Vision Research 49 (2009) 2790–2799

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

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

S.J. Harrison [a,*], J. Feldman [b]

[a] SUNY State College of Optometry, Vision Sciences, 33 West 42nd Street, New York, NY 10036, United States
[b] 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.

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 deployment 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 (S.J. Harrison).

0042-6989/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2009.08.014

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-