stimuli consisted of geometrically well-formed closed shapes, whereas the open contour part stimuli did not. For the same reason, it is also possible that the difference in performance between open contour and volumetric components might solely reflect additional time taken for edge completion with the open contour forms (e.g., Edelman, 1999). If this were the case,

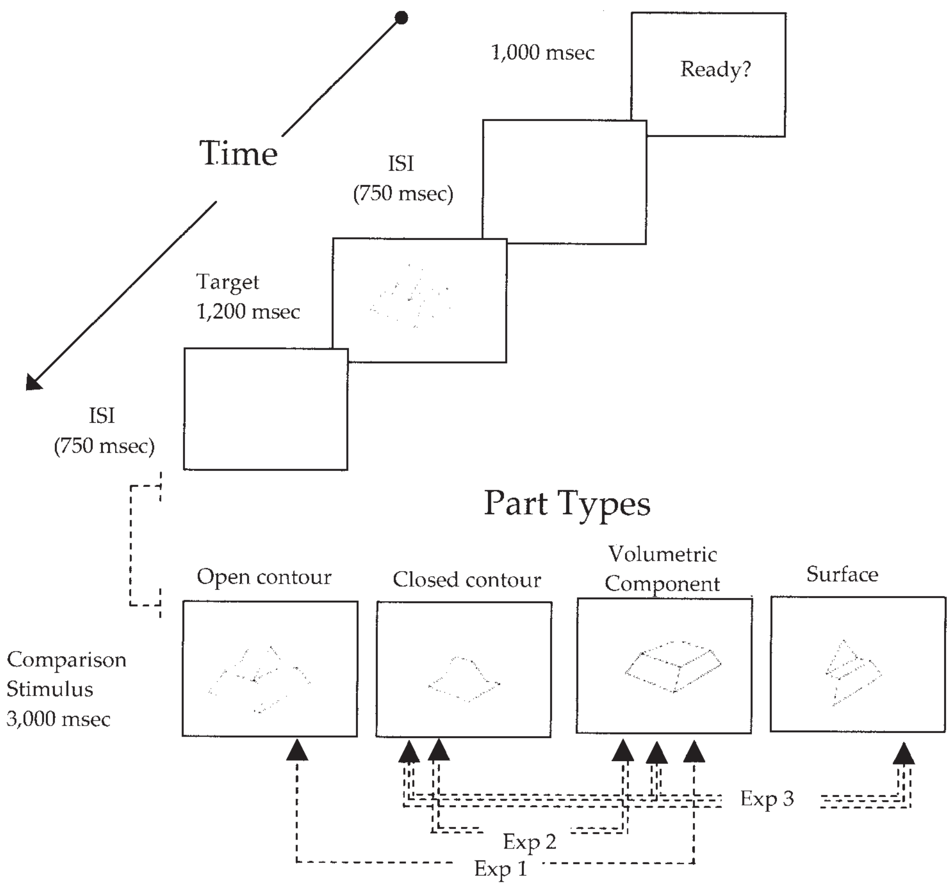

Figure 2. The procedure for the whole–part matching paradigm used in Experiments (Exp) 1, 2, and 3. The bottom of the figure illustrates the part comparison stimulus types used in each experiment. ISI = interstimulus interval.

then the advantage in matching volumetric parts should disappear if they are contrasted with geometrically closed but nonvolumetric groups of edge contour.

In addition, although the open contour and volumetric components were matched in terms of the number of L, Y, and T vertices shown, only open contour parts contained endpoints at the terminations of deleted edges. Arguably, this factor may also have contributed to the apparent advantage for volumetric components by increasing stimulus complexity in the open contour parts.[3] These issues were examined in Experiment 2.

Experiment 2

The aim of Experiment 2 was to determine whether the apparent advantage for matching volumetric parts found in Experiment 1 can be accounted for in terms of an effect of part stimulus closure. Experiment 2 also allowed us to examine whether the presence of endpoints at the terminations of deleted vertices in the open contour stimuli contributed to the apparent advantage for matching volumetric parts. These issues were examined by contrasting performance in whole–part matching with volumetric components versus nonvolumetric closed regions of partial edge contour.

Method

Participants. Nineteen undergraduate students from the University of Wales, Bangor, United Kingdom, participated in the experiment for course credit. All participants had normal or corrected-to-normal visual acuity.

Apparatus and stimuli. The apparatus and stimuli were the same as those used in Experiment 1, except for the following modifications. In addition to the volumetric components, a new set of part stimuli were created for each of the 12 novel objects. These consisted of closed contour stimuli made by deleting regions of edge contour under the constraint that the resulting stimulus consisted of a closed form containing edge contour from both volumetric components of the whole object (see Figure 2). There were two subsets of closed contour stimuli: large ($n = 12$) and small ($n = 12$; see Table 2).

The large closed contour set contained a mean of 60.99% ($SD = 6.68\%$) of total edge contour. The small closed contour set contained a mean of 47.58% ($SD = 5.93\%$) of total edge contour. Both the large closed contour and volumetric component and small closed contour and volumetric component part stimuli were matched in terms of visible edge contour, $t(22) = 1.82$, *ns*, and $t(22) = 1.69$, *ns*, respectively. It was not possible to match between conditions in terms of the number of edge vertices in the part displays. The number of edge vertices in the closed contour stimuli was

[3] We thank an anonymous reviewer for this observation.

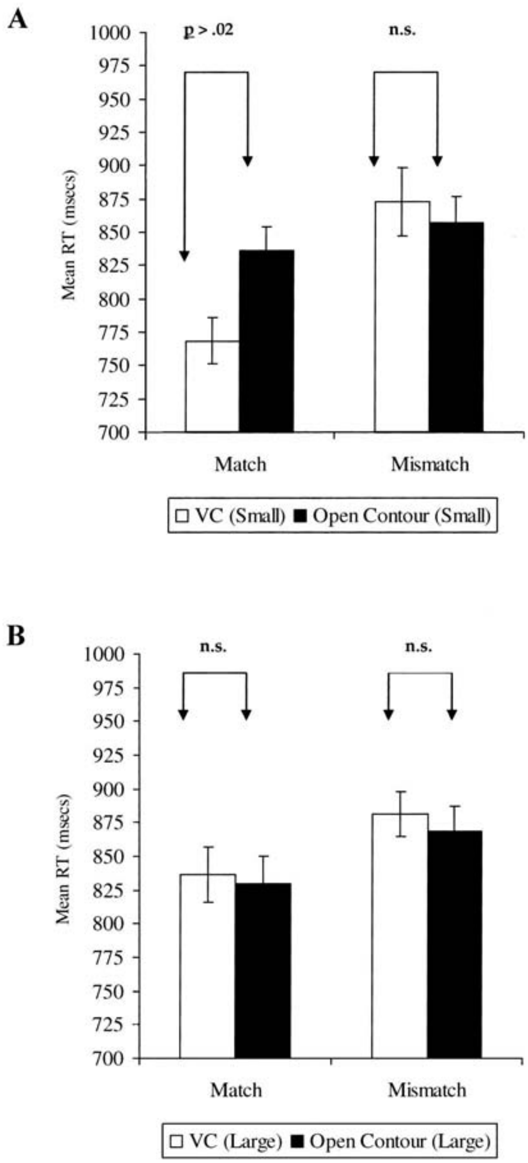

Figure 3. Mean reaction times (RTs) per condition for match and mismatch trials for small (A) and large (B) part types in Experiment 1. Error bars show standard error of the mean. VC = volumetric component.

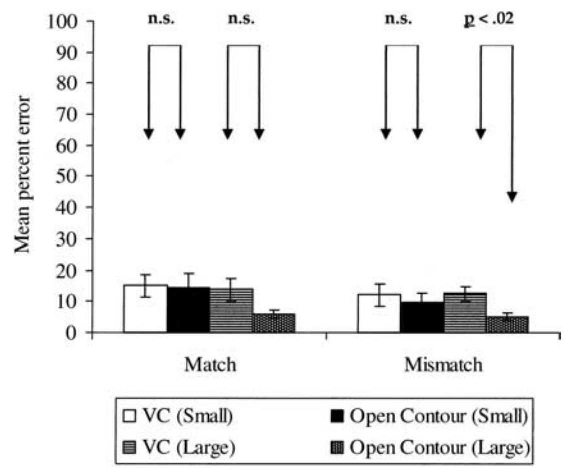

Figure 4. Mean percentage error rates per condition in Experiment 1. Error bars show standard error of the mean. VC = volumetric component.

significantly greater than in the volumetric components for both large, $t(22) = 7.00$, $p < .0001$, and small, $t(22) = 3.26$, $p < .003$, stimulus sets. We examine the possible effects of this difference in the analyses. Neither the closed contour nor volumetric components contained endpoints at the terminations of deleted vertices.

Design and procedure. The design and procedure were identical to Experiment 1, except for the factor of part type in which volumetric parts were contrasted with closed contour rather than open contour parts.

Results

Analyses of RTs. RT analyses were conducted on mean RTs per object for correct responses across conditions. RTs were trimmed to within ± 2 *SD*s from the mean per condition.

A 2 (match vs. mismatch) × 2 (part type: volumetric components vs. closed contour) × 2 (part size: large vs. small) repeated measures ANOVA showed a significant main effect of part type, $F(1, 11) = 165.76$, $p < .00001$. There were also significant two-way interactions of Match–Mismatch × Part Type, $F(1, 11) = 16.61$, $p < .001$, and of Part Size × Part Type, $F(1, 11) = 5.72$, $p < .03$. Figures 5A and 5B show mean RTs in the match and mismatch conditions for large and small volumetric components and large and small closed contour.

The interactions were explored by conducting separate 2 (match vs. mismatch) × 2 (part type) repeated measures ANOVAs on mean RTs for the small and large part conditions. For the small part conditions, there was a significant main effect of part type, $F(1, 11) = 57.12$, $p < .00001$, and a significant interaction of Match–Mismatch × Part Type, $F(1, 11) = 24.16$, $p < .0004$, indicating a larger advantage for volumetric components over closed contour stimuli in the match than mismatch trials. Planned comparisons between conditions showed that the difference in RTs between volumetric components and closed contours was significant in both match, $t(11) = 9.53$, $p < .00001$, and mismatch, $t(11) = 3.54$, $p < .004$, trials. For the large part conditions, there was a significant main effect of part type, $F(1, 11) = 9.62$, $p < .01$. There was also a significant interaction between the two factors, $F(1, 11) = 7.4$, $p < .01$. Simple effects analyses showed that the advantage for volumetric components over closed contours was significant in the match trials, $t(11) = 3.34$, $p < .006$, but not in the mismatch trials, $t(11) = 0.62$, *ns*.

Table 2
Properties of the Closed Contour and Volumetric Part Stimuli Used in Experiment 2

	Edge contour (cm)		% total edge contour		*N* edge vertices	
Part type	*M*	*SD*	*M*	*SD*	*M*	*SD*
Closed contour						
Large	16.36	2.99	60.99	6.68	9.83	0.83
Small	12.88	3.22	47.58	5.93	8.17	2.08
Volumetric						
Large	19.03	4.10	70.21	3.85	7.50	0.80
Small	10.86	2.60	40.19	5.02	5.83	1.34

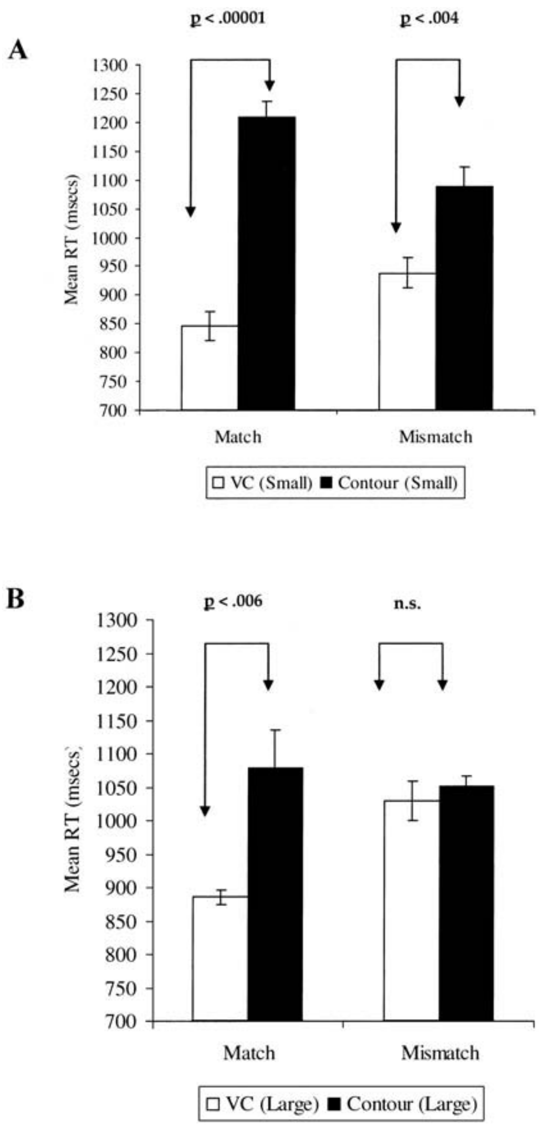

Figure 5. Mean reaction times (RTs) per condition for match and mismatch trials for small (A) and large (B) part types in Experiment 2. Error bars show standard error of the mean. VC = volumetric component.

Further analyses. As noted earlier, closed contour and volumetric component part stimuli were matched in terms of mean visible edge contour, but closed contour parts did contain significantly more edge vertices than the volumetric components. This difference may potentially account for the apparent advantage for matching volumetric components. Indeed, there were significant correlations between mean RTs and the number of edge vertices per stimulus for both large ($r^2 = .34$), $F(1, 22) = 12.86$, $p < .001$, and small ($r^2 = .43$), $F(1, 22) = 16.78$, $p < .0004$, part stimuli. The contribution of vertices to the difference in RTs was examined using an analysis of covariance (ANCOVA), with mean RTs per part stimulus (match trials only) as the dependent variable, part condition as the categorical predictor, and the number of edge vertices as the covariate. For the large part displays, neither the main effects of part type, $F(1, 21) = 0.67$, *ns*, or vertices, $F(1, 22) = 1.68$, *ns*, were significant. However, post hoc analyses using the Fisher least significant difference (LSD) test showed that the difference in mean RTs between closed contour and volumetric component parts was significant when the factor of vertices was partialed out ($p < .002$). For small part displays, there was a significant effect of part type, $F(1, 21) = 41.20$, $p < .00001$, and a marginally significant effect of edge vertices, $F(1, 21) = 3.93$, $p < .06$. Post hoc analyses using the Fisher LSD test showed a significant difference between closed contour and volumetric components partialing out the covariate ($p < .0001$).

Analyses of error rates. The mean percentage error rate per condition was 17.0% of trials ($SD = 17.5\%$). Mean error rates per condition are shown in Figure 6.

An analysis, across conditions, using the Friedman nonparametric test for multiple-dependent groups by ranks was significant, $\chi^2(7, N = 12) = 26.99$, $p < .0003$. This was further explored using pairwise Wilcoxon signed-ranks test. In the match trials, there were significant differences in the error rates between closed contour and volumetric component parts in the small part displays, $Z(12) = 2.93$, $p < .003$, but not in the large part displays, $Z(12) = 1.56$, *ns*. In the mismatch trials, there were no significant differences in error rates between conditions. There was a significant positive correlation between RTs and error rates ($r^2 = .35$), $F(1, 94) = 52.50$, $p < .0001$. There was no indication of a speed–accuracy trade-off.

Discussion

The main findings from Experiment 2 were as follows: First, for both the small and large part stimuli, whole–part matching was faster for volumetric components than for nonvolumetric configurations of closed edge contour. Second, further analyses showed that the advantage for matching volumetric components was present even when differences between conditions, in terms of the number of edge vertices shown, were factored out. Third, for the large part stimuli, the difference between conditions was found

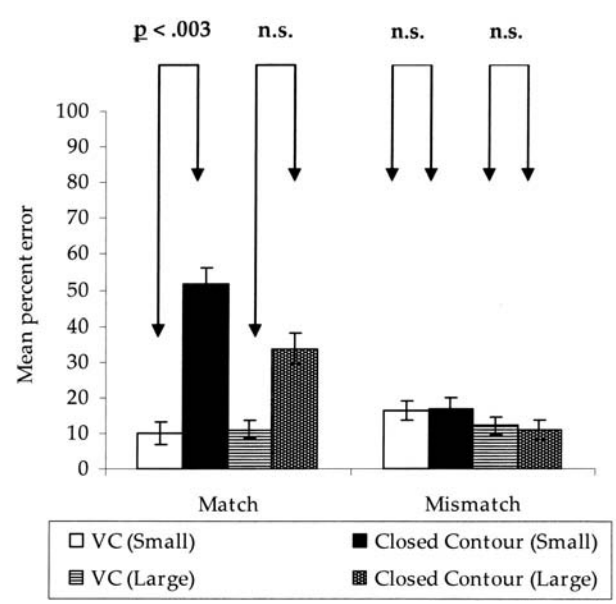

Figure 6. Mean percentage error rates per condition in Experiment 2. Error bars show standard error of the mean. VC = volumetric component.

only in the match trials, showing that the volumetric part advantage is not due to differences in the relative ease of discrimination between match and mismatch trials.

These findings appear to rule out the possibility that the volumetric part advantage found previously in Experiment 1 reflects an effect of shape closure per se and therefore is due to additional time costs associated with edge completion for open contour displays. In Experiment 2, the small closed contour and volumetric parts were matched in terms of visible edge contour, and both part types consisted of geometrically closed forms.

The results also show that the apparent advantage for volumetric components could not be accounted for by differences, between conditions, in terms of either total edge contour presented or the number of vertices or endpoints shown in the part displays. In terms of edge contour, the stimuli were matched across conditions (as in Experiment 1), but closed contour parts did contain more vertices than the volumetric components. The analyses suggest that this difference could not account for the volumetric part advantage.

However, although the results of Experiment 2 show that shape closure as well as the amount of edge contour vertices and endpoints shown in the part displays does not underlie the advantage for volumetric components, there remain several other factors that could also potentially contribute to this effect other than volumetric structure. One factor concerns changes in vertex type that result from the deletion of edge contours in the part displays. Although we have so far considered the influence of the number of vertices present in each display type, the deletion of edge contour in the part displays also often results in changes in the types of vertices—or edge intersections—that occur at the terminations (i.e., endpoints) of edge segments (see Figure 1). One change is that T junctions in whole objects frequently become L junctions at the same locations in the contour-deleted part stimuli. Such changes are less frequent in the volumetric part displays. This change in vertex type could contribute to the apparent volumetric part advantage found in Experiments 1 and 2. We examined this possibility in Experiment 3.

Another factor, noted at the beginning of this article, concerns a potential confound between the deletion of volumetric components and the deletion of other kinds of shape information that is covariant with volumetric part structure. One potential covariate is surface shape. Object surfaces can carry important information that facilitates perception (e.g., Barrow & Tennenbaum, 1981; Camps et al., 1998; Lee & Park, 2002; Leek & Arguin, 2000; Leek, Reppa, & Arguin, 2002, 2003; Marr, 1982; Marr & Nishihara, 1978; Nakayama et al., 1995; Pentland, 1989). In the previous two experiments, the contour and volumetric part stimuli also differed in terms of the information they contain about shape surfaces and their spatial configuration. The volumetric part stimuli, unlike the open and closed contours, preserved information about bounded regions of edge contour that corresponds to object surfaces. Thus, the volumetric part advantage might reflect an effect of surface, rather than volumetric, part structure. We also examined this possibility in Experiment 3.

Experiment 3

In Experiment 3, we contrasted performance in whole–part matching for part stimuli consisting of either (a) regions of closed edge contour, (b) volumetric components, or (c) bounded regions of edge contour that correspond to object surfaces. To control for effects of surface structure, we matched the surface and the volumetric component part stimuli for the number of visible object surfaces that they contain. For example, for a volumetric component containing three visible surfaces, a surface part stimulus was created that contained three spatially adjacent surfaces from the same object but in a nonvolumetric configuration. If the pattern of results found in Experiments 1 and 2 is a genuine effect of volumetric structure, volumetric parts should show an advantage over both closed contour and surface part stimuli—neither of which contain edges that are grouped into volumetric components.

Method

Participants. Twenty undergraduate students from the University of Wales, Bangor, United Kingdom, participated in the experiment for course credit. All participants had normal or corrected-to-normal visual acuity.

Apparatus and stimuli. The apparatus and stimuli were the same as those used in Experiment 2, except for the following modifications. An additional set of part stimuli were created for each of the 12 novel objects. These were matched to the closed contour and volumetric component stimuli for size (6×6 cm) and subtended visual angle (6.86° from a viewing distance of 50 cm). The new stimulus set consisted of two-dimensional (2D) surface patches (defined as polygonal bounded regions of edge contour; see Figure 2). Two sets of surface stimuli were created: large and small. The large surface set contained a mean of 72.81% (SD = 9.41%) of total edge contour. The small surface set contained a mean of 53.29% (SD = 8.07%) of total edge contour.

The surface part sets were constructed using two constraints: (a) For each object, the large and small surface parts were matched, exactly, to the respective volumetric components in terms of the number of visible surfaces shown (see below), and (b) the surfaces were contiguous (i.e., spatially adjacent) as in the volumetric parts. Because of the constraints on the construction of the surface stimuli, it was not possible to always match the surface parts to the volumetric component and closed contour stimuli in terms of the proportion of total edge contour or the number of edge vertices (see Table 3).

Proportion of total visible edge contour. For the small part stimuli, closed contour and volumetric components did not significantly differ in terms of the mean proportion of visible edge contour shown, $t(22) = 1.69$, *ns*, and neither did closed contour and surface parts, $t(22) = 1.09$, *ns*. Volumetric components contained significantly less edge contour than surface parts (M = 40.19% and SD = 5.02% vs. M = 53.29% and SD = 8.07%, respectively), $t(22) = 3.25$, $p < .004$.

For the large part stimuli, closed contour and volumetric components did not significantly differ in terms of the mean proportion of visible edge contour, $t(22) = 1.82$, *ns*, and neither did volumetric components and surface parts, $t(22) = 0.25$, *ns*. Closed contour parts contained a significantly smaller proportion of edge contour than surface parts (M = 60.99% and SD = 6.68% vs. M = 72.81% and SD = 9.41%, respectively), $t(22) = 2.68$, $p < .01$.

Number of edge vertices. In the small part stimulus set, closed contours contained significantly more edge vertices than the volumetric components ($M = 8.17$ and $SD = 2.08$ vs. $M = 5.83$ and $SD = 1.34$, respectively), $t(22) = 3.26$, $p < .003$, but did not significantly differ from surface parts, $t(22) = 0.65$, *ns*. Surface parts contained significantly more vertices than volumetric components ($M = 7.67$ and $SD = 1.61$ vs. $M = 5.83$ and $SD = 1.34$, respectively), $t(22) = 3.02$, $p < .001$.

In the large part stimulus set, volumetric components contained significantly fewer vertices than both the closed contour parts ($M = 7.50$ and $SD = 0.80$ vs. $M = 9.83$ and $SD = 0.83$, respectively), $t(22) = 7.00$, $p < .0001$, and surface parts ($M = 7.50$ and $SD = 0.80$ vs. $M = 10.17$ and $SD = 1.40$, respectively), $t(22) = 5.72$, $p < .0001$. Closed contours and surface parts were not significantly different, $t(22) = 0.70$, *ns*.

Proportion of vertex-type deletion. As noted earlier, a further relevant factor that may contribute to the efficiency of performance in whole–part matching concerns potential effects related to changes in the type of edge

Table 3
Properties of the Closed Contour, Volumetric Part, and Surface Part Stimuli Used in Experiment 3

Part type	Edge contour (cm)		% total edge contour		*N* edge vertices		% vertex change		*N* surfaces	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Closed contour										
Large	16.36	2.99	60.99	6.68	9.83	0.83	72.83	7.28		
Small	12.88	3.22	47.58	5.93	8.17	2.08	77.86	12.47		
Volumetric										
Large	19.03	4.10	70.21	3.85	7.50	0.80	16.87	18.90	3.25	0.45
Small	10.86	2.60	40.19	5.02	5.83	1.34	22.38	21.27	2.50	0.52
Surface										
Large	19.31	2.36	72.81	9.41	10.17	1.40	44.53	11.22	3.25	0.45
Small	14.19	2.58	53.29	8.07	7.67	1.61	47.21	13.31	2.50	0.52

vertices between whole and part stimulus displays that occur through the deletion of edge contour. The proportion of vertices shown in the part stimuli that have changed type as a result of edge deletion varies between conditions (see Table 3). Notably, closed contours have a higher proportion of vertex-type changes than volumetric components in both large, $t(11) = 10.54$, $p < .0001$, and small, $t(11) = 8.66$, $p < .0001$, part displays as well as a higher proportion than surfaces: large, $t(11) = 6.97$, $p < .0001$, and small, $t(11) = 6.10$, $p < .0007$. Surface parts have a higher proportion of vertex-type changes than volumetric components in both large, $t(11) = 4.07$, $p < .001$, and small, $t(11) = 3.00$, $p < .01$, displays. These vertex-type changes may be expected to increase matching difficulty. We examined the influence of this factor in the analyses (see below).

Number of surfaces. Both small and large volumetric component and surface part stimuli were matched exactly for the number of visible surfaces shown (see Table 3). Potential issues arising from the differences between conditions for proportion of total edge contour and the number of edge vertices shown are examined in the analyses (see below).

Design and procedure. Experiment 3 was based on a 2 (match vs. mismatch) × 3 (part type: closed contour vs. volumetric component vs. surface) × 2 (large vs. small part) design. The experiment consisted of 144 trials (72 match and 72 mismatch) comprising 12 trials per condition.

In total, each whole object stimulus was presented 12 times, and each part stimulus twice (once in a match trial and once in a mismatch trial). Trials were presented in two blocks of 72 trials each containing the same number of trials per condition. Participants were also shown 12 practice trials (6 match and 6 mismatch). The experiment lasted about 40 min. In all other respects, the design and procedure were identical to Experiments 1 and 2.

Results

Analyses of RTs. RT analyses were conducted on mean RTs per object for correct responses across conditions. RTs were trimmed to ± 2 *SD*s from the mean per condition.[4]

A 2 (match vs. mismatch) × 3 (part type: closed contours vs. volumetric components vs. surfaces) × 2 (part size: large vs. small) repeated measures ANOVA showed significant main effects of match–mismatch, $F(1, 10) = 22.17$, $p < .0008$, and of part type, $F(2, 20) = 9.10$, $p < .001$. There was also a significant two-way interaction of Match–Mismatch × Part Type, $F(2, 20) = 8.94$, $p < .001$. There was no main effect of part size, $F(1, 10) = 0.09$, *ns*, and no interactions of part size with any other factor.

In the absence of any main effect or interactions involving part size, RTs were collapsed across this factor. Figure 7 shows mean RTs in the match and mismatch trials as a function of part condition collapsed across part size.

The interaction between match–mismatch and part type was further examined using one-way ANOVAs on match and mismatch trials. For match trials, there was a significant effect of part type, $F(2, 20) = 16.17$, $p < .00006$. Planned comparisons showed significant differences in mean RTs between closed contours and volumetric components, $t(10) = 4.35$, $p < .001$, and between closed contours and surfaces, $t(10) = 4.50$, $p < .001$. There was no significant difference in mean RTs between volumetric components and surfaces, $t(10) = .82$, $p = .42$. There were no significant effects in the analyses of data from mismatch trials.

Further analyses. In order to further assess the factors underlying the apparent advantage for matching volumetric components and bounded surfaces regions in whole–part matching, we also examined the relative contribution of differences between conditions in terms of the amount of edge contour and vertices as well as the effects of changes in vertex type arising through edge deletion in the part displays.

Visible edge contour. Although the surface and volumetric component part stimuli were matched for the number of visible surfaces, as previously noted, it was not possible to precisely match part stimuli across conditions on all factors including the amount of edge contour shown. The effect of this factor in accounting for the pattern of differences in RTs between conditions was examined using an ANCOVA, with mean RTs per part stimulus (match trials only) as the dependent variable, part condition as the categorical predictor, and total visible edge contour (cm) as the covariate. There was a significant effect of part type, $F(2, 62) = 12.89$, $p < .0001$, but not of edge contour, $F(1, 62) = 0.40$, *ns*. Post hoc analyses using the Fisher LSD test showed significant differences between volumetric components and closed contour ($p < .0001$), between surfaces and closed contour ($p < .0001$), but not between surfaces and volumetric components ($p = .57$), when differences in visible edge contour were factored out. There was also no overall correlation between mean RTs and edge contour ($r^2 = .02$), $F(1, 64) = 1.60$, *ns*.

Number of edge vertices. As in Experiment 2, we also examined the possible contribution of differences between conditions in

[4] Data from one stimulus (Object 4) were discarded from the analyses of Experiment 3 because the global error rate exceeded 2 *SD*s from the mean error rate across objects.

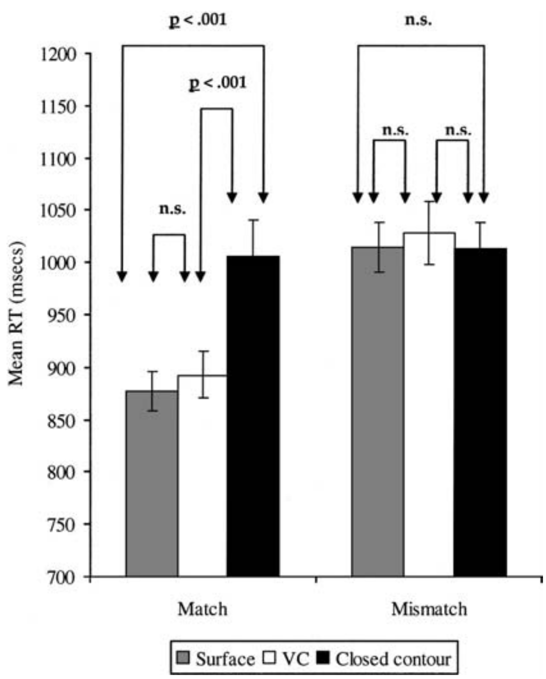

Figure 7. Mean reaction times (RTs) per condition for match and mismatch trials collapsed across part size in Experiment 3. Error bars show standard error of the mean. VC = volumetric component.

terms of the number of edge vertices shown in the part displays (i.e., the number of edge vertices contained in the part stimuli relative to the whole object). An ANCOVA was conducted, with mean RTs per part stimulus (match trials only) as the dependent variable, part condition as the categorical predictor, and the number of edge vertices as the covariate. The results showed a significant effect of part type, $F(2, 62) = 13.02$, $p < .00001$, but no effect of edge vertices, $F(1, 62) = 0.01$, *ns*. Post hoc analyses using the Fisher LSD test showed significant differences between volumetric components and closed contour ($p < .0001$), between surfaces and closed contour ($p < .0001$), but not between surfaces and volumetric components ($p = .57$). There was also no overall correlation between mean RTs and vertices ($r^2 = .01$), $F(1, 64) = 0.95$, *ns*.

Proportion of vertex-type deletion. Part stimuli also differed in the proportion of vertex-type changes resulting from contour deletion (i.e., changes in the type of vertex shown at a particular junction of edges between the part and whole object displays). These changes may also be expected to disrupt whole–part matching. Unlike the proportion of edge contour and the number of vertices in the part stimuli, there was a significant correlation between the proportion of vertex-type changes and mean RTs ($r^2 = .14$), $F(1, 64) = 11.04$, $p < .001$, suggesting that higher proportions of vertex-type changes lead to slower responses. It is relevant then that closed contour parts contained a significantly higher proportion of vertex-type changes than either volumetric component or surface parts. This difference, rather than part type per se, could underlie the pattern of RT data that was found. In order to examine this possibility, we conducted an ANCOVA, with mean RT per part stimulus (match trials only) as the dependent variable, part condition as the categorical predictor, and the proportion of vertex-type changes as the covariate. The results showed a significant effect of part type, $F(2, 62) = 7.12$, $p < .001$, but no effect of vertex-type change, $F(1, 62) = 0.02$, *ns*. Post hoc analyses using the Fisher LSD test showed significant differences between volumetric components and closed contour ($p < .0001$), between surfaces and closed contour ($p < .0001$), but not between surfaces and volumetric components ($p = .57$) when variation in the proportion of vertex-type change between conditions is factored out.

This analysis shows that the proportion of vertex-type changes in the feature-deleted part displays cannot account for the differences in performance between part–match conditions. That is, the relative advantage in matching surface and volumetric components over closed contour parts does not derive from the higher proportion of vertex-type changes in the closed contour displays.

Analyses of error rates. The mean percentage error rate per condition was 23.60% of trials ($SD = 3.73\%$). Mean error rates per condition are shown in Figure 8.

A Friedman nonparametric test for multiple-dependent groups by ranks, across conditions, was not significant, $\chi^2(5, N = 11) = 7.16$, *ns*. Additional contrasts using pairwise Wilcoxon signed-ranks test showed a significant difference in error rates for match trials between closed contours and volumetric components, $Z(12) = 2.03$, $p < .04$, but no other significant differences between conditions. There was a significant positive correlation between error rates and RTs ($r^2 = .08$), $F(1, 64) = 6.01$, $p < .01$. There was no indication of a speed–accuracy trade-off.

Discussion

The main findings of Experiment 3 were as follows: First, both large and small surface and volumetric component parts were matched more efficiently than nonvolumetric closed contour parts. Second, this advantage was found only in match trials—there were no significant differences between conditions in the mismatch trials. Third, this pattern of results was still found when differences between conditions in terms of visible edge contour, the number of edge vertices, and the proportion of vertex-type changes shown were partialed out.

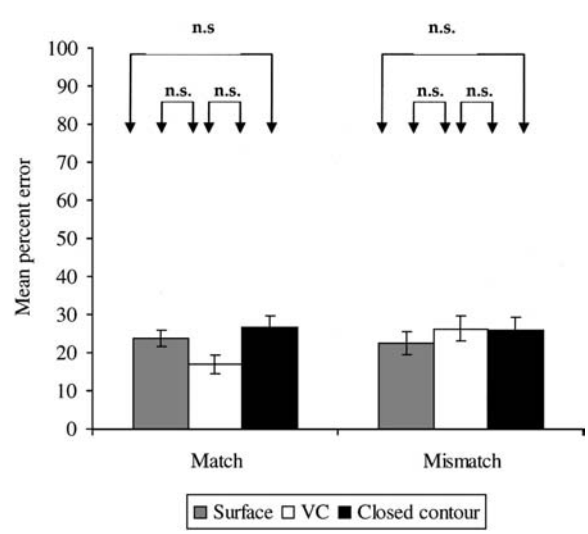

Figure 8. Mean percentage error rates per condition in Experiment 3. Error bars show standard error of the mean. VC = volumetric component.

General Discussion

The results suggest that the configuration of edge contour that is presented in the part stimulus displays can affect performance in whole–part matching. Participants were faster at matching edge contour that corresponds to volumetric components or surfaces than to either open or closed nonvolumetric regions of edge contour that cannot be directly mapped onto volumetric or surface parts. The study was based on the rationale that differences between conditions in the efficiency of whole–part matching may be taken to reflect the functional status of specific configurations of edge information in the shape representations mediating task performance. At a general level, the results support the view that mental representations of object shape are part based (e.g., Biederman, 1987; Hoffman & Richards, 1984; Marr & Nishihara, 1978; Palmer, 1977). They also provide new constraints on hypotheses about the nature of the primitive shape elements in mental representations of objects.

We first consider a number of methodological issues. We then discuss the theoretical implications of the results for hypotheses about the structure of shape representations of 3D solid objects.

Methodological Issues

Opponents of part-based models of shape representation might argue that task requirements of the whole–part matching paradigm somehow bias observers toward shape decomposition (Cave & Kosslyn, 1993; Edelman, 1999). This possibility could potentially undermine the generality of the current results and those of other studies showing part-based effects in whole–part matching tasks (e.g., Ankrum & Palmer, 1991; Palmer, 1977). The objection is unsatisfactory for a number of reasons. First, it does not explain the patterns of differences between conditions found across experiments. That is, even if the task did bias perception toward image segmentation, this alone leaves unanswered the theoretically relevant question of why performance was dependent on the specific configuration of edge information in the part displays. Second, the objection only carries force to the extent that it can also account for the whole range of data from other studies reporting evidence for part-based representations from other paradigms (e.g., Biederman, 1987; Biederman & Cooper, 1991; Bower & Glass, 1976; Hoffman & Richards, 1984; Hoffman & Singh, 1997; Leek & Arguin, 2000; Leek, Reppa, & Arguin, 2003; Leek, Reppa, & Tipper, 2003; Reed, 1974; Reed & Johnsen, 1975; Reppa & Leek, 2003; Siddiqi, Tresness, & Kimia, 1996; Stankiewicz, 2002; Vecera, Behrmann, & Filapek, 2001; Vecera, Behrmann, & McGoldrick, 2000; Xu & Singh, 2002).

Another possible challenge is that the whole–part paradigm encourages an artificial strategy of template matching that could preclude the encoding of more abstract (possibly nonpart-based) shape representations. A template-matching strategy might bias performance against the open contour conditions because of increased demands on short-term memory for the encoding and sequential matching of discrete edge segments rather than groups of bounded edge regions. There are also several counterarguments against this possibility. First, in the current paradigm, participants could not directly match the part- and whole-object displays because they differed in size by 150%. This was done precisely to eliminate image overlap and to reduce the likelihood of direct template matching. Second, as noted above, the argument does not fully account for the pattern of differences that were found between conditions. For example, it does not explain why template matching should be more efficient for bounded regions that correspond to surfaces or volumetric components. To account for this finding, one needs to consider the functional significance of the regions bounded by the edges in each condition. This, in turn, requires making reference to the internal structure of shape representations—precisely the goal of the current study. It is also relevant to note that this pattern of differences between conditions could not be accounted for in terms of the amount of edge contour or vertices shown in the part displays. This finding also argues against an explanation solely, for example, in terms of short-term memory load during template matching.

Finally, one might argue that the results could be explained in terms of differences, between conditions, in the ease of discriminability of match and mismatch trials, rather than because of the efficiency of matching particular types of primitives to shape representations. For example, it might have been easier to discriminate match from mismatch trials in the volumetric component and surface conditions than in the closed contour condition—perhaps on the basis of their global shape outline or some other property. This possibility also seems unlikely because a similar pattern of differences would be expected between conditions in the match and mismatch trials. With the exception of one condition in Experiment 2, this was not the case. Therefore, differences between conditions in the ease of discriminability of match and mismatch trials cannot account for the patterns of results found across experiments.

None of these possibilities provides a satisfactory explanation for the results reported here. Rather, it seems that an adequate account must appeal to theoretical hypotheses about the structural primitives of 3D shape representations. We now discuss the implications of the data for theories of object shape representation.

Implications for Theories of 3D Shape Representation

In Experiments 1 and 2, an apparent advantage was found for matching displays consisting of edge contour that is grouped into volumetric component parts over both open and closed nonvolumetric configurations of edge contour. This finding is consistent with volumetric models of 3D shape representation, such as RBC (Biederman, 1985, 1987). According to this hypothesis, whole–part matching should be more efficient when part stimuli consist of edge contour that is grouped into volumetric components over nonvolumetric bounded regions. This is because the part stimuli should be more easily matched to volumetric primitives that are, by hypothesis, encoded in the shape representations mediating task performance. Arguably, the open and closed contour stimuli cannot be directly mapped onto such primitives without additional processing related to edge completion, image segmentation, and the resolution of spurious relations among vertices and edge contours (e.g., Biederman, 1987; Dickinson et al., 1997).[5] However, in Experiment 3, we also found an equivalent advantage over closed contour parts for stimuli consisting of edges that are grouped into

[5] The results also showed that these differences between conditions are sometimes only found when part stimuli contain a relatively small proportion of total edge contour from the whole object—an effect that was found in Experiment 1 and which has also been reported in some other studies using feature-deletion paradigms (e.g., Biederman, 1987). This presumably reflects the fact that feature deletion only disrupts performance differentially when there is insufficient visible edge contour to prevent the recovery of critical shape information.

nonvolumetric bounded regions. The analyses showed that the advantage for these nonvolumetric bounded regions over the closed contour parts could not be accounted for by differences between these conditions in terms of the complexity of edge features (i.e., mean edge contour, number of vertices, and percentage vertex change).[6] A relevant question, then, concerns the functional status of these nonvolumetric bounded regions. Our hypothesis is that they correspond to 2D edge-bounded primitives that are used to approximate surface shapes in mental representations of 3D objects. This proposal is consistent with a nonvolumetric surface-based model of shape representation, which we outline below (Leek & Arguin, 2000; Leek, Reppa, & Arguin, 2002, 2003).

A Surface-Based Model of Shape Representation

The hypothesis is schematically outlined in Figure 9. We assume that the basic elements mediating high-level shape representation consist of 2D edge-bounded polygons that are used to approximate the shapes of surfaces in 3D objects (Leek & Arguin, 2000; Leek, Reppa, & Arguin, 2002, 2003). Image scene segmentation is based on the detection of discontinuities from multiple cues, including luminance (light), chrominance (color), texture, and retinal disparity (i.e., local depth) gradients (as well as motion in nonstatic environments). Surfaces belonging to individual objects are approximated by bounded 2D regions defined by these discontinuities (e.g., Barrow & Tennenbaum, 1981; Beck, 1972; Binford, 1981; Grossberg & Swaminathan, 2004; Julesz, 1981; Leek & Arguin, 2000; Leek, Reppa, & Arguin, 2003; Palmer & Rock, 1994; Potmesil, 1983; Sajda & Finkel, 1995; Witkin, 1981). The model distinguishes between perceptual representations computed from the visual stimulus and stored representations of familiar object shapes. At both levels of representation, the spatial configuration of object surfaces is described by a surface configuration map (e.g., Lee & Park, 2002). Other attributes of individual surfaces (such as shape, color, and texture where available) are encoded separately in surface feature layers linked to each surface node in the configuration map (shown only for the stored representation in Figure 9). During perception, a 2D surface configuration map is computed that encodes patterns of spatial adjacency among visible surface polygons in the stimulus. Thus, object representations at this level are necessarily viewpoint specific. In contrast, the surface configuration maps for stored representations contain a description of the patterns of spatial adjacency for all known surfaces in the object. In this sense, the stored configuration maps may be considered to be 3D model representations.[7,8] Metric properties of surface shape and orientation are specified for pairs of spatially adjacent surfaces in a surface-centered 3D coordinate system whose axes are defined in relation to a kernel surface. Recognition depends on 2D geometric transformations that are used to determine correspondences between surfaces in the perceptual representation and surfaces encoded at the level of the stored 3D surface model. These transformations permit access to stored shape representations despite the inherent perspective deformation and viewpoint specificity of 3D object recognition. Initial matching of a subset of kernel surfaces in the stored 3D surface model is then used to further constrain geometric transformation and interpretation of the 2D surface configuration map computed from the stimulus.

A key feature of this hypothesis—described here in outline form only—is that it does not contain volumetric primitives, such as geons or generalized cylinders (e.g., Biederman, 1985, 1987; Hummel, 2001; Marr, 1982). Instead, the basic units of shape description consist of 2D edge-bounded primitives that are used to approximate object surface shapes. Another key feature is that object spatial configuration is represented in terms of local pairwise relations among surface primitives and does not depend on the computation of global shape attributes such as principal axes of elongation and symmetry (e.g., Marr, 1982; Marr & Nishihara, 1978).[9] As such, the model differs from RBC and other volumetric models of 3D shape representation. It also provides an account for the pattern of results found in the current study. As the model does not contain volumetric components but surface-based primitives, the surface and volumetric part stimuli would be expected to show equivalent performance in the current study because these two

[6] It has been suggested by one reviewer that the difference between the closed contour and surface part stimuli might also be characterized by some measure of simplicity related, for example, to their respective Fourier spectra. We do not address this possibility here.

[7] Although the variant of the surface representations hypothesis outlined here proposes the use of 3D model representations at the level of the stored surface configuration map, we do not rule out other possibilities. One such possibility is the use of viewer-centered stored representations in which sets of possible object views are encoded in a finite set of 2D surface aspects (e.g., Beymer & Poggio, 1996; Chakravarty & Freeman, 1982; Koenderink & van Doorn, 1979; Ullman & Basri, 1991). This solution to the problem of perspective deformation in object constancy and representational complexity is not incompatible with the hypothesis about surface primitives in shape representation outlined in this article. It is possible to construct an aspect graph in which each aspect specifies sets of topological invariants of constituent surface primitives of the kind proposed here. The hypothesis does not preclude the further organization of stored 3D object models in terms of structured viewpoint-specific surface aspect graphs.

[8] It should not be assumed that the stored 3D surface configuration map predicts viewpoint-invariant recognition. On the contrary, the processes mediating the transformation and matching of perceptual and stored representations in the model are predicted to be highly viewpoint dependent, consistent with empirical evidence on object constancy (e.g., Tarr & Bulthoff, 1998). One reason for this is that changes in viewpoint induce perspective deformations of individual surfaces that must be compensated for by geometric transformation. Another reason is that variation in the frequency at which particular views of objects are seen will result in some surfaces in the stored representations being activated more frequently than others. Such frequency effects would be expected to affect the times taken to access stored representations from particular viewpoints on subsequent presentations (i.e., more frequently seen surfaces will be activated more quickly).

[9] It has been suggested that one advantage of volumetric models of shape representation over the surface-based model outlined here is the ease with which observers can recognize stick figures (Marr & Nishihara, 1978). It should be noted, however, that the class of objects that can be identified using stick figure representations is very limited. Consider, for example, a door, a chair, a carpet, a mug, a table, a TV, a keyboard, and a book, to name a few. In fact, the efficiency of identifying stick figures seems largely restricted to basic-level distinctions among animals. In any event, stick figure recognition, in itself, does not provide evidence in favor of volumetric parts. We agree that it does demonstrate that the visual system is able to compute relatively abstract descriptions of some shape properties, such as the elongation of certain shape components. In the model outlined above, such abstract shape properties could also be computed. For example, surface elongation could be derived from the 3D coordinate frame of reference in which metric surface features of shape and orientation are encoded.

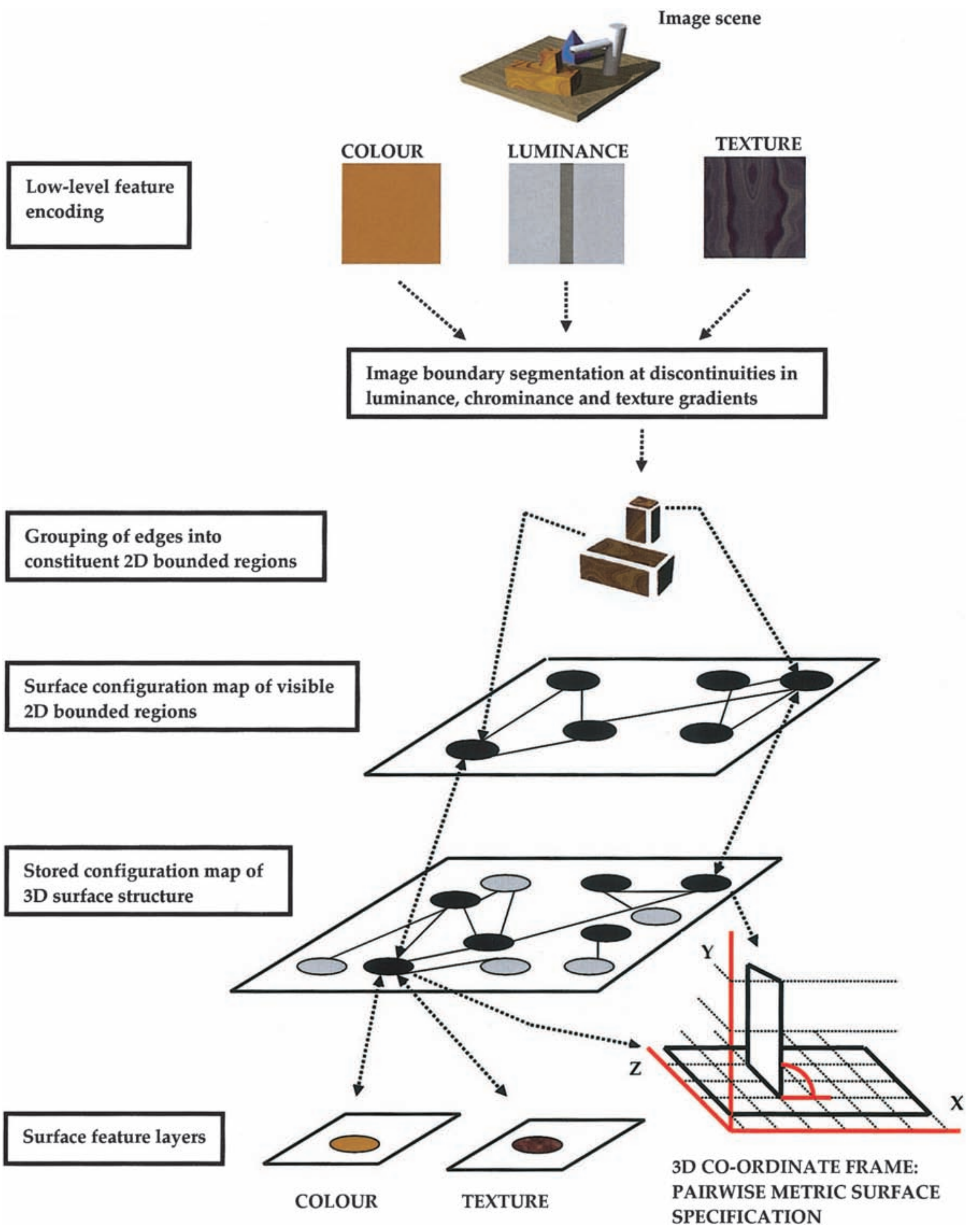

Figure 9. A schematic outline of the surface-based representations hypothesis. Two-dimensional (2D) edge-bounded polygons are used to approximate the shapes of object surfaces. The spatial configuration of visible surfaces in the stimulus (black circles) is encoded in a perceptual surface configuration map. Individual surface attributes (e.g., color and texture) are encoded separately in feature layers linked to each surface (shown only for stored shape representations). A three-dimensional (3D) surface configuration map is used to encode the configuration of all known surfaces (gray circles represent known but currently occluded surfaces). Metric surface attributes of shape and orientation are specified for each pair of surfaces using a surface-centered 3D coordinate reference frame.

conditions were matched in terms of the information they contain about visible object surfaces in the part stimuli.[10,11] It is also relevant to note that the advantage for surface parts found in Experiment 3 cannot be explained in terms of a general advantage for matching edge-bounded or closed regions or taken as evidence that any bounded region constitutes a primitive of shape representation. It is not the case that performance is equally efficient for all types of closed parts (e.g., Tversky et al., 2003). The differences in performance between the closed contour and both surface and volumetric component conditions show that it is not closure or boundedness that matters but rather to what features of object shape the bounded regions correspond. The pattern of results supports the hypothesis that edge-bounded regions derive functional significance from their correspondence to object surfaces.[12]

The model also motivates an alternative explanation for the results from some previous feature-deletion studies that have been used to support volumetric models of representation (e.g., Biederman, 1987; Biederman & Cooper, 1991). In those studies, the deletion of volumetric components (or the deletion of vertices that are assumed to support the recovery of volumetric parts) is typically confounded with the disruption of object surface structure. That is, the apparent advantage for the recovery of volumetric components in these feature-deletion studies may not reflect a genuine effect of volumetric structure but rather result from the preservation of information related to the efficient recovery of edge-bounded regions that can be mapped onto object surfaces.

Surfaces in Shape Perception

Other studies, in both perceptual psychophysics and machine vision, support a role for surfaces in the interpretation of visual scenes and the representation of object shape (e.g., Barrow & Tenenbaum, 1981; Binford, 1981; Cunningham, Shipley, & Kellman, 1998; Grossberg, 2000; Grossberg & Swaminathan, 2004; He & Nakayama, 1992, 1995; Hummel, 2000; Kellman & Shipley, 1991; Lee & Park, 2002; Leek & Arguin, 2000; Leek, Reppa, & Arguin, 2002, 2003; Lehky & Sejnowski, 1988, 1990; Liu, Collin, & Chaudhuri, 2000; Marr, 1982; Marr & Nishihara, 1978; Nakayama et al., 1995; Nakayama & Shimojo, 1992; Nicholson & Humphrey, 2001; Nishihara, 1981; Pentland, 1989; Potmesil, 1983; Ramachandra, 1988; Sajda & Finkel, 1995; Sanocki, Bowyer, Heath, & Sarkar, 1998; Witkin, 1981). For example, in some studies, it has been shown that observers are better at recognizing gray scale images of objects (which contain information about shading and surface luminosity gradients) than line drawings (Brodie, Wallace, & Sharrat, 1991; Price & Humphreys, 1989; Sanocki et al., 1998)—although line drawings, like those used in the current study, can be sufficient to support rapid stimulus identification (Biederman & Ju, 1988).

There is also supporting evidence from studies of visual attention. He and Nakayama (1992) manipulated binocular disparity in a visual search task. Participants searched stereogram displays for a target among distracters that could either appear on a plane in front of or behind another planar surface. Although the shape of the target element was always identical, the target, when perceived in the foreground, was interpreted as a white *L* shape. However, the same display could also be perceived as the protruding corners of an occluded white square under certain manipulations of binocular disparity. The results showed that search latencies for target detection were drastically increased when the target displays seemingly appeared behind an occluding surface. According to He and Nakayama (1992), this time cost arose because of surface completion. The *L*-shape targets were no longer perceived as an *L* shape per se but as the protruding corner of an occluded white square. These findings, like those reported in the current study, suggest that the grouping of objects and image features is influenced by surface structure computed from disparity gradients (see also He & Nakayama, 1995; Nakayama et al., 1995; Nakayama & Shimojo, 1992).

Additional evidence also comes from the neuropsychological literature on acquired deficits in visual object recognition. The object recognition ability of some individuals with acquired visual agnosia—particularly those stemming from damage to stored shape representations or access to those representations—can be shown to be sensitive to surface properties of objects (e.g., Chainay & Humphreys, 2001; Humphrey, Goodale, Jakobson, & Servos, 1994; Servos, Goodale, & Humphrey, 1993). For example, in object identification, patients often make fewer errors with real objects and photographs than with line drawings (e.g., Chainay & Humphreys, 2001; Davidoff & Wilson, 1985; Farah, 1990). This finding could be explained within the context of a surface-based model of 3D shape representation: Forms of visual input such as real objects and photographs contain more information about the surface structure of objects (e.g., their planar orientation and texture) than line drawings (in addition to other forms of informa-

[10] We do not claim, however, that volumetric structure has no status or role in shape representation. Although the surface model does not explicitly contain volumetric primitives, it potentially shows an implicit, emergent, volumetric structure. This follows from the way in which surface connectivity is encoded among groups of spatially adjacent surfaces at the level of the surface connectivity map (Leek, Reppa, & Arguin, 2003). This connectivity is expressed in terms of a weight matrix that specifies the frequency at which pairs of surfaces occur in spatial proximity to each other for any given 2D aspect projection. As a result, groups of spatially adjacent surfaces tend to develop relatively high intercorrelations through Hebbian association, whereas nonadjacent surfaces develop low intercorrelations. These regions of high intercorrelation lead to an emergent volumetric structure for groups of spatially adjacent surfaces, despite the fact that the representation has no volumetric components. This emergent volumetric structure also provides a potentially important link between surface shape primitives and conceptual–linguistic distinctions among volumetric object parts such as the legs of a table and the handle of a cup (e.g., Jackendoff, 1990).

[11] The hypothesis outlined here also potentially provides solutions to some limitations of volumetric approaches (e.g., Edelman, 1999; Kurbat, 1994). For example, the kinds of 2D bounded regions that are proposed in the model to approximate object surfaces are likely to be very useful in the representation of geometrically irregular shapes and for solid objects that consist of a single putative volumetric component, such as a shoe (which may be represented as a configuration of textured 2D surface patches). Also, as noted earlier, the use of surface primitives provides a means of encoding and binding representations of object shapes with other important visual features such as texture and color—surface attributes that are likely to play an important role in image segmentation, object recognition, and other types of visuomotor function such as reaching and grasping.

[12] It is relevant to note, also, that an equivalent advantage was found for the volumetric and surface parts even though several of the surface part stimuli were perceptually bistable (see Figure 2) and therefore arguably more ambiguous. This apparent lack of an effect of bistability may indicate that the perception of object shape does not depend on the recovery of visible surface depth as in, for example, the 2.5D sketch of Marr (1982) because bistability presumably arises from contrasting 3D interpretations of 2D images.

tion such as relative object size and color). It is interesting to note that Chainay and Humphreys (2001) have shown that the object recognition performance of integrative agnosic patient HJA was better with stimuli containing surface luminance gradients than with line drawings. The presence of surface information, in this case, facilitated the segmentation and recognition of the stimulus displays.

Volumetric Parts in Shape Perception

Although we have argued that the current data support a role for edge-bounded 2D surface primitives in shape representation, they cannot be taken as evidence against all volumetric models. On the one hand, it could be argued that both surface and volumetric models can account equally well for the pattern of results found in Experiment 3. For example, matching in the surface part condition could potentially be mediated, not by the approximation of 2D edge-bounded regions to object surfaces but by the recovery of volumetric parts (such as geons in RBC) from the surface displays. The current data do not allow us to definitively rule out this possibility. However, if this were the case, one might also expect that any additional computation required to approximate volumetric components from the surface parts would result in an additional time cost—that is, one would predict an advantage for volumetric components over surface parts. This is not supported by the pattern of results that showed an equivalent advantage for surfaces and volumes over nonvolumetric closed parts.

However, some volumetric models do propose intermediate levels of shape representation that could also provide a basis for explaining the current findings (e.g., Hummel, 2001; Marr, 1982; Marr & Nishihara, 1978). For example, Marr (1982) proposed a structural decomposition model with a level of surface-based representation—the 2.5D sketch, intermediate between low-level edge detection and the computation of a volumetric 3D model. In this model, the equivalent advantage for matching volumetric components and surface parts could be accounted for by assuming that whole–part matching is performed on the basis of information derived from the 2.5D sketch rather than from image segmentation at the level of the 3D (volumetric) model representation.

More recently, Hummel (2001) has outlined a revised version of RBC (JIM3) that also contains a level of surface representation—a significant departure from the earlier RBC model (Biederman, 1987). On this later account, surface structure derives from the grouping of edges bounding each surface, and this level of representation, in turn, outputs activation both to a map of surface attributes and a geon-based object shape model. The current findings could also be interpreted within the context of this formulation of RBC.[13]

Summary

In this article, we have shown that performance in 3D whole–part object matching depends on the way in which edge information is grouped in part displays. The results showed an advantage for matching part displays consisting of volumetric components over nonvolumetric configurations of open and closed part stimuli. An equivalent advantage was found for part displays containing 2D bounded regions that correspond to object surfaces. We have argued that this pattern of results challenges volumetric models of shape representation that do not contain intermediate surface-based levels of representation. We have also outlined an alternative surface-based model of 3D shape representation whose basic structural elements consist of edge-bounded 2D polygons that are used to approximate object surface shape. This hypothesis provides a new framework for future studies of 3D shape representation in human vision.

[13] It is relevant to note, however, that the status of surfaces in JIM3 is not entirely clear. According to Hummel (2001), “the shift to surface properties in the current version of the model is more a matter of convenience . . . than a strong theoretical claim” (p. 496).

References

Ankrum, C., & Palmer, J. (1991). Memory for objects and parts. *Perception & Psychophysics, 50,* 141–156.
Arguin, M., & Leek, E. C. (2003). Orientation invariance in visual object priming depends on prime-target asynchrony. *Perception & Psychophysics, 65,* 469–477.
Barrrow, H. G., & Tenenbaum, J. M. (1981). Interpreting line drawings as three-dimensional surfaces. *Artificial Intelligence, 17,* 75–116.
Barrow, H. G., & Tenenbaum, J. M. (1993). Retrospective on “Interpreting line drawings as three-dimensional surfaces.” *Artificial Intelligence, 59,* 71–80.
Beck, J. (1972). Similarity grouping and peripheral discriminability under uncertainty. *American Journal of Psychology, 85,* 1–19.
Beymer, D., & Poggio, T. (1996, June 28). Image representations for visual learning. *Science, 272,* 1905–1909.
Biederman, I. (1985). Human image understanding: Recent research and a theory. *Computer Vision, Graphics, and Image Processing, 32,* 29–73.
Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94,* 115–147.
Biederman, I., & Cooper, E. (1991). Priming contour-deleted images: Evidence for intermediate representations in visual object recognition. *Cognitive Psychology, 23,* 393–419.
Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 1162–1182.
Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology, 20,* 38–64.
Binford, T. O. (1981). Inferring surfaces from images. *Artificial Intelligence, 17,* 205–244.
Bower, G. H., & Glass, A. L. (1976). Structural units and the reintegrative power of picture fragments. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 456–466.
Brodie, E. E., Wallace, A. M., & Sharrat, B. (1991). Effects of surface characteristics and style of production on naming and verification of pictorial stimuli. *American Journal of Psychology, 104,* 517–545.
Brooks, R. (1981). Symbolic reasoning among 3-dimensional and 2-dimensional images. *Artificial Intelligence, 17,* 285–349.
Camps, O. I., Huang, C.-Y., & Kanungo, T. (1998). Hierarchical organization of appearance-based parts and relations for object recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 685–691). Santa Barbara, CA: Institute of Electrical and Electronics Engineers.
Cave, C. B., & Kosslyn, S. M. (1993). The role of parts and spatial relations in object identification. *Perception, 22,* 229–248.
Chainay, H., & Humphreys, G. W. (2001). The real object advantage in agnosia: Evidence for a role of surface and depth information in object recognition. *Cognitive Neuropsychology, 18,* 175–191.
Chakravarty, I., & Freeman, H. (1982). Characteristic views as a basis for three-dimensional object recognition. In *Proceedings of the Society for*

Photo-Optical Instrumentation Engineers Conference on Robot Vision (Vol. 336, pp. 37–45). Bellingham, WA: Society for Photo-Optical Instrumentation Engineers.

Cunningham, D. W., Shipley, T. F., & Kellman, P. J. (1998). The dynamic specification of surfaces and boundaries. *Perception, 27,* 403–415.

Davidoff, J., & Wilson, B. (1985). A case of visual agnosia showing a disorder of presemantic visual classification. *Cortex, 21,* 121–134.

Dickinson, S. J., Bergevin, R., Biederman, I., Eklundh, J.-O., Muck-Fairwood., R., Jain, A. K., & Pentland, A. (1997). Panel report: The potential of geons for generic 3-D object recognition. *Image and Vision Computing, 15,* 277–292.

Edelman, S. (1997). Computational theories of object recognition. *Trends in Cognitive Sciences, 1,* 296–304.

Edelman, S. (1999). *Representation and recognition in vision.* Cambridge, MA: MIT Press.

Enns, J. T., & Rensink, R. A. (1991). Preattentive recovery of three-dimensional orientation from line drawings. *Psychological Review, 98,* 335–351.

Farah, M. J. (1990). *Visual agnosia: Disorders of object recognition and what they tell us about normal vision.* Cambridge, MA: MIT Press.

Feldman, J. (2003). What is a visual object? *Trends in Cognitive Science, 7,* 252–256.

Grossberg, S. (2000). The complementary brain: Unifying brain dynamics and modularity. *Trends in Cognitive Science, 4,* 233–246.

Grossberg, S., & Mingolla, E. (1985). Neural dynamics of perceptual grouping: Textures, boundaries and emergent segmentation. *Perception & Psychophysics, 38,* 141–171.

Grossberg, S., & Swaminathan, G. (2004). A laminar cortical model of 3D perception of slanted and curved surfaces and of 2D images: Development, attention and bistability. *Vision Research, 44,* 1147–1187.

Gum, T. (1995). *Psyclab v2.2: Macintosh applications programme* [Computer software]. Montreal, Ontario, Canada: Psyclab.

Guzman, A. (1968). Decomposition of a visual scene into three-dimensional bodies. *AFIPS Proceedings of the Fall Joint Computer Conference, 33,* 291–304.

He, Z. J., & Nakayama, K. (1992, September 17). Surfaces versus features in visual search. *Nature, 359,* 231–233.

He, Z. J., & Nakayama, K. (1995). Visual attention to surfaces in 3D space. *Proceedings of the National Academy of Sciences, USA, 92,* 11155–11159.

Hoffman, D. D., & Richards, W. (1984). Parts of recognition. *Cognition, 18,* 65–96.

Hoffman, D. D., & Singh, M. (1997). Salience of visual parts. *Cognition, 63,* 29–78.

Hummel, J. E. (1997). Structure and binding in object perception. In J. W. Donahoe & V. Packard (Eds.), *Neural-network models of cognition: Biobehavioral foundations. Advances in psychology* (Vol. 121, pp. 203–219). Amsterdam: Elsevier Science.

Hummel, J. (2000). Where view-based theories break down: The role of structure in human shape perception. In E. Dietrich & B. Arthur (Eds.), *Cognitive dynamics: Conceptual and representational change in humans and machines* (pp. 157–185). Mahwah, NJ: Erlbaum.

Hummel, J. E. (2001). Complementary solutions to the binding problem in vision: Implications for shape perception and object recognition. *Visual Cognition, 8,* 489–517.

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review, 99,* 480–517.

Hummel, J. E., & Stankiewicz, B. J. (1996). An architecture for rapid, hierarchical structural description. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 93–121). Cambridge, MA: MIT Press.

Humphrey, G. K., Goodale, M. A., Jakobson, L. S., & Servos, P. (1994). The role of surface information in object recognition: Studies of a visual form agnosic and normal subjects. *Perception, 23,* 1457–1481.

Jackendoff, R. (1990). *Semantic structures.* Cambridge, MA: MIT Press.

Julesz, B. (1981, March 12). Textons, the elements of texture perception, and their interaction. *Nature, 290,* 91–97.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolations in object perception. *Cognitive Psychology, 23,* 141–221.

Koenderink, J. J., & van Doorn, A. J. (1979). Internal representation of solid shape with respect to vision. *Biological Cybernetics, 32,* 211–216.

Kurbat, M. A. (1994). Structural description theories: Is RBC/JIM a general purpose theory of human entry-level object recognition? *Perception, 23,* 1339–1368.

Lee, Y.-L., & Park, R.-H. (2002). A surface-based approach to 3-D object recognition using a mean field annealing neural network. *Pattern Recognition, 35,* 299–316.

Leek, E. C. (1998a). The analysis of orientation-dependent time costs in visual recognition. *Perception, 27,* 803–816.

Leek, E. C. (1998b). Effects of stimulus orientation on the identification of common polyoriented objects. *Psychonomic Bulletin & Review, 5,* 650–658.

Leek, E. C., & Arguin, M. (2000, November). *Surface-based shape descriptions in object recognition: Evidence from whole–part matching*. Paper presented at the 41st Annual Meeting of the Psychonomics Society, New Orleans, LA.

Leek, E. C., Reppa, I., & Arguin, M. (2002, November). *Surface-based shape representations mediate object recognition: Further evidence from part–whole priming*. Paper presented at the 43rd Annual Meeting of the Psychonomics Society, Kansas City, KS.

Leek, E. C., Reppa, I., & Arguin, M. (2003). The structure of object shape representations in visual recognition. *Perception, 32*(Suppl.), 120.

Leek, E. C., Reppa, I., & Tipper, S. (2003). Inhibition-of-return for objects and locations in static displays. *Perception & Psychophysics, 65,* 388–395.

Lehky, S. R., & Sejnowski, T. J. (1988, June 2). Network model of shape-from-shading: Neural function arises from both receptive and projective fields. *Nature, 333,* 452–454.

Lehky, S. R., & Sejnowski, T. J. (1990). Neural network model of visual cortex for determining surface structure from images of shaded surfaces. *Proceedings of the Royal Society of London B: Biological Sciences, 240,* 251–278.

Liu, C. H., Collin, C. A., & Chaudhuri, A. (2000). Does face recognition rely on encoding of 3-D surface? Examining the role of shape-from-shading and shape-from-stereo. *Perception, 29,* 729–743.

Lowe, D. G. (1987). Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence, 31,* 355–395.

Marr, D. (1982). *Vision.* San Francisco: Freeman.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London B: Biological Sciences, 200,* 269–294.

Nakayama, K., He, Z. J., & Shimojo, S. (1995). Visual surface representation: A critical link between lower-level and higher-level vision. In S. M. Kosslyn & D. Osherson (Eds.), *Visual cognition: Vol. 2. An invitation to cognitive science* (2nd ed., pp. 1–70). Cambridge, MA: MIT Press.

Nakayama, K., & Shimojo, S. (1992, September 4). Experiencing and perceiving visual surfaces. *Science, 257,* 1357–1363.

Nicholson, K. G., & Humphrey, G. K. (2001). Surface cues reduce the latency to name rotated images of objects. *Perception, 30,* 1057–1081.

Nishihara, H. K. (1981). Intensity, visible surface, and volumetric representations. *Artificial Intelligence, 17,* 265–284.

Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology, 9,* 441–474.

Palmer, S. E., & Rock, I. (1994). Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin & Review, 1,* 515–519.

Pentland, A. (1989). Shape information from shading: A theory about human perception. *Spatial Vision, 4,* 165–182.

Potmesil, M. (1983). Generating models of solid objects by matching 3D surface segments. In *Proceedings of the 8th International Joint Conference on Artificial Intelligence* (pp. 1089–1093). Karlsruhe, Germany: William Kaufmann.

Price, C. J., & Humphreys, G. W. (1989). Effects of surface detail on object categorization and naming. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 41*(A), 797–828.

Ramachandra, V. J. (1988). Perceiving shape from shading. *Scientific American, 121,* 58–65.

Reed, S. K. (1974). Structural descriptions and the limitations of visual images. *Memory & Cognition, 2,* 329–336.

Reed, S. K., & Johnsen, J. A. (1975). Detection of parts in patterns and images. *Memory & Cognition, 3,* 569–575.

Reppa, I., & Leek, E. C. (2003). The modulation of inhibition-of-return across object-internal structure: Implications for theories of object-based attention. *Psychonomic Bulletin & Review, 10,* 493–502.

Sajda, P., & Finkel, L. H. (1995). Intermediate level visual representations and the construction of surface perception. *Journal of Cognitive Neuroscience, 7,* 267–291.

Sanocki, T., Bowyer, K. W., Heath, M. D., & Sarkar, S. (1998). Are edges sufficient for object recognition? *Journal of Experimental Psychology: Human Perception and Performance, 24,* 340–349.

Servos, P., Goodale, M. A., & Humphrey, G. K. (1993). The drawing of objects by a visual form agnosic: Contribution of surface properties and memorial representation. *Neuropsychologia, 31,* 251–259.

Siddiqi, K., Tresness, J., & Kimia, B. B. (1996). Parts of visual form: Psychophysical aspects. *Perception, 25,* 399–424.

Stankiewicz, B. J. (2002). Empirical evidence for independent dimensions in the visual representation of three-dimensional shape. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 913–932.

Tarr, M. J., & Bulthoff, H. H. (1998). *Object recognition in man, monkey, and machine.* Cambridge, MA: MIT Press.

Tversky, T., Geisler, W. S., & Perry, J. (2003). Contour grouping: Is there something special about closed contours? [Abstract]. *Journal of Vision, 3,* 121a.

Ullman, S., & Basri, R. (1991). Recognition by linear combinations of models. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 13,* 992–1005.

Vecera, S. P., Behrmann, M., & Filapek, J. C. (2001). Attending to the parts of a single object: Part-based selection limitations. *Perception & Psychophysics, 63,* 308–321.

Vecera, S. P., Behrmann, M., & McGoldrick, J. (2000). Selective attention to the parts of objects. *Psychonomic Bulletin & Review, 7,* 301–308.

Witkin, A. P. (1981). Recovering surface shape and orientation from texture. *Artificial Intelligence, 17,* 17–45.

Xu, Y., & Singh, M. (2002). Early computation of part structure: Evidence from visual search. *Perception & Psychophysics, 64,* 1039–1054.

Received September 30, 2003
Revision received September 8, 2004
Accepted December 16, 2004 ■