



## Perceptual comparison of features within and between objects: A new look

S.J. Harrison<sup>a,\*</sup>, J. Feldman<sup>b</sup>

<sup>a</sup> SUNY State College of Optometry, Vision Sciences, 33 West 42nd Street, New York, NY 10036, United States

<sup>b</sup> Department of Psychology, Center for Cognitive Science, Rutgers University, New Brunswick, NJ 08854, United States

### ARTICLE INFO

#### Article history:

Received 11 July 2008

Received in revised form 10 August 2009

#### Keywords:

Object formation  
Perceptual grouping  
Object benefit  
Feature comparison  
Axis orientation

### ABSTRACT

The integration of spatially distinct elements into coherent objects is a fundamental process of vision. Yet notwithstanding an extensive literature on perceptual grouping, we still lack a clear understanding of the representational consequences of grouping disparate visual locations. We investigated this question in a feature comparison task; subjects identified matching features that belonged either to the same apparent object (within-object condition) or to different apparent objects (between-object condition). The stimulus was backward-masked at a variable SOA, to examine the consequences of changes in the perceptual organization of the segments over time. Critical to our aims, the two objects composing our stimulus were occluded to a variable extent, so that differences in within-object and between-object performance could be unequivocally related to the formation of objects. For certain stimulus arrangements, we found superior performance for within-object matches. The pattern of performance was, however, highly dependent on the stimulus orientation and was not related to the strength of the object percept. Using an oblique stimulus arrangement, we observed superior between-object comparisons that *did* vary with the object percept. We conclude that performance in our feature comparison task is strongly influenced by spatial relations between features that are independent of object properties. Indeed, this dominating effect may hide an underlying mechanism whereby formation of a visual object suppresses comparison of distinct features within the object.

© 2009 Elsevier Ltd. All rights reserved.

### 1. Introduction

Rapid grouping of disparate visual elements into coherent objects underlies our interpretation of everyday scenes, and hence the ease with which we navigate the world around us. Objects are formed when local and separable components are perceived as belonging to a common body. Notwithstanding a long history of research into the cues and mechanisms underlying perceptual grouping, crucial aspects of it are still not understood. In this paper we consider the consequences of grouping on the representation of the composite elements—the way that perception of visual elements changes as a result of their being grouped together.

The influence of object formation is more often considered in terms of the allocation of attentional resources: while it has long been known that the operation of attention in visual space has a spatial component, likened to a “spotlight” of variable size (Downing & Pinker, 1985; Eriksen & Eriksen, 1974; Eriksen & Hoffman, 1973; Posner, Snyder, & Davidson, 1980), studies in recent years have addressed the role of objects in the operation of attention. Object representations wield an often-decisive effect on the deploy-

ment and consequent spread of attention within structured visual scenes, with object boundaries delimiting the spatial extent of the area over which the effects of attention are observed (Driver, Davis, Russell, Turatto, & Freeman, 2001). For instance, one body of work has demonstrated that two properties of the same object are identified faster (e.g. Behrmann, Zemel, & Mozer, 1998), or more accurately (e.g. Duncan, 1984), than if the two properties belong to different objects. Another approach has shown that detection of a new feature is faster when the feature appears within the object to which attention was already directed than if the feature appears within a different object (e.g. Egly, Driver, & Rafal, 1994). These types of benefits for features located within the same object compared to those located in different objects have also been shown to hold for 3D objects that extend in depth (Atchley & Kramer, 2001), and for occluded objects (Moore, Yantis, & Vaughan, 1998).

The studies described above have underscored the effect that perceptual grouping and consequent object formation have on the operation of attention. In this paper we sought to elucidate the complementary effect that object formation has on the underlying representation of the grouped elements. Prior to grouping, visual elements are (at least fleetingly) represented as distinct and independent entities, each with their own identifiable properties. Once elements have been grouped and interpreted as belonging to an object, do we still perceive the elements in the same way,

\* Corresponding author. Fax: +1 212 938 5760.

E-mail address: [sharrison@sunyopt.edu](mailto:sharrison@sunyopt.edu) (S.J. Harrison).

or are their representations irrevocably altered? What is the essence of grouping in terms of perceptual consequences?

It could be anticipated, extrapolating from the established within-object benefit for attentional resources, that multiple elements belonging to the same object would be more accurately perceived. This would imply that grouping results in a stronger representation of composite elements. However under some circumstances, within-object benefit is not observed: In a task requiring subjects to judge two features as “same” or “different”, Davis et al. (Davis, 2001b; Davis & Holmes, 2005; Davis, Welch, Holmes, & Shepherd, 2001) found, unusually, that comparisons *between* objects were facilitated. They concluded that observations of within-object benefit were specific to conditions common to most previous studies, and did not apply to the stimuli and methodology that they had used. Their work suggests that within-object benefits may not generalize to all cases, and so we surmise that it is not a foregone conclusion that multiple elements belonging to the same object are more accurately distinguished than elements belonging to different objects.

What basis is there for questioning the applicability of the established within-object benefit to feature representation? Within what framework might impairments of feature comparisons within objects be understood? When image regions are grouped together, their properties take on new meaning in the context of the newly-formed object. First of all, as in conventional Gestalt rhetoric, the object itself has properties distinct from those of any of its constituent elements—its centroid, size, orientation, and so forth. Additionally, in common with the notion of object-centered representations, the properties of each constituent element may now be represented only in relation to the object’s own properties, with some consequent loss of precision or accessibility of the absolute values of component properties. Indeed one of the principal “benefits” of object representation, from the perspective of representational efficiency, would be the suppression of the heterogeneous and highly redundant properties of separate elements in favor of a much smaller set of summary object properties.

Moreover, it is not hard to envision situations where it would make ecological sense for comparisons *between* objects to be facilitated. In naturalistic tasks it is often the differences between objects that drive meaningful object comparisons, and that determine the division between what are perceived as separate objects in the first place. In contrast, variations of a property within a single object are often perceived as less functionally meaningful (unless they are on perceptually distinct object parts, in which case the parts are treated much as are separate objects; (see Barenholtz & Feldman, 2003)). Consider as an example the majority of smooth or random variations in texture (Ben-Shahar, 2006; Ben-Shahar & Zucker, 2004; Nothdurft, 1992); the non-homogeneous nature of the texture orientation is not interpreted by the visual system as salient.

These considerations lead to the perspective that perceptual formation of an object could be expected to *inhibit* performance in tasks requiring overt comparisons between the visual elements within the object. In the following experiments, we measured the accuracy of comparison of features within vs. between objects, paying careful attention to factors that may have confounded previous studies. Note that our motivation differed from that of previous studies at a very fundamental level. We were not seeking to observe how attention moves within a visual scene, or is constrained by object representations; we wanted to illuminate the representational consequences when disparate elements are grouped into an object. This motivation determined many details of our stimulus and task design, which will be detailed below. However given that our approach is broadly similar to that in much of the object-based attention literature, perhaps the most notable and novel aspect of our study was that we modulated the strength

of object formation, in order to firmly establish whether observed differences in performance in within- and between-object conditions were directly related to this factor.

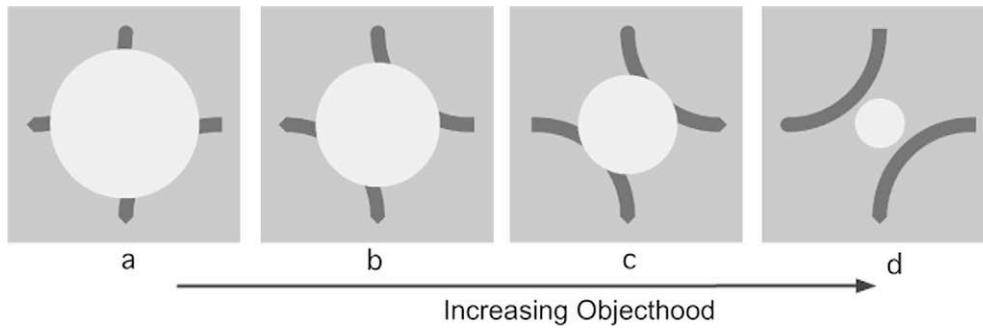
## 2. Experimental approach

Our study aimed to measure the perceptual consequences of grouping on the representations of component visual elements. We asked whether representations of the component elements, assessed via subjects’ performance in a comparison task, are heightened, suppressed, or perhaps unchanged, by the formation of “whole” objects. To this end, we designed a stimulus where four segments could group into two objects behind a central occluder (Fig. 1). Each distal end of a segment had a distinct shape. Prior to presentation of the stimulus, a spatial cue indicated to observers one location out of the four imminent end-shape locations. At stimulus onset, the observer then found a matching “feature” at one of the other three locations, a task that requires comparison of multiple features. We measured performance as a function of whether the two critical locations (the cued feature location and the matching feature location) did or did not belong to the same perceptual object.

When measuring the ability of subjects to correctly locate matching features that did or did not belong to the same object as the cued feature, we wanted to directly relate any difference in performance to the formation of coherent objects in the display. To accomplish this, we varied the degree to which our object segments perceptually “grouped,” as we reasoned that any effect that is due to the grouping of parts should vary with the strength of grouping of those parts. Additionally, in order to investigate the temporal evolution of effects related to object-formation, we presented our stimuli for brief exposures at a range of durations, and masked them immediately at offset. Backward-masking was chosen over a reaction time methodology, as speeded tasks reflect the time required for a final perceptual decision to be reached; in contrast, we wanted to track the evolution of the effect of grouping on the feature representations, as reflected in our feature-matching task.

By convention, we will refer to the two equidistant feature-match conditions as “within” and “between”, as this terminology is widely used. However in some respects the “within/between” terminology is a misnomer in our task, as correctly locating the matching feature required assessment of features not just within an object or between objects according to the condition, but likely across the whole stimulus. Moreover, it should be born in mind that when perceptual grouping of the four segments is weak (Fig. 1a and b), the term “within” is particularly misleading, as the percept of two objects may not occur. Nevertheless for consistency with established work on the effects of object formation on attention, we will use the term “within” to refer to the condition where matching features within the stimulus configuration would be perceived as belonging to the same object, were grouping strong enough to support object formation.

Aside from the previously stated aims of this study, an outstanding experimental factor that we wanted to address was the effect of stimulus orientation. Many previous studies of object-based effects have used two-object displays, with object axes lying either vertically or horizontally. Under this stimulus arrangement, half of the cases of the within-object condition consist of vertically-separated features and half consist of horizontally-separated features, and likewise for between-object cases. Some authors have found no effect of the orientation of object axes (Egly et al., 1994; Moore et al., 1998; Watson & Kramer, 1999), while others have made no mention of orientation effects (Atchley & Kramer, 2001; Chen, 1998); yet others have reported a strong effect or an interaction of object orientation with the within- and between-



**Fig. 1.** The stimulus consisted of two arc objects, here oriented obliquely, which were occluded by a central disk. (a–d) As the diameter of the occluding disk decreased, the strength of grouping of the four arc-segments into two objects increased. At the highest objecthood level, the two objects were fully visible.

object conditions (Davis, Driver, Pavani, & Shepherd, 2000; Davis et al., 2001), corroborating our concern that this issue has not been given sufficient consideration. The orientation of object axes is potentially critical because comparison of horizontally-aligned elements may benefit from early symmetry detection mechanisms, which favor comparisons across the vertical midline (Wagemans, 1997). This could create a serious confound for studies seeking to investigate object-based mechanisms, in particular when using feature-comparison paradigms, as some of the performance patterns found might derive from global symmetry detection mechanisms that are unrelated to object formation. The issue of stimulus orientation was therefore also investigated in our experiments.

### 3. Method

#### 3.1. Stimuli

Stimuli consisted of two arc-shaped objects and a central disk. On most trials, the central disk occluded the two arcs such that they appeared as four segments, which could be perceived as completing amodally behind the disk (Fig. 1a–d). The background luminance was  $80 \text{ cdm}^{-1}$ , and stimuli were composed of either white parts at  $130 \text{ cdm}^{-1}$  and a black disk at  $30 \text{ cdm}^{-1}$  or vice versa. Stimuli were generated and presented via Matlab (The MathWorks, Natick, MA) utilizing functions provided in the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997), and were viewed on a G4 eMac (viewable screen size  $30 \times 22 \text{ cm}$ ,  $1152 \times 864$  pixels, 80 Hz refresh rate).

Stimuli were presented centrally and were viewed at a distance of 45 cm, which resulted in a monitor pixel resolution of 30 pixels  $\text{deg}^{-1}$ . The two arc-shaped objects were portions of circular annuli; they had a width of 30 pixels, an outer radius of 200 pixels, and each subtended  $\pi/2$  of circular arc. The arcs were positioned back-to-back such that the four arc endings fell at the vertices of an (imaginary) square with sides of length 304 pixels (just over  $10^\circ$  visual angle). This geometry, in combination with the  $\pi/2$  arc length, results in the four terminating faces of the arcs being equidistant from the midpoint of the display and facing directly outwards i.e. the locations of the terminating faces of the arcs were  $90^\circ$  rotational translations of each other. A shape segment was added to each arc ending so as to modify the arc outline. The segment could be a rectangle, a semi-circle or a triangle, and these shapes were the features of comparison in our task. The exact dimensions of the three shapes were chosen such that they were perceived as having the same spatial extent; their total area was approximately, but not exactly, equal.

The central occluding disk had one of four possible sizes, with a radius of 50, 100, 125 or 150 pixels ( $1.67^\circ$ ,  $3.33^\circ$ ,  $4.12^\circ$ ,  $5^\circ$  of visual angle). The size of the occluding disk determined the support ratio for the four arc-segments' grouping into two objects; the smaller the disk, the larger the support ratio and the stronger the perceived

grouping into two objects. Hence the four sizes of disk will be referred to from here on as objecthood levels. The largest disk, which occluded the largest portion of the arcs, corresponded to the first, lowest, objecthood level (Fig. 1a). At the lowest objecthood level, only the very ends of the two objects were visible, hence the direction of their curvature and their underlying grouping could not be determined from the four visible segments without prolonged scrutiny. As the grouping cues were very weak at this objecthood level, it was ambiguous which features belonged to common coherent objects (or indeed whether any underlying objects existed); we would not expect to find differences between the within-object and between-object conditions at this objecthood level. At the highest objecthood level, the two objects were fully visible and it was completely unambiguous which features belonged to common objects (Fig. 1d). In this case differences between the within and between conditions, should they exist, would be most apparent. The four objecthood levels therefore spanned the full continuum of possible strengths of the object percepts, and hence the full range of object-related variation in performance should be observed.

To prevent the absolute area or length of the four segments from being used to assess similarity of the shape features at the brief presentation times used here, the location of the central occluding disk was jittered in the horizontal and vertical directions, with the offset in each direction drawn randomly from a rectangular distribution of  $\pm 10$  pixels. (The location of the two arcs remained unchanged.) This meant that the visible area of the four arc-segments relative to each other varied randomly, but without affecting the visibility of the shape features at the arc terminations.

#### 3.2. Procedure

The presentation sequence of a single trial is illustrated in Fig. 2. Stimulus presentation was preceded with a fixation cross. Subjects initiated each trial by pressing the space bar. The central disk was then presented for 1 s, together with a small red dot indicating one of the four arc segment locations. The four arc-segments then appeared, as if behind the occluding disk, and with the location cue still present. The entire stimulus display remained for 50, 62, 88, 138 or 188 ms (4, 5, 7, 11 or 15 screen refreshes); these times were chosen following pilot work to be the most potentially revealing. Stimuli were masked with a checkerboard grid of contrasting blocks, which was presented immediately at stimulus offset for 138 ms (11 screen refreshes). The mask covered an area slightly greater than that subtended by the stimulus.

Subjects were instructed to maintain fixation during stimulus presentation and to indicate the location of the feature that matched the feature at the cued location. (Subjects were not required to report the identity of the feature.) In order to successfully perform the task, observers needed to identify the feature at the cued location whilst maintaining a broad focus so as to assess

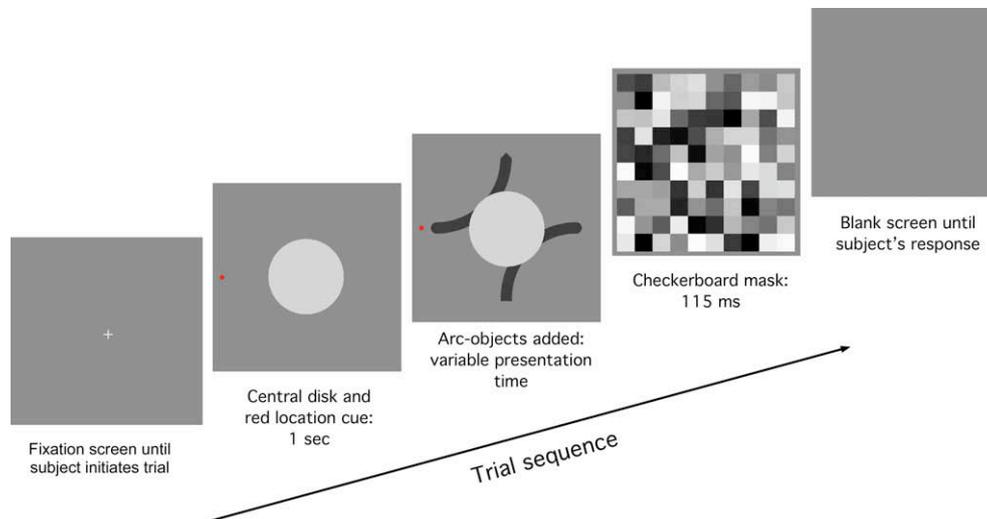


Fig. 2. Sequence of a single trial. Note that the stimulus images are cropped here for illustrative purposes.

the features at the remaining locations. The cue always indicated the location of the feature shape to be matched, but gave no indication as to which of the three other possible locations would contain the match. (Note this is quite different from a conventional object-based attention cueing paradigm, where a cue indicates the probable location of the target. Here, the cue never indicated the location of the target.) There was always one, and only one, matching feature present in the stimulus. Eye movements were not monitored, however eye movement towards the cued location prior to feature presentation, had they occurred would not invalidate our comparison of within- and between-object performance, as both critical locations were equidistant from the cued location.

The spatial relation of the matching feature to the cued feature varied among three conditions: the matching feature could be at the other end of the same arc (“within-object”), at the equidistant location on the other arc (“between-object”) or at the most distant, “diagonal”, location. (Trials where the observer responded that the match location was the same as the cued location were extremely rare, and were discarded as they indicated that the observer was unaware of the cue and had most likely been momentarily distracted.) The shape of the feature to be matched was chosen randomly on each trial. The two non-match features occupied the remaining two arc-termination locations.

Comparison of performance in the within-object and between-object conditions was our primary interest, but the diagonal-match condition was included because otherwise subjects could achieve correct performance by inspecting only one potential feature match location in addition to the cued location. In essence, inclusion of the diagonal-match condition encouraged subjects to evaluate the entire stimulus area. Our stimulus and task were designed to assess the simultaneous representations of multiple features—subjects’ success at the task would be achieved by monitoring all stimulus features, both matching/relevant and non-matching/irrelevant, with the four feature locations delimiting the spatial extent of the display. This is unlike previous feature-comparison paradigms where typically only two features are presented. Additionally, we would expect any observable object-based effects to be strongest under conditions of spread attention (e.g. Goldsmith & Yeari, 2003), such as we encouraged here.

#### 4. Experiment 1

In this first experiment investigating the impact of grouping on the representations of component features, our stimulus was

oriented so that the four features fell at the vertices of a square tilted by 45°. This “oblique” configuration leads to two possible groupings of the four arc segments, into two objects tilted either left from vertical or right from vertical (here termed left-oblique and right-oblique respectively), and so allowed us to address our concerns about the possible special status of horizontally- and vertically-oriented objects. In the oblique configuration, the critical within- and between-object features are obliquely displaced from the cued feature, i.e. comparisons are made between obliquely-displaced features, rather than vertically- and horizontally-displaced features. Subjects indicated the location of the matching feature using the keys “8”, “6”, “2” and “4”, which correspond spatially to the four arc segments.

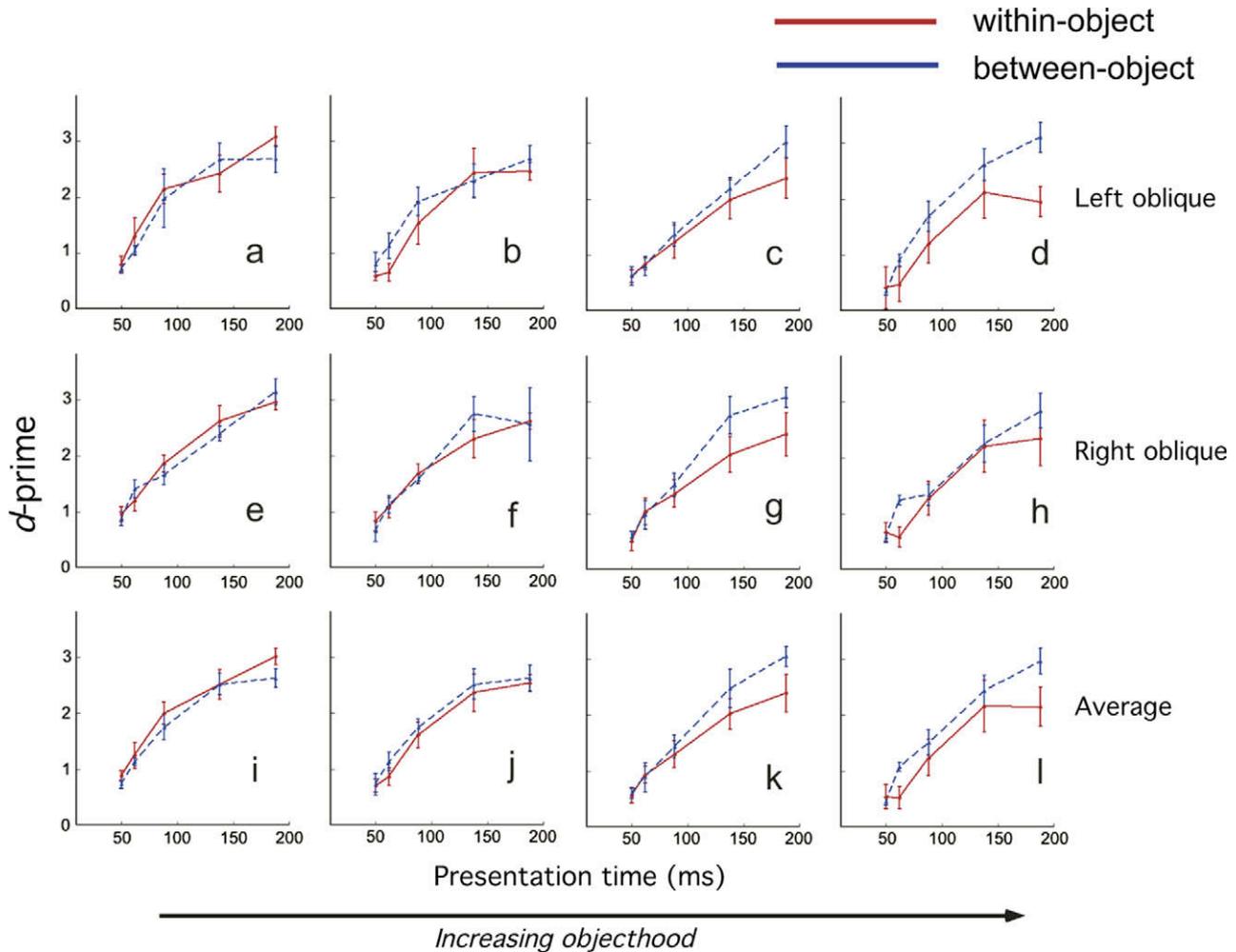
##### 4.1. Subjects

Five naïve subjects were paid for their participation in Experiment 1. Each took part in five testing sessions of approximately 1 h each. Each subject generated a total of 4800 responses, which resulted in 80 responses for each permutation of within/between/diagonal condition, objecthood level and presentation time.

##### 4.2. Results

Percent correct data was converted to a normally-distributed  $d$ -prime performance measure for subsequent statistical analysis, using the Smith (1982) approximation for a 3-AFC task. However it should be emphasized that we duplicated our analyses using percent-correct, and all findings and conclusions were identical.  $d$ -prime values were analyzed with a 3-way repeated-measures analysis of variance (ANOVA); within/between condition (2)  $\times$  object orientation (2)  $\times$  objecthood level (4)  $\times$  presentation time (5). Results averaged over observers are shown in Fig. 3, split into the left- and right-oblique orientations (top two rows; a–d, e–h respectively) and averaged over both orientations (bottom row, i–l).

The within/between condition was not a significant factor in its own right ( $F(1, 4) = 2.606, p = .182$ ), and likewise the two object orientations were equivalent ( $F(1, 4) = .436, p = .545$ ). As could be expected, presentation time was highly significant ( $F(4, 16) = 135.631, p < .001$ ). Objecthood level was also significant ( $F(3, 12) = 6.177, p = .009$ ); increasing the strength of the object percept decreased performance levels. The effect of objecthood level on performance was unchanged when the highest, unoccluded objects, objecthood level was removed from the analysis ( $F(2, 8) = 9.902, p = .009$ ).



**Fig. 3.** Results from Experiment 1, showing  $d$ -prime values in the within- and between-object conditions, at increasing objecthood levels (left to right). Results are separated into left-oblique object completion (top row; a–d), right-oblique object completion (middle row; e–h), and averaged over both orientations (bottom row, i–l). Error bars are standard error of subjects' mean  $d$ -prime values.

The interaction between within/between condition and presentation time was significant: Although performance in the within- and between-object conditions was very similar at the shortest presentation times, performance improved more rapidly and reached higher final performance levels in the between-object condition ( $F(4, 16) = 3.722, p = .025$ ). In addition, objecthood level affected within- and between-object performance differently: formation of object percepts was detrimental to performance when the matching features occurred within the same object, but not when they were part of different objects ( $F(3, 12) = 4.771, p = .021$ ). Again, this effect was significant even without inclusion of the highest objecthood level ( $F(2, 8) = 4.995, p = .040$ ). Overall, the nature of the performance pattern could be summarized by saying that as the segments became more strongly grouped into perceptual objects, due to both stronger grouping cues and longer presentation times, within-object performance was increasingly impaired.

#### 4.3. Discussion

In Experiment 1, where matching features were obliquely displaced from each other, we found no evidence of a within-object benefit for identifying matching features. This is not altogether surprising, as at our lowest objecthood level there was negligible grouping between arc segments and we would not expect to see

any difference for within- compared to between-object performance in these “ungrouped” stimuli: Any benefit for the within-object condition could be expected to increasingly emerge as grouping cues became stronger. However, instead we found that identification of features was *hindered* when they were perceived as belonging to the same object. Note that because this difference between the two conditions increased with objecthood level, we can surmise that it is related to object formation. Additionally, the difference between the two conditions was not immediately apparent but was only seen at the longer presentation times, as would be expected if it was related to the evolving process of grouping the occluded arcs so as to form two perceptual objects.

Our finding of a between-object advantage for comparison of feature representations could be attributed to many factors: our solid, rather than outline, non-overlapping objects and simultaneous presentation of features and objects, rather than delayed presentation of features on top of existing objects, are both factors that have been cited as being unfavorable for observing within-object benefits, perhaps even promoting between-object benefits (Davis & Holmes, 2005). Additionally, our task was different to any used previously, in that we encouraged a holistic assessment of the entire stimulus rather than limited processing of just two salient locations, and observed the ongoing process of object formation rather than using a reaction-time measure. Finally, features were not compared across the horizontal and vertical midlines. The

horizontal/vertical effect has generally been considered to have negligible influence on the otherwise-robust within-object benefit. We decided to assess the influence of this factor using our existing stimulus, to see if introducing horizontal and vertical feature comparisons would be sufficient to reproduce the well-known within-object benefit, despite other differences in our task and stimulus.

## 5. Experiment 2

In our second experiment, the stimulus was presented such that the four features fell at the vertices of an *upright* square. The task itself was as in Experiment 1. Subjects indicated the location of the matching feature using the keys “4”, “5”, “1” and “2”, which correspond spatially with the new locations of the arc segments. The “upright” configuration leads to two possible groupings of the four arc segments, into either two horizontal objects or two vertical objects. Hence the within- and between-object conditions correspond to matching features being horizontally or vertically displaced from each other.

It rapidly became apparent that a second change was necessary if we were to observe any differential performance between conditions: when feature comparisons were made in the vertical and horizontal directions, rather than in the oblique directions, shorter

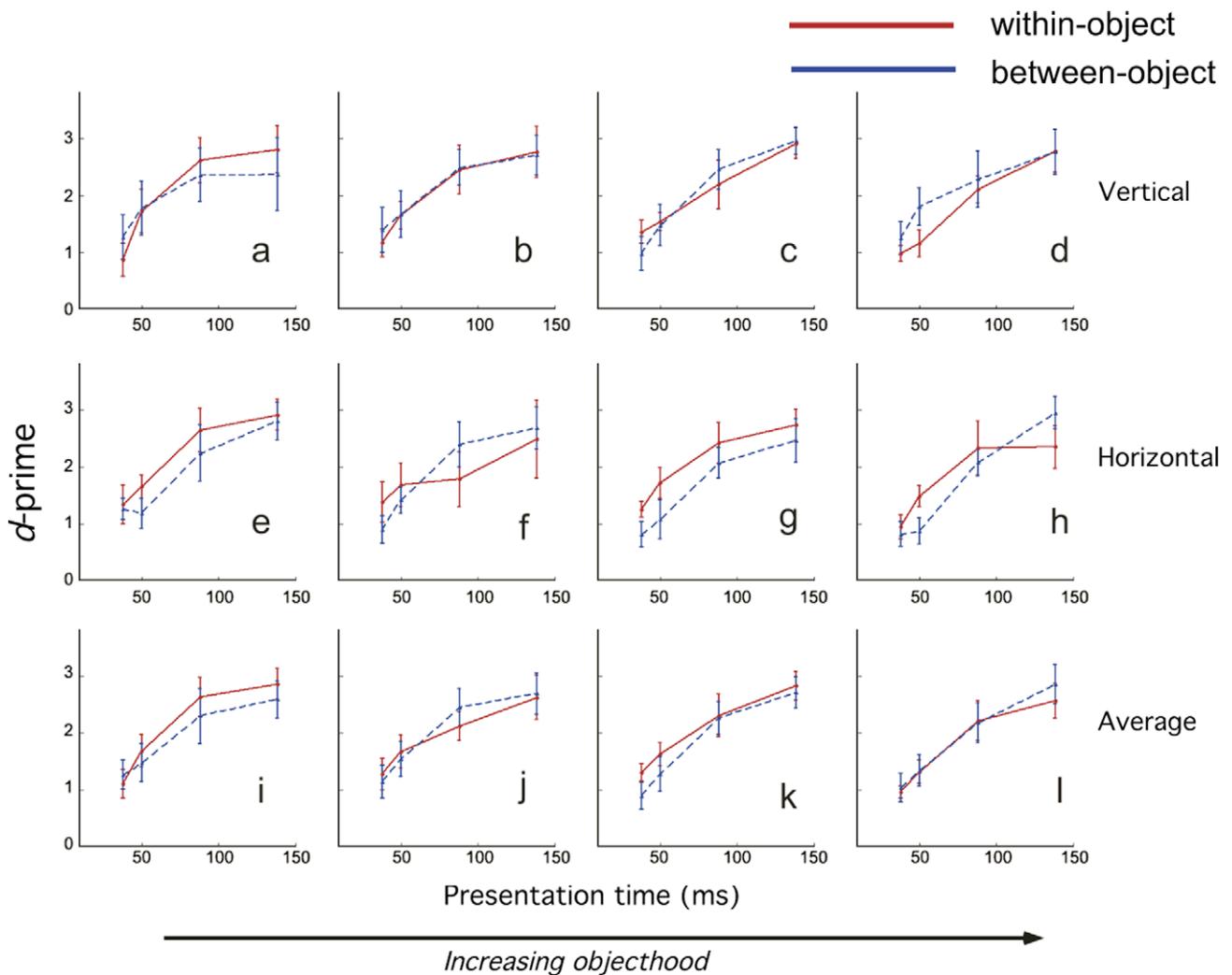
presentation times were needed if we were to observe non-ceiling performance. The presentation times ultimately used in this experiment were 38, 50, 88 and 138 ms (3, 4, 7 and 11 screen refreshes).

### 5.1. Subjects

Four naïve subjects were paid for their participation in Experiment 2, which was also completed by one of the authors. The four subjects each took part in 5 testing sessions of approximately 1 h. Each subject generated a total of 3840 responses, which resulted in 80 responses per subject for each permutation of within/between/diagonal condition, object orientation, objecthood level and presentation time.

### 5.2. Results

As in Experiment 1, percent correct data was converted to a normally-distributed  $d$ -prime performance measure for subsequent statistical analysis, using the Smith (1982) approximation for a 3-AFC task. Again our findings and conclusions were identical when performance was analyzed in terms of percent-correct.  $d$ -prime values were analyzed with a 4-way repeated-measures analysis of variance (ANOVA); within/between condition (2)  $\times$  object orien-



**Fig. 4.** Results from Experiment 2, showing  $d$ -prime values in the within- and between-object conditions, at increasing objecthood levels (left to right). Results are separated into vertical object completion (top row; a–d), horizontal object completion (middle row; e–h), and averaged over both orientations (bottom row, i–l). Error bars are standard error of subjects' mean  $d$ -prime values.

tation (2)  $\times$  objecthood level (4)  $\times$  presentation time (5). Results are plotted in Fig. 4, split into the vertical and horizontal orientations (top two rows; a–d, e–h respectively) and averaged over both orientations (bottom row, i–l).

As before, the within-object and between-object conditions were not significantly different overall ( $F(1, 4) = 1.113$ ,  $p = .351$ ), and there was a highly significant effect of presentation time, with performance improving with presentation time ( $F(3, 12) = 42.831$ ,  $p < 0.001$ ). However in contrast to the oblique stimulus arrangement, performance in the within- and between-object conditions was affected differently by object orientation: When the segments grouped so as to form two vertically-oriented objects, between-object comparisons were slightly more accurate than within-object comparisons; conversely when the objects formed horizontally, between-object comparisons were notably impaired i.e. a within-object benefit ( $F(1, 4) = 10.365$ ,  $p = .032$ ). This pattern of performance accounts for the trend for superior performance for vertical objects ( $F(1, 4) = 4.907$ ,  $p = .091$ ). In the absence of the highest, unoccluded objects, objecthood level, the interaction of within/between condition and object orientation did not reach significance.

Critically, there was no effect of objecthood level ( $F(3, 12) = .826$ ,  $p = 0.505$ ), and moreover the effect of object formation was the same for both within- and between-object conditions ( $F(3, 12) = 1.283$ ,  $p = 0.325$ ). Even when horizontal objects, which showed an apparent within-object benefit, were analyzed separately, there was still no significant interaction of within/between condition and objecthood level. This means that neither the higher performance levels for vertically-oriented objects, nor the apparent within-object benefit for horizontally-oriented objects, can be attributed to the formation of object representations from the arc segments.

### 5.3. Discussion

The stimulus configuration in Experiment 2 was similar to that used in many previous studies in the object-based attention literature, with the two perceptual objects oriented either vertically or horizontally. However, unlike many previous studies, we could not conclude that inclusion within the same object conveys a benefit on comparison of feature representations. As we noted earlier, weak grouping cues at our lower objecthood levels mean that any differences between the within- and between-object conditions would only emerge as grouping cues became stronger, and would be observable as an interaction between condition and objecthood level. However no such interaction was seen. As the percept of two objects emerged with increasing objecthood level, there was no systematic change in performance for within-object relative to between-object conditions.

Importantly, we did observe an apparent within-object benefit, but only for horizontally-oriented objects. This performance pattern showed no effect of objecthood level, a strong suggestion that the underlying mechanism may not be strongly related to object formation and may instead rely only on the relative spatial location of the features. Hence our main finding in Experiment 2 was that performance was strongly affected by object orientation; evidence was lacking that this effect related to the status of the features as belonging to objects.

We note that the longer presentation times necessary for performance to plateau in Experiment 1, in comparison to Experiment 2, are in closer agreement with the known time course of perceptual completion of occluded objects (e.g. Ringach & Shapley, 1996). (Despite assessing performance at shorter presentation times, overall performance was still higher in Experiment 2 than in Experiment 1.) The shift in the time course of perceptual processing to shorter durations is additional evidence that performance in Experiment 2 benefited substantially from the “upright” orientation of the stimulus, in a way that enabled faster comparisons of

features independently of, and conceivably even prior to, their grouping into objects.

We discuss possible reasons for the “upright” benefit later. But mechanisms aside, this finding confirms that our concerns of possible orientation confounds in some earlier studies were warranted. Our results in Experiment 2 raise the possibility that differences in performance between within- and between-object conditions are being hidden by an unrelated effect of object orientation. The influence of vertical and horizontal object formation should therefore be controlled or otherwise neutralized in studies comparing features within and between objects.

## 6. General discussion

The studies presented here used a novel target-matching paradigm to investigate the representational consequences of processes that group parts into objects. We wanted to know whether the representations of features were altered when they were perceived as belonging to a unified object as compared to when they were perceived as being independent of all other features. In particular, we aimed to evaluate how representations of features belonging to the same object compared to representations of features belonging to different objects. Importantly, we were able to evaluate whether performance differences were related to formation of object percepts, due to our manipulation of the strength of the object percept. To summarize, our results show that (a) representation of perceptual features within obliquely-oriented objects, as reflected in subjects' ability to compare and match features, is superior for between- compared to within-object conditions, but that (b) relative performance for the two conditions is highly dependent on stimulus orientation, with vertical and horizontal orientations leading to effects that are not clearly related to the formation of object percepts.

In contrast to the horizontal and vertical object orientations, our results for oblique stimulus orientations were related to the strength of object formation. This leads us to suggest that within-object suppression may be the true nature of the effect of object formation on feature representation. However any such underlying pattern of object-related performance appears to be easily swamped by a much larger effect of stimulus orientation. Critically, this latter effect has been played down or disregarded in many related studies of object-based attention, and indeed, has often been confounded with the within/between factor. Therefore orientation effects may have hindered an accurate assessment of object-based attentional effects as well.

As mentioned above, we are not the first to find within-object feature suppression rather than facilitation. In a series of two-feature-comparison experiments, Davis et al. found some cases of better performance for between-object tasks than for within-object tasks (Davis, 2001a; Davis & Holmes, 2005; Davis et al., 2001). While pointing towards the same general conclusion, our studies go beyond theirs in several respects. We used an objective *d*-prime measure rather than reaction times, avoiding some of the interpretive difficulties of RTs such as the possibility of speed-accuracy tradeoffs. Our task required visual assessment of multiple feature locations encompassing the entire figure rather than simply two cued locations, suggesting that the task involves basic object representations rather than isolated points of attention (see discussion below). We carefully controlled the orientation of the comparison (again see discussion below). Finally, we consistently varied the strength of object representations, allowing us to assess whether any effect was truly object-based. Hence our results confirm several of Davis's conclusions and extend them to a broader setting. Indeed, many of our key conclusions echo those of Davis and his collaborators.

However, the significance of these results and the doubt they throw on the conventional view of object-based effects does not seem yet to be widely appreciated. The vast majority of studies have found within-object benefits, even when using feature-comparison paradigms that, on the face of it, seem to resemble the present study. But, as suggested by Davis et al. and confirmed by our finding, apparently subtle differences in methodology seem to influence the presence and even the direction of object-based effects in a way that is not consistent with the conventional view of a ubiquitous within-object benefit. One such methodological detail in some earlier studies is the extensive pre-exposure of objects *prior* to the appearance of the features of interest. For example, many previous studies have used variations on the methodology of Egly et al. (1994), whereby salient features are flashed on top of pre-existing objects after a spatial cue to attention, and reaction times for feature detection are then measured. This methodology places strong demands on temporal aspects of the movement of exogenous attention, and has demonstrated that attention spreads within objects before spreading to other nearby objects.

In contrast to presenting objects prior to feature onset, we suggest that when the features of comparison are inherently contained in the objects, this may probe a subtly different aspect of object perception. Additionally, we feel that there is an important distinction to be made between feature comparison tasks where only two features are present—and hence it is implicit which parts of the image are to be compared—and the task used in the current study where a “match” had to be found among four features. The latter discourages a narrow focus of spatially-allocated attention and instead promotes a holistic, object-based, representation of the entire stimulus. Our interpretation is that the methodology used here, whereby the task-relevant features do not elicit exogenous attention, probes the accuracy or accessibility of the object-part representations. However, given that we could alter the relative performance in within- and between-object conditions by the simple manipulation of rotating our stimulus, in our studies orientation seems to have played the most substantial role, overwhelming other possible factors of stimulus design or task.

As well as the extensive literature documenting within-object superiority in tasks such as those described in the Introduction, such effects have also been used as an indicator of whether parts are grouped into a single object or multiple object representations (Ben-Shahar, Scholl, & Zucker, 2007; Feldman, 2007; Zemel, Behrmann, Mozer, & Bavelier, 2002), and at what level of representation objects and parts are perceived (Barenholtz & Feldman, 2003; Chen, 1998; Vecera, Behrmann, & McGoldrick, 2000; Watson & Kramer, 1999). The rationale underlying such studies is that attention selects the spatial region corresponding to an object for more detailed analysis than the rest of an image. But if our conclusions and those of Davis et al. are correct, such logic may only be valid within specific methodological constraints and with respect to certain aspects of perception.

### 6.1. Grouping and feature suppression

Another aspect of our results that calls for consideration is the direction of the performance change when parts were grouped into objects. In Experiment 1, where performance variation was related to object formation, performance decreased as the strength of grouping increased. Our results presumably depend on the brief stimulus exposures; these short intervals support “first glance” comparisons only and preclude serial inspection of distinct locations. By design, this time pressure allowed us to assess the distinctness of the simultaneous representations, as they enable comparison of the two critical spatial features, as opposed to a serial comparison process such as the sequential allocation of focused attention. Our results suggest that this distinctness

diminishes as features are increasingly perceptually bound, implying the creation of unified perceptual objects within which featural differences are no longer as strongly represented.

These observations raise more general issues about the processes underlying perceptual grouping. In our study, the location of all four features could be inferred in advance of feature presentation, and feature size was unchanged across conditions. If observers were able to restrict the locus of their attention to solely the features themselves, or the four informative spatial locations, then presumably this would result in the best performance (as the lowest objecthood condition demonstrated). Yet subjects' representations of features were affected by automatic processing of “irrelevant” portions of the stimulus. Features that were locally unchanged across all objecthood levels became “weaker” as the strength of their context increased. Viewed in these terms, our result that performance decreased with strength of grouping reflects the degradation of separate feature representations (or at least their accessibility in a comparison task) under stimulus conditions that induce grouping. This echoes the study of Rensink and Enns (1998), who observed in their search task that when fragments perceptually group, the groups become the basic units of attentional access, with subsequent processes unable to access the constituent fragments without increased attentional effort. Similarly, Mapelli, Cherubini, and Umiltà (2002) considered grouped and ungrouped conditions in a speeded feature comparison task. They also found that the grouping of features into objects was detrimental to performance. Finally, a recent study of vernier acuity (Malania, Herzog, & Westheimer, 2007) found that judgment of the offset of the two critical line elements was increasingly impaired as the lines were perceptually grouped with surrounding line elements. This could be explained if perceptual access to the representations of the two line elements composing the vernier was subsumed as the lines became part of a larger group.

A “loss” of information in the process of object formation is potentially a “gain” in representational efficiency, and in retrospect it is surprising that this aspect of object formation has not, to our knowledge, been directly investigated earlier. While our brief presentations are of course not typical of everyday visual experience, object formation processes are ubiquitous in visual function. Our investigation of the perceptual consequences at these early, “automatic”, stages of processing suggests that in the first instance the visual system packages features into objects, with a consequent weakening of the individual feature representations. We would suggest that changes in feature representations within the context of identifiable objects, is an area worthy of further investigation.

### 6.2. The influence of stimulus orientation

Our finding that the strength of grouping into two “upright”, vertical, objects did not significantly affect the pattern of performance indicates that other aspects of the stimulus were more influential in subjects' perception. It appears that this particular physical arrangement of features inherently facilitates comparisons. Other authors have previously noted faster same/different judgment of two features in speeded comparison tasks when features were horizontally displaced (Ben-Shahar et al., 2007; Davis & Holmes, 2005). The pattern of results in their study led Davis and Holmes (2005) to suggest there was an interaction between detection of feature symmetry and object formation, with the benefit of feature symmetry applying when a single object lay across the vertical meridian, but not when two objects lay either side of the vertical meridian as in the between-objects condition. Indeed, an extensive range of studies has demonstrated that detection of mirror symmetry across a vertical axis is faster and more accurate than across any other axis (e.g. Pashler, 1990; Wagemans, 1997). It may be that the visual system has a bias to detect symmetry within

rather than between objects, and that this bias is heightened when the symmetry is across the vertical meridian (Baylis & Driver, 1993; Koning & Wagemans, 2009).

Unfortunately such an explanation does not fit with our results, as we found very similar performance for matches across the vertical meridian, regardless of whether the features belonged to two objects or one. However in agreement with the general premise of these authors, we would suggest that symmetry effects can swamp the intended-to-be-observed effects of object formation. In support of this conjecture, further analysis of Experiment 1 (“oblique” stimulus) showed that the “diagonal” condition gave higher performance than both the within- and between-object conditions when the matching features were horizontally displaced (1.817 vs. 1.488 and 1.733 respectively) despite the greater spatial separation of features, but not when vertically displaced (1.520 vs. 1.698 and 1.783 respectively), though this interaction was not significant. This performance pattern did not appear to vary with objecthood level, arguing against the possibility that subjects in fact grouped the segments in the diagonal direction at the lowest objecthood levels where arc segments were small and the intended grouping cues were weak. Instead the pattern appears to be further confirmation of the non-object based horizontal advantage for feature comparisons: The diagonal match performance was again low relative to the within- and between-conditions in Experiment 2, when the feature comparisons were on a diagonal relative to the meridians, as well as to the stimulus. The previously-found facilitation of symmetry detection across the meridians has been attributed to attentional scanning strategies, whereby certain axes are default reference axes when the visual system is assessing the existence of an axis of symmetry (Pashler, 1990; Wenderoth, 1994).

Overall, our results do confirm that the spatial direction in which comparisons are made is an important factor, which should be treated very cautiously in future studies. The issue is problematic, as creating equivalence of within- and between-object conditions appears to necessitate an axis of symmetry in the stimulus. However, in the case of symmetry detection, it has been suggested that the perceptual inequality of symmetry axes at different orientations may be overcome by blocking stimuli by axis orientation (Wenderoth, 1994). This removes the factors of axis familiarity and uncertainty, and so may also be an effective strategy when studying part- vs. object-representation.

## 7. Conclusion

We began this paper by asking what the consequences of perceptual grouping are. Our experiments required subjects to match two spatially remote features, while the degree to which the two locations were incorporated into a common grouped object was manipulated. We find that distinct representation of spatially remote features is not generally facilitated when the features are grouped together into a coherent object—at least to the extent that it supports accurate assessment of their matching or non-matching status in our task. To the contrary, the true object-related effect may be for *between-object* facilitation in this aspect of perception. However the overriding factor in the variation of performance that we observed was object orientation. Hence whether feature representations differ when comparing within- and between-objects clearly requires more extensive investigation, always with careful consideration for the details of the stimulus. Additionally, our study suggests that however features are accessed for the purposes of comparison, this access is generally impeded rather than facilitated when features become part of larger units.

The object benefit and related phenomena now have an extensive literature. But clear conclusions about underlying processes of object formation are still lacking, perhaps in part due to the many

different methodologies that have been used to investigate it. This has resulted in conflicting and sometimes only loosely-related results being squeezed together under the general label of “object-benefits”. The choice of stimulus and task must surely be led by the question under investigation, and so an assumption of universal “within-object benefits” would seem unwise.

## Acknowledgments

This research was supported by NIH Grant No. EY15888.

## References

- Atchley, P., & Kramer, A. F. (2001). Object and space-based attentional selection in three-dimensional space. *Visual Cognition*, 8(1), 1–32.
- Barenholtz, E., & Feldman, J. (2003). Visual comparisons within and between object parts: Evidence for a single-part superiority effect. *Vision Research*, 43(15), 1655–1666.
- Baylis, G. C., & Driver, J. (1993). Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception and Performance*, 19(3), 451–470.
- Behrmann, M., Zemel, R. S., & Mozer, M. C. (1998). Object-based attention and occlusion: Evidence from normal participants and a computational model. *Journal of Experimental Psychology: Human Perception and Performance*, 24(4), 1011–1036.
- Ben-Shahar, O. (2006). Visual saliency and texture segregation without feature gradient. *Proceedings of the National Academy of Sciences of the United States of America*, 103(42), 15704–15709.
- Ben-Shahar, O., Scholl, B. J., & Zucker, S. W. (2007). Attention, segregation, and textons: Bridging the gap between object-based attention and texton-based segregation. *Vision Research*, 47(6), 845–860.
- Ben-Shahar, O., & Zucker, S. W. (2004). Sensitivity to curvatures in orientation-based texture segmentation. *Vision Research*, 44(3), 257–277.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436.
- Chen, Z. (1998). Switching attention within and between objects: The role of subjective organization. *Canadian Journal of Psychology*, 52(1), 7–16.
- Davis, G. (2001a). Between-object binding and visual attention. *Visual Cognition*, 8(3/4/5), 411–430.
- Davis, G. (2001b). Between-object binding and visual attention. *Visual Cognition*, 8(3–5), 411–430.
- Davis, G., Driver, J., Pavani, F., & Shepherd, A. (2000). Reappraising the apparent costs of attending to two separate visual objects. *Vision Research*, 40(10–12), 1323–1332.
- Davis, G., & Holmes, A. (2005). Reversal of object-based benefits in visual attention. *Visual Cognition*, 12(5), 817–846.
- Davis, G., Welch, V. L., Holmes, A., & Shepherd, A. (2001). Can attention select only a fixed number of objects at a time? *Perception*, 30(10), 1227–1248.
- Downing, C., & Pinker, S. (1985). The spatial structure of visual attention. In M. Posner & O. S. M. Marin (Eds.), *Attention and performance* (Vol. 11, pp. 171–187). London: Erlbaum.
- Driver, J., Davis, G., Russell, C., Turatto, M., & Freeman, E. (2001). Segmentation, attention and phenomenal visual objects. *Cognition*, 80(1–2), 61–95.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, 113(4), 501–517.
- Egley, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, 123(2), 161–177.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics*, 16(1), 143–149.
- Eriksen, C. W., & Hoffman, J. E. (1973). The extent of processing of noise elements during selective encoding from visual displays. *Perception and Psychophysics*, 14(1), 155–160.
- Feldman, J. (2007). Formation of visual “objects” in the early computation of spatial relations. *Perception and Psychophysics*, 69(5), 816–827.
- Goldsmith, M., & Yeari, M. (2003). Modulation of object-based attention by spatial focus under endogenous and exogenous orienting. *Journal of Experimental Psychology: Human Perception and Performance*, 29(5), 897–918.
- Koning, A., & Wagemans, J. (2009). Detection of symmetry and repetition in one and two objects. Structures versus strategies. *Journal of Experimental Psychology*, 56(1), 5–17.
- Malania, M., Herzog, M. H., & Westheimer, G. (2007). Grouping of contextual elements that affect vernier thresholds. *Journal of Vision*, 7(2), 1–7.
- Mapelli, D., Cherubini, P., & Umiltà, C. (2002). Attending to objects: Costs or benefits? *Acta Psychologica (Amsterdam)*, 109(1), 57–74.
- Moore, C. M., Yantis, S., & Vaughan, B. (1998). Object-based visual selection: Evidence from perceptual completion. *Psychological Science*, 9(2), 104–110.
- Nothdurft, H. C. (1992). Feature analysis and the role of similarity in preattentive vision. *Perception and Psychophysics*, 52(4), 355–375.
- Pashler, H. (1990). Coordinate frame for symmetry detection and object recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 16(1), 150–163.

- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437–442.
- Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology*, 109(2), 160–174.
- Rensink, R. A., & Enns, J. T. (1998). Early completion of occluded objects. *Vision Research*, 38(15–16), 2489–2505.
- Ringach, D. L., & Shapley, R. (1996). Spatial and temporal properties of illusory contours and amodal boundary completion. *Vision Research*, 36(19), 3037–3050.
- Smith, J. E. (1982). Simple algorithms for M-alternative forced-choice calculations. *Perception and Psychophysics*, 31(1), 95–96.
- Vecera, S. P., Behrmann, M., & McGoldrick, J. (2000). Selective attention to the parts of an object. *Psychon Bull Rev*, 7(2), 301–308.
- Wagemans, J. (1997). Characteristics and models of human symmetry detection. *Trends in Cognitive Sciences*, 1(9), 346–352.
- Watson, S. E., & Kramer, A. F. (1999). Object-based visual selective attention and perceptual organization. *Perception and Psychophysics*, 61(1), 31–49.
- Wenderoth, P. (1994). The salience of vertical symmetry. *Perception*, 23(2), 221–236.
- Zemel, R. S., Behrmann, M., Mozer, M. C., & Bavelier, D. (2002). Experience-dependent perceptual grouping and object-based attention. *Journal of Experimental Psychology: Human Perception and Performance*, 28(1), 202–217.