Vision Research 50 (2010) 229–241

Contents lists available at ScienceDirect

Vision Research

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

# Exploring the mechanisms underlying surface-based stimulus selection

Gene R. Stoner *, Georgina Blanc

*Vision Center Laboratory, The Salk Institute for Biological Studies, La Jolla, CA, USA*

## ARTICLE INFO

*Article history:*
Received 10 September 2009
Received in revised form 4 November 2009

*Keywords:*
Object-based attention
Surface-based attention
Transparent motion
Biased competition
Area MT

## ABSTRACT

Valdes-Sosa, et al. (2000) introduced a transparent-motion design that provides evidence of surface-based processing of visual motion. We show that this design suffers from a motion-duration confound that admits an alternative explanation based on neuronal adaptation and competition. We tested this explanation by reversing the relationship between motion duration and which perceptual surface was "cued". We also examined the role of color duration. Our findings support the surface-based account and, more specifically, demonstrate that this type of surface-based selection involves selective spatial processing at the scale of the texture elements that define the transparent surfaces.

## 1. Introduction

Visual processing is selective – few stimuli that impinge the retinae reach perceptual awareness and/or elicit behavioral responses. Selective processing based on location (e.g. Posner, 1980; Posner & Cohen, 1984; Treisman & Gelade, 1980) or individual features (e.g. Aine & Harter, 1986; Anllo-Vento & Hillyard, 1996) is well-established and easy to reconcile with the organization of the visual cortex into retinotopic maps and feature columns. There is growing evidence, however, that whole objects or surfaces can be selectively processed (e.g. Blaser, Pylyshyn, & Holcombe, 2000; Duncan, 1984; O'Craven, Downing, & Kanwisher, 1999; Valdes-Sosa, Cobo, and Pinilla, 2000). The mechanisms underlying such object- or surface-based selection are unclear. We have argued (e.g. Mitchell, Stoner, Fallah, & Reynolds, 2003; Mitchell, Stoner, & Reynolds, 2004; Reynolds, Alborzian, & Stoner, 2003) that the transparent-motion design offered by Valdes-Sosa et al. (2000) has provided the best evidence of object- or surface-based selection (we mostly use the later phrase hereafter) to date, but in this study we tested an alternative (i.e. non-surface-based) account of that design. Our findings are consistent with surface-based selection and shed light on the underlying mechanisms.

### 1.1. Transparent motion and surface-based attention

The transparent-motion design introduced by Valdes-Sosa et al. has been adapted to study various aspects of surface-based selection including perceptual mechanisms (Lopez, Rodriguez, & Valdes-Sosa, 2004; Mitchell et al., 2003; Reynolds et al., 2003; Rodriguez, Valdes-Sosa, & Freiwald, 2002), single-unit correlates in the non-human primate (Fallah, Stoner, & Reynolds, 2007), event-related potentials (ERPs) in humans (Khoe, Mitchell, Reynolds, & Hillyard, 2005; Pinilla, Cobo, Torres, & Valdes-Sosa, 2001; Valdes-Sosa, Bobes, Rodriguez, & Pinilla, 1998), and interactions between selective attention and binocular rivalry (Mitchell et al., 2004). While stimulus and behavioral details in the above studies varied somewhat, in all of those studies (except for Fallah et al. in which there was no behavioral component) subjects were asked to judge brief translations of one of two superimposed dot fields, which (excepting the brief translations) rotated in opposite directions (i.e. clockwise and counterclockwise). It has been found that translations of dot fields that are "cued" (endogenously or exogenously) are judged accurately relative to translations of the other ("uncued") dot field. This design is meant to rule out both spatial and feature-based selection as an explanation for the performance bias. Spatial selection (at least at a coarse scale) is ruled out by spatial superimposition of the two dot fields. Motion-based selection is ruled out since the direction of the translation is unpredictable. By removing the color differences between the two dot fields, Mitchell et al. (2003) have also demonstrated that the performance bias is not color-based. Instead, these results have been taken as evidence of surface-based selection whereby the successive motions (i.e. the rotation followed by the translation) of a cued perceptual surface are preferentially processed relative to motions of an uncued surface.

### 1.2. The motion-duration confound

In considering how preferential processing of a cued dot field's rotation direction might give rise to preferential processing of that dot field's translation direction, we realized that previous designs

---

* Corresponding author.
*E-mail address:* gene@salk.edu (G.R. Stoner).

0042-6989/$ - see front matter
doi:10.1016/j.visres.2009.11.015

suffered from a motion-duration confound that admitted an explanation not requiring surface-based selection. This confound applies to both the original "two-translation" design devised by Valdes-Sosa et al. (2000) and to the "delayed-onset" design introduced by Reynolds et al. (2003) in which there is only a single translation. This confound provides a challenge to the interpretation of the numerous studies that have used these designs (cited above). We first illustrate this confound with the delayed-onset design, since it is more obvious in that design and because we use the delayed-onset design in this study.

Fig. 1 shows two complementary depictions of the delayed-onset design. Fig. 1A illustrates the appearance of the transparent-motion stimuli as two counter-rotating perceptual surfaces. The depiction in Fig. 1B, conversely, explicitly shows the relative duration of each type of motion, thereby more clearly revealing the motion-duration confound. In this later depiction, the dots of the two dot fields are distinguished by line style (dashed or solid), dot color is given by line color, and type of motion (i.e. clockwise, counterclockwise, or translation) is given by vertical line placement. The difference in the onset times of the two dot fields yields the motion-duration confound (gray region): translations of the "*cued*" (i.e. delayed) dot field occur in the presence of the *older* rotation direction, whereas translations of the "*uncued*" (i.e. non-delayed) dot field occur in the presence of the *newer* rotation direction.

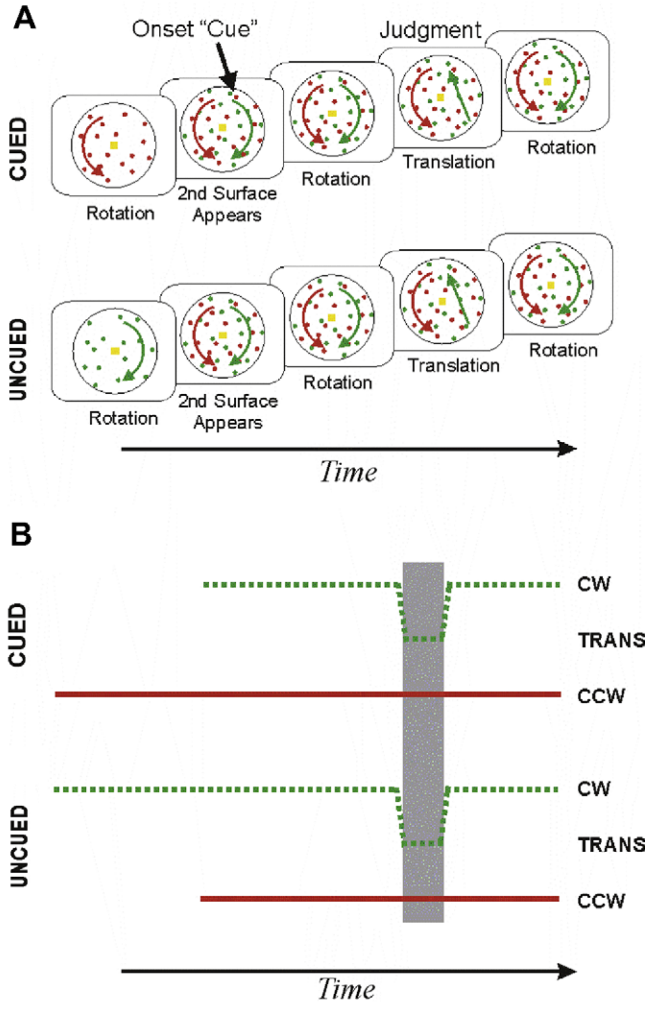

**Fig. 1.** Delayed-onset design. (A) Conventional depiction. Two superimposed dot fields rotate in opposite directions (about a central fixation target) yielding a perception of two transparent surfaces. One rotating dot field appears first followed by the "delayed" dot field. Subsequently, either the delayed (i.e. "cued") or non-delayed (i.e. "uncued") dot field translates briefly. After translation, both dot fields rotate. Subjects report the direction of the translation. Translations of the cued dot field are judged more accurately than translations of the uncued dot field. (B) Feature-based depiction. Dot fields are distinguished by line style (dashed or solid), dot field color is given by line color, and type of motion (i.e. clockwise, counterclockwise, or translation) is given by vertical line placement. The onset differences in this design result in "cued" translations occurring in the presence of the older rotation direction and "uncued" translations occurring in the presence of the newer rotation direction (gray region).

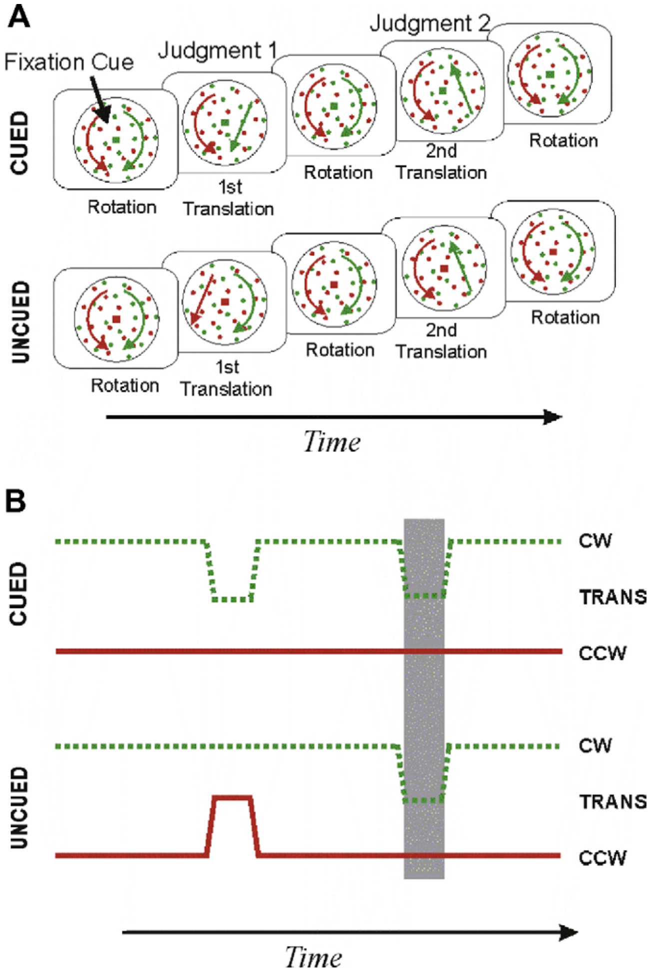

**Fig. 2.** Two-translation task of Valdes-Sosa et al. (2000). (A) Conventional depiction. Fixation-point color (green as in upper panels or red as in lower panels) indicates which surface translates first. Following a period of rotation, the cued dot field translates briefly, while the other field continues to rotate. The dot fields then continue to rotate for a variable delay, at which point one dot field, chosen randomly, translates briefly. After this second translation, both surfaces rotate. Observers report the direction of each translation. It was found that the first translation is judged accurately, but the *second* translation is only judged accurately if it is of the same dot field that translated first. (B) Feature-based depiction. Conventions are same as in Fig 1B. The first translation has been proposed to exogenously cue attention to the translating dot field (Reynolds et al., 2003). This first translation also leads to a motion-duration confound (gray region): cued second translations occur in the presence of the older (i.e. non-interrupted) rotation and uncued second translations occur in the presence of the newer (i.e. interrupted) rotation.

The original design of Valdes-Sosa et al. (2000) had two successive translations (Fig. 2A) and, as described below, it also suffers from a motion-duration confound. In this design, fixation target color (red or green)[1] cues subjects as to which dot field translates

---

[1] For interpretation of color in Figs. 1–8, the reader is referred to the web version of this article.