



Segmentation, attention and phenomenal visual objects

Jon Driver^{a,*}, Greg Davis^b, Charlotte Russell^a,
Massimo Turatto^c, Elliot Freeman^a

^a*Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1N 3AR, UK*

^b*Birkbeck College, London, UK*

^c*University of Padova, Padova, Italy*

Received 8 March 2000; accepted 17 November 2000

Abstract

Issues concerning selective attention provoke new questions about visual segmentation, and vice-versa. We illustrate this by describing our recent work on grouping under conditions of inattention, on change blindness for background events and the residual processing of undetected background changes, on modal versus amodal completion in visual search, and the differential effects of these two forms of completion on attentional processes, and on attentional modulation of lateral interactions thought to arise in early visual cortex. Many of these results indicate that segmentation processes substantially constrain attentional processes, but the reverse influence is also apparent, suggesting an interactive architecture. We discuss how the ‘proto-objects’ revealed by studies of segmentation and attention (i.e. the segmented perceptual units which constrain selectivity) may relate to other object-based notions in cognitive science, and we wrestle with their relation to phenomenal visual awareness. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Segmentation; Attention Phenomenal visual objects

1. Introduction

The recent literature on so-called ‘object-based’ visual attention represents fertile new cross-talk between two traditionally separate research fields, one concerning visual segmentation and grouping processes, and the other concerning selective

* Corresponding author. Fax: +44-207-813-2835.

E-mail address: j.driver@ucl.ac.uk (J. Driver).

attention. We shall not provide an exhaustive summary of research on object-based visual attention here (for reviews see Driver & Baylis, 1998; Kanwisher & Driver, 1992; Scholl, 2001). Instead, we shall use selected examples from our own recent work, seeking to illustrate how research on visual segmentation and on selective attention can be complementary. When the two fields are considered together, new questions arise that neither research tradition would otherwise have considered. Moreover, methods from the two fields can be fruitfully used in combination.

2. Image segmentation

The ‘objects’ with which we will be concerned are segmented perceptual units. These might more properly be called ‘proto-objects’, since they need not correspond exactly with conceptual or recognizable objects. Instead, they reflect the visual system’s segmentation of current visual input into candidate objects (i.e. grouping together those parts of the retinal input which are likely to correspond to parts of the same object in the real world, separately from those which are likely to belong to other objects). There is a long research tradition on visual segmentation, dating back to seminal work by the Gestaltists (e.g. Wertheimer, 1923). Image segmentation is traditionally viewed as a fundamental problem that the visual system must solve quite early (e.g. Marr, 1982; Nakayama, He, & Shimojo, 1995; Nakayama, Shimojo, & Silverman, 1989; Neisser, 1967; Rock, 1975), before proceeding to further processes such as object recognition (although see Peterson, 1994). Rock, Nijhawan, Palmer, and Tudor (1992, p. 779) state that “it is widely acknowledged that a precondition for the perception of the world of objects and events is an early process of organization in the visual system”. Thus, a standard idea is that real-world objects can only be recognized appropriately if fragments of the retinal image are bundled together appropriately, before matching against long-term memories of known visual objects. For instance, when faced with an animal partly occluded behind a tree, the retinal fragments corresponding to the animal should ideally be bundled together for matching against representations of known objects in long-term memory, separately from those fragments corresponding to the tree. Trying to recognize a bundle of features which combined part of the animal with part of the tree would be unlikely to lead to an appropriate match with a single known object! From this perspective, appropriate image segmentation seems to be an essential first step, and so it has traditionally been attributed to quite early visual processes (although many current accounts refer to it as reflecting ‘mid-level’ vision, with this term implying an intermediate stage between initial description of raw image statistics, and subsequent matching against a knowledge base).

The starting point for much research on image segmentation was phenomenology, as in the Gestaltists’ many demonstrations of grouping (e.g. Wertheimer, 1923) or figure-ground segmentation (e.g. Rubin, 1921). These demonstrations, reproduced in countless textbooks since, rely entirely on the phenomenology of the observer for their appeal. For instance, the reader may be asked to inspect an array of dots in order to confirm the introspection that subsets of these dots seem to go together (as

when organized into columns or rows by principles such as proximity or similarity). Because of the theoretical perspective outlined above (on which image segmentation is typically considered as an essential first step for the visual system), it has commonly been assumed that the phenomenology of Gestalt grouping displays somehow affords a privileged view of quite early visual processes. However, it should be remembered that since the classic Gestalt demonstrations rely entirely on phenomenology, they may presumably tap into many stages of vision, or even reflect an end-product. Phenomenology is certainly a useful starting point for much research into perception, and the phenomenal appearance of visual displays remains one of the things that researchers must explain (and which we will wrestle with here). However, phenomenology can be a crude or even inappropriate tool for addressing some questions, especially those of the sort thrown up by the literature on selective attention.

3. Selective attention

Just as for image segmentation, there is a long but entirely separate research tradition on selective attention, our ability to process incoming information selectively. Again phenomenology provides one starting point, as many of the fundamental issues about selective attention can be prompted simply by phenomenal demonstrations that involve two streams of information (traditional examples include two concurrent spoken messages in audition, or two superimposed pictures in vision, e.g. Cherry, 1953; Neisser & Becklen, 1975; Rock & Gutman, 1981). People can simply be asked to confirm the introspection that they may concentrate on one or the other of the two streams at will, even when both streams are equally clear at the ear or eye. Phenomenologically, there is a much richer awareness of the attended stream. Moreover, objective tests confirm that there is little or no explicit knowledge of the unattended stream (e.g. Cherry, 1953; Mack & Rock, 1998; Neisser & Becklen, 1975; Rock & Gutman, 1981). Several issues are immediately prompted by the phenomenology of attention in such settings: for instance, on what basis can we attend to one stream versus the other, and what are the consequences for the two streams? Such questions have now been addressed with many objective measures.

Ever since the seminal account of Broadbent (1958), it has been traditional to contrast two hypothetical stages of perceptual processing, thought to be qualitatively distinct, for such situations. ‘Preattentive’ processing is traditionally thought to be automatic, parallel, and to extract relatively simple stimulus properties. This stage would operate for both streams of input in the examples above, regardless of which one the person attended at will. This ‘preattentive’ stage by definition is traditionally thought to precede a subsequent ‘attentive’ stage, with the latter by definition depending on the attentional state of the observer. The attentive stage would only operate for one of the two streams in the examples above, and has been variously characterized as serial or of limited capacity, and as extracting more complex

stimulus properties (e.g. Broadbent, 1958; Treisman & Gelade, 1980), although there have been many subsequent qualifications of this (e.g. see Pashler, 1998).

Space limits preclude a full consideration of the substantial and complex literature motivated by this preattentive/attentive dichotomy. Suffice to say that while it remains influential, and has proved a useful heuristic for generating research, the dichotomy is clearly a gross oversimplification of biological reality. For instance, recent neuroscience studies of selective attention have shown, both with single-cell recording in animals (e.g. Desimone & Duncan, 1995) and with functional imaging in people (e.g. Kastner, Pinsk, De-Weerd, Desimone, & Ungerleider, 1999), that the attentional state of the observer can modulate *many* levels of perceptual processing, rather than kicking in only at one specific ‘attentive’ stage, following an initial ‘preattentive’ stage. It does remain roughly true that later stages of vision can show greater attentional modulation than earlier stages (e.g. Kastner et al., 1999), but even primary visual cortex can be modulated to some extent (e.g. Somers, Dale, Seiffert, & Tootell, 1999). Our own view is that ‘attention’ is best thought of as the umbrella-term for a general topic, subsuming a host of questions about selective processing, *not* as a single explanatory process, which affects just a single stage of perception. If we occasionally slip into referring to ‘attention’ as if this were a singular process, then we do so merely for expository convenience.

4. Attention and image segmentation

The kind of questions which an attention researcher might have about visual segmentation include the following. Does such-and-such a segmentation process occur ‘automatically’, regardless of the attentional state of an observer? Does it still proceed if the observer attends elsewhere to perform some other task, of low or high demand? Can it be strengthened by attending appropriately? Can it be weakened by attending in a different way? Does it constrain what an observer can efficiently select for judgement from a visual display (e.g. with better performance when judging two visual attributes if these two attributes are segmented together, rather than apart)? Does it constrain what an observer can ignore (e.g. are distractors more distracting when grouped together with a target)? These various questions differ in their specifics, but many relate to whether the visual task that an observer is currently performing, or is about to perform, can modify how an image gets segmented (i.e. can attention affect segmentation?). Conversely, many concern whether image segmentation constrains what can be efficiently selected in a particular visual task (i.e. does segmentation affect attention?).

The Gestaltists tended to view their grouping phenomena as an illustration of the perceiver imposing a seemingly arbitrary (albeit systematic) organization upon the stimuli (e.g. Wertheimer, 1923). Nowadays the more typical view of such grouping demonstrations would be that they reflect non-arbitrary properties within the stimuli (similarity, etc.), which the visual system exploits heuristically because these properties are likely to reflect divisions into distinct objects in the real world. The attentional questions about segmentation, as outlined above, provide a new frame-

work for asking whether an observer can, at will, make a contribution to how a given display will be segmented, or whether instead, the influence works mainly in the opposite direction, with image segmentation constraining what an observer can attend at will.

The traditional view from attention research, within the preattentive/attentive dichotomy, emphasizes the latter direction of influence, with segmentation processes thought to arise preattentively, to yield ‘proto-objects’ (or feature-bundles) which may then be attended in turn for further processing (e.g. Driver & Baylis, 1998; Duncan, 1984; Neisser, 1967; Treisman, 1986). Palmer and Rock (1994, p. 37) went so far as to suggest that this may be a logical necessity, arguing that “...logic dictates that some amount of visual organization must occur at an early stage in visual processing and that it must occur preattentively. As Neisser (1967) and Treisman (1986) have put it, discrete perceptual elements of some sort must be present to serve as candidates for further element-based processing. Only after such elements are present can we attend selectively to one or another.”

Note that although somewhat different language is used in this quotation (now gesturing towards the preattentive/attentive dichotomy), this line of argument accords with the traditional view within the segmentation literature as described earlier, on which segmentation is considered an essential first step of organizing the input into feature-bundles, before further processing of each bundle. Note also that many questions are begged when using attentional language to convey this general story. In particular, exactly which kinds of segmentation precede selective processing (note the ‘some amount’ in the above quotation), and for what kinds of selective processing, given that rather than modulating vision at just a single discrete stage of selectivity, attentional factors are now known to modulate visual processing at many stages (cf. the neuroscience evidence above)?

Despite these uncertainties, some of us have long been sympathetic (e.g. Driver & Baylis, 1989, 1998) to the general view that some forms of segmentation can precede attentional selectivity. In part, we sympathize with this because we consider it to be broadly consistent with the actual *phenomenology* of Gestalt grouping displays (see Driver & Baylis, 1998). As explained earlier, the phenomenology of Gestalt grouping is often naively taken to tap directly into early segmentation processes, but we suggest that it may reflect attentional processes also. Indeed, the introspection that a subset of dots in a Gestalt display ‘belongs together’ may arise precisely because when trying to pick out one of these dots, you tend to pick out also those dots which are grouped with it. In other words, they may seem to ‘belong together’ precisely because you tend to attend them together. This would fit the general story (Driver & Baylis, 1998; Neisser, 1967) that initial grouping processes constrain the perceptual units that can be attended. Thus, it may be hard to attend separate groups at the same time, and also hard to attend a single element within a group, without also selecting the other elements that are grouped with it. Much evidence from objective performance measures (see Driver & Baylis, 1998, for review) now supports this general account, in addition to it being broadly consistent with the phenomenology of grouping displays.

On the above perspective, phenomenology may reveal some useful facts about

attentional issues, not merely about segmentation processes. Nevertheless, it is virtually impossible to address some of the more specific ‘attentional’ issues about segmentation that we raised above, using phenomenology alone. In particular, it is extremely awkward to assess whether a particular segmentation process still takes place when the observer does *not* wish it to do so, using mere phenomenology. The problem here is similar to that induced by asking someone *not* to think about spotted pink elephants! Asking people to introspect about whether a particular grouping phenomenon still arises, when they attend elsewhere, runs into the obvious problem that merely asking about the grouping phenomenon is likely to induce attention to the associated stimuli. This is a familiar methodological problem in the selective attention literature (i.e. how to measure unattended processing without turning it into attended processing) and many standard solutions have been developed (e.g. see Driver, *in press*; Pashler, 1998, for reviews), as described later. However, these have only recently been applied to attentional issues concerning segmentation processes.

The late Irvin Rock was one of the first researchers to note explicitly that the way a visual display gets segmented may depend on the question asked of an observer, and thus on what they ‘attend’ in this sense (e.g. Rock, 1975, 1983). He realized that the very act of asking observers to confirm an introspection (e.g. that certain arrays of dots appear to group into rows or columns, as in the standard Gestalt demonstrations) may well change how particular visual stimuli are processed. Together with Arian Mack and their collaborators, he developed a seemingly straightforward method of measuring whether traditional Gestalt grouping still takes place under conditions of true ‘inattention’. In other words, under conditions where the observer has no specific intention with respect to the grouping display, being engaged with a different demanding visual task instead, which involves a completely separate visual stimulus at another location (e.g. Mack & Rock, 1998; Mack, Tang, Tuma, Kahn, & Rock, 1992; Rock, Linnett, Grant, & Mack, 1992). From the perspective of attention research, we think their studies posed exactly the right kind of question about segmentation processes. But as discussed below, we are less satisfied with the methods they used, and the conclusions reached. Nevertheless, some of their conclusions from their early studies were certainly provocative. Indeed, they appear to be in direct opposition to the position some of us had adopted in the past. Whereas Baylis and Driver (1992) argued that “visual attention is directed to groups derived from a preattentive segmentation of the scene according to Gestalt principles”, in the same year Mack et al. (1992, p. 498) concluded that “no perception of either texture segregation of Gestalt grouping” takes place for unattended stimuli!

5. Gestalt grouping without attention: ‘inattentional blindness’ reconsidered

Mack et al. (1992) and Rock, Linnett et al. (1992) correctly noted that some of the paradigms traditionally thought to tap ‘preattentive’ processing may in fact involve diffuse attention, plus some intention by the observer to see particular visual properties (see also Joseph, Chun, & Nakayama, 1996). For instance, in visual search

studies (e.g. Treisman, 1986), observers typically look for a specific target, and know that all locations on the screen may be task-relevant. Mack et al. (1992) suggested that in order to test whether segmentation processes can truly operate without attention, one needs a situation where the observer has no intention to perform the segmentation, being engaged in some other demanding visual task for another stimulus at a different location (see Joseph et al., 1996; Lavie, 1995, for arguments that the level of demand may be critical).

Mack et al. (1992) presented subjects with displays in which a central cross appeared, with a Gestalt grouping display (or texture segmentation display) as its surrounding background. The instructed task concerned the central cross (was its horizontal or vertical limb longer?), so the background display was initially irrelevant to the task. But after several trials, observers were suddenly asked surprise questions about the immediately preceding background display, concerning how it appeared to group. This actually represents one of the oldest ways to try and measure unattended processing, by asking surprise retrospective questions (e.g. Broadbent, 1958; Cherry, 1953). The basic finding, across several experiments (see Mack & Rock, 1998, for recent extensions), was that observers seemed to know surprisingly little about the segmentation of the background display, in response to the surprise question. From this basic result, plus various control observations, Mack et al. (1992) reached their radical conclusion that no Gestalt grouping takes place without attention.

The results of Mack et al. (1992) accord with many classic demonstrations that little is known about unattended stimuli on subsequent surprise questioning (e.g. Broadbent, 1958; Cherry, 1953). In the visual domain, such results are now often discussed under the general heading of ‘inattention blindness’ (e.g. Mack & Rock, 1998). However, there is a standard objection to findings of little knowledge for unattended stimuli on surprise retrospective questioning, which dates back to the dawn of attention research (e.g. Cherry, 1953). The poor knowledge shown may reflect poor explicit memory, rather than the absence of on-line processing when the unattended stimulus was presented (i.e. inattentional ‘amnesia’ rather than inattentional ‘blindness’; Holender, 1986; Wolfe, 1999). Although Mack et al. (1992) and also Mack and Rock (1998) do discuss this possibility, and attempt to reject it, the problem is intrinsic to the method of surprise retrospective questioning. But posing the direct questions in advance, or at the time of presentation, would clearly suffer from the ‘spotted-pink-elephant’ problem, being likely to make a supposedly unattended stimulus become task-relevant and thus attended. The traditional solution to this methodological problem, within mainstream attention research (e.g. Lewis, 1970; Stroop, 1935), has been to devise *indirect* on-line measures of unattended processing, which do not require the subject to explicitly judge or respond to the unattended information. In prototypical form, one measures instead whether unattended information can influence responses to attended information.

Moore and Egeth (1997) applied this general approach to the issue of whether Gestalt grouping can arise under ‘inattention’. They took advantage of some standard geometric illusions (i.e. the Ponzo and Muller-Lyer illusions), manipulating whether background dots could be grouped (by similarity in contrast polarity) into

appropriate inducers for biasing the perception of the horizontal line-length of two target lines via geometric illusions. In a task of judging the relative length of the two target lines, they found that the background organization could produce the conventional illusions for the target lines, even when subjects were unable to report this background organization retrospectively, as tested by the Mack et al. (1992) surprise retrospective questions. The implication is that some degree of background grouping by similarity in contrast polarity can still take place, even when the background is task-irrelevant.

This is an ingenious study, but the implications may be restricted in several respects. First, the effective ‘grouping’ of the very proximal dots by common contrast polarity may have been simply due to blurring in low spatial frequency channels, and so would represent only a very crude form of grouping at best. Indeed, it has often been suggested that the Ponzo and Muller-Lyer illusions can have a low spatial frequency basis (e.g. Rock, 1983). Second, many of the dots inducing the foreground illusion were very close to the target lines, and so a sceptic might argue that their effectiveness could have been due to their falling within an attended spatial area.

We (Driver, Russell, & Howlett, unpublished data) have recently developed a new method which can also provide an indirect on-line measure of whether grouping may still take place under ‘inattention’, but while avoiding these limitations. Unlike the Moore and Egeth (1997) study, the method is not tied to the requirements of a particular illusion. As in Mack et al. (1992), subjects are presented with displays that each comprise a central target (now a small square made up of random black or white pixels) plus a surrounding background Gestalt grouping display of dots (which can be arranged into regular columns or rows by common colour, or can have a random organization by the two equiluminant colours used within each display). Each trial comprises two successive displays of this kind (see Fig. 1), separated by 150 ms. The task is to judge as rapidly as possible whether the two successive central target squares are the same or different (when different, only a single pixel changes). Independently of whether these successive targets are the same or different, we also manipulate whether the background has the same organization between two successive displays (e.g. Fig. 1A) or a different organization (e.g. Fig. 1B). The two colours used for the dots in the two successive backgrounds always change (e.g. red and green dots in the first display, but yellow and blue in the second display), so that same versus different background organization can be unconfounded from same versus different colours for each dot between the two successive displays. We measure whether reaction time (RT) and errors in the same/different task for the central square are influenced by the background organization being the same or different over the pair of successive displays. Then, at the end of the experiment, we test for explicit knowledge of the immediately preceding background display, using the surprise retrospective questioning method of Mack et al. (1992).

In a series of studies of this kind, we have repeatedly found that the surprise retrospective questions at the end replicate the findings of Mack et al. (1992), from which they had concluded that no Gestalt grouping takes place under inattention. However, the RT and error data from the on-line same/different task (performed on

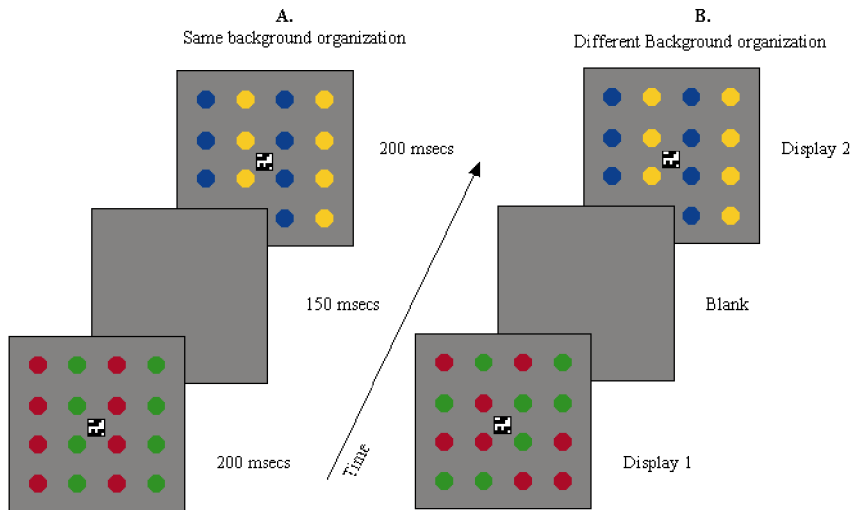


Fig. 1. Two examples of the display sequence for a single trial in the studies of Driver et al. (unpublished data). The task is to decide whether the central square matrix, comprising randomly black or white pixels, is the same or different for Display 1 and Display 2 (it is the same in both the example sequences shown). The surrounding grid of dots can be organized by common colour into columns, or into rows (not shown), or pseudo-randomly (e.g. see Display 1 at bottom right of B). This background organization can change (as in B) or be the same (as in A) across Display 1 and Display 2 within a trial, independently of whether the central target matrix changes or not. The two colours used for the background dots always changed between Display 1 and Display 2 (in the illustration, from red and green to yellow and blue), so that colour change per se could be disentangled from changes in grouping by colour.

the central square matrix) suggest that some grouping of the background dots had in fact taken place. Subjects are faster and/or more accurate with their target-different judgements if the background organization also changes (e.g. from random into columnar; see Fig. 1B), and target-same judgements can also be faster and/or more accurate if the background organization stays the same. We find such results even if the critical change in background organization depends only on dots which are relatively far away from the central square target (e.g. those in the outer locations).

These data suggest that some degree of grouping by common colour can arise under conditions of 'inattention', as Mack et al. (1992) define this, contrary to the strong conclusion drawn in their original paper. In our experimental situation, the effective background grouping is irrelevant to the demanding central target task. It arises in a dimension (i.e. equiluminant colour) which is different from the relevant properties for the central task (which concerns luminance, and possibly shape). The effective background grouping can involve only elements which are some distance away from the central target. Finally, subjects cannot explicitly report the background organization, using Mack et al.'s (1992) surprise question method. This situation would therefore seem to fit Mack et al.'s definition of true 'inattention', and yet some evidence for Gestalt grouping is still found.

On the other hand, the nature of this evidence tells us little about phenomenology, which has traditionally been of considerable (often primary) interest in Gestalt demonstrations, as discussed above. The RT and error pattern we find in the same/different task implies that the background displays were grouped by common colour to some extent, despite the conditions being set up to induce ‘inattention’ for them. However, these data certainly do not imply that the subjects necessarily perceived the background grouping consciously, in the manner of a standard Gestalt demonstration. Indeed, on a Mack et al. (1992) interpretation of the results for the surprise direct questioning, one would conclude that subjects were not aware of the background grouping, even though it could evidently still influence their target judgements. Our results might therefore be reconciled with the idea of ‘inattentional blindness’ for Gestalt grouping, but only as regards conscious vision (see Mack & Rock, 1998), not in the sense of no Gestalt grouping taking place whatsoever, as Mack et al. (1992) had originally claimed. However, the retrospective nature of Mack et al.’s surprise questioning method can again be criticized, this time as an index of visual awareness, rather than of visual processing in general (i.e. responses to a surprise question after a display may underestimate what is consciously seen during a display).

6. ‘Change blindness’ reconsidered

The method depicted in Fig. 1 may relate to another paradigm which, like ‘inattentional blindness’, has recently been taken as evidence that our representation of the visual world can be surprisingly sparse. We refer here to the recent demonstrations of ‘change blindness’ in the flicker paradigm introduced by Rensink, O’Regan, and Clark (1997) (see Simons, 2000; Simons & Levin, 1997, for reviews). In the original version of this, subjects are presented with complex images of natural scenes. Successive displays show the same scene, separated by a substantial transient (a brief ‘flicker’) in which the whole screen goes blank, thus changing luminance and other properties everywhere. The subjects’ task is typically to report whether anything has changed in the scene when it returns after the flicker, or to press a button when they detect this change. Often, two versions of the same scene, differing in one aspect (e.g. deletion of an object), may be cycled repeatedly with a flicker separating them, until the change to the scene is reported.

The basic finding (Rensink et al., 1997) is that it can take a remarkably long time (i.e. many cycles) before the change to the scene is noticed, even for quite substantial changes (e.g. a large object disappearing and reappearing). Of course this depends on the flicker which separates the successive versions of the scene, as this provides a large global transient to obscure the local transient that the change to the scene would otherwise produce. Nevertheless, when the change is detected, subjects can hardly believe it went undetected for so long; once detected, the change seems absurdly salient across subsequent repeated cycles. We (Russell & Driver, unpublished data) have confirmed that similar phenomena remain when eye-position

is held fixed, and that an otherwise undetectable change becomes noticeable as soon as its location is cued, even without saccades.

Such ‘change blindness’ has been interpreted as strong evidence for surprisingly sparse internal representations of visual scenes, similarly to ‘inattention blindness’ claims, but now primarily concerning the issue of what is brought forward across successive displays, rather than what is extracted from a single display. It has been interpreted (see O’Regan, Rensink, & Clark, 1999; Simons & Levin, 1997) as suggesting that standard introspections of perceiving a rich visual world are an ‘illusion’. We agree that change blindness provides a particularly compelling example of the influence that attentional factors can have on phenomenal visual experience, as when a previously undetectable change becomes overwhelmingly salient once its location is cued. However, we are more sceptical about some of the radical conclusions that have been drawn from change blindness. In particular, one may question whether it really implies exceptionally sparse representation, or a total failure to bring visual information forward across successive samples. The method depicted in Fig. 1 could be usefully adapted to test whether various undetected changes in the flicker paradigm might still be extracted unconsciously by the visual system (see Fernandez-Duque & Thornton, 2000; Rensink, 2000; Smilek, Eastwood, & Merikle, 2000). This could be revealed by any effects of background changes on RT or accuracy for same/different judgements to central targets, similar to those we have already observed for changes in background grouping.

We have already found some initial evidence (Turatto, Russell, & Driver, unpublished data) that unseen changes may nevertheless be extracted in the flicker paradigm. We used displays (see examples in Fig. 2) comprising dots on a background of vertical stripes. As in the standard flicker paradigm (e.g. Rensink et al., 1997), successive displays were separated by an intervening blank (of intermediate grey), and subjects had to report simply whether anything changed (and if so what) between successive displays. In fact, only two types of change were possible. Either one or more of the grey dots could change luminance (from dark grey to light grey, or vice-versa; e.g. Fig. 2A), or all the background stripes could change their luminance (from dark to light, or vice-versa; e.g. Fig. 2B). Even though the latter change was physically much more substantial (in terms of actual luminance change and spatial area affected) than any change to the dots within the same luminance domain, we found that subjects typically reported virtually all the dot changes, yet none of the background changes.

This appears to confirm that change blindness in the flicker paradigm is indeed more pronounced for background than foreground changes (see Rensink et al., 1997), even when the background change is physically stronger. As soon as the possibility of background changes was pointed out to subjects, they actually became faster at detecting those than for changes to the foreground dots, consistent with the larger physical signal for the background stripes. The initial failure to detect any background changes may relate both to a default tendency to attend foreground items, and also to the Gestalt psychologists’ proposals that foregrounds are more phenomenally compelling than backgrounds (e.g. Rubin, 1921). Indeed, attentional factors may conceivably provide part of the explanation for the latter phenomenology.

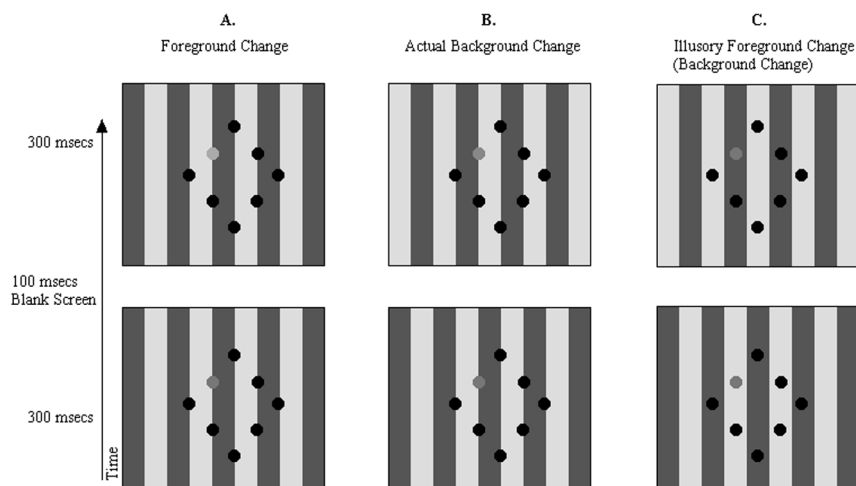


Fig. 2. Three examples of two successive displays, which were separated by a brief blank screen, in the change blindness study of Turatto et al. (unpublished data). Time runs from bottom to top. (A) The grey circle in the second display brightened; such changes were reported by naive observers. (B) The background stripes reversed contrast; such changes were not noticed until pointed out to naive observers, even though the luminance change was larger than for the circle in (A). Note that in (A) and (B), each circle falls half on a light stripe and half on a dark stripe. (C) Each circle now falls on a single stripe. When the background stripes reversed contrast, illusory changes in the brightness of the grey circle (due to simultaneous contrast with the surrounding stripe) were reported, though the background change itself went unseen again.

The second question we addressed in our study was whether unseen background changes might nevertheless still be coded by the visual system. We addressed this by exploiting a well known simultaneous contrast illusion (see Rock, 1975). When a grey dot appears with a darker stripe as its surrounding background, it appears lighter than when appearing against a light background. Some of our displays were arranged such that the background change altered whether a given grey dot appeared on a dark stripe or on a light stripe between the two successive displays (see Fig. 2C). In such cases, naive observers always reported any change as concerning the brightness of the foreground grey dot, not as the alteration of the background stripes which took place physically. They did so even though the stripes went from dark to light (or vice-versa) physically, while the dot only underwent an illusory change in lightness that was entirely due to the change in the background stripes. This shows that unseen background changes can produce illusory phenomenal foreground changes in the flicker paradigm, which would seem to indicate that some of the unseen background changes must still be extracted (and the changed properties ‘brought forward’) by the visual system. We are currently adapting this paradigm to test whether several particular Gestalt grouping processes can operate in the guise of unseen background changes that produce foreground illusions.

7. Studying specific forms of segmentation in relation to attention: modal and amodal completion

The extensive literature on ‘object-based’ attention (see Driver & Baylis, 1998; Kanwisher & Driver, 1992; Scholl, 2001, for reviews) has tended to lump all image segmentation processes together. However, there is no guarantee that all forms of image segmentation (e.g. by different static Gestalt factors, by stereo cues, by dynamic motion cues, etc.) will have exactly the same relation to attentional processes. A more productive strategy may be to examine individual segmentation processes in particular, in relation to specific attentional issues. In this way, attention researchers might exploit many of the advances which have been made by the vision research literature in understanding specific forms of image segmentation (e.g. Nakayama et al., 1989, 1995; Palmer & Rock, 1994; Pomerantz & Kubovy, 1986).

Kanizsa subjective figures (e.g. Fig. 3A) provide one of the most intensively investigated examples of human image segmentation. Not only are they phenomenally intriguing (Kanizsa, 1979), but they are by now relatively well understood (see Leshner, 1995; Petry & Meyer, 1987). Moreover, they are often considered (e.g. Fahle & Koch, 1995) to represent a prototypical case of the visual system binding together separate elements in the image (such as the ‘pacmen’ inducers in Fig. 3A) to produce a single subjective object. They certainly illustrate several principles of image segmentation. These include edge assignment (e.g. the straight edges on the pacmen inducers in Fig. 3A,B typically get assigned to the subjective square, not to the pacmen) plus several completion processes, as follows. When a subjective figure is seen, the inducing pacmen are typically seen as ‘amodally’ complete circles (i.e. complete circles that are partially occluded; see Fig. 3C), while the subjective square is ‘modally’ completed (i.e. with accompanying illusory colour and/or brightness in the completed region). Such figures thus also provide compelling examples of subjective ‘filling-in’ of illusory colour and brightness (as for the seemingly grey and transparent regions near the central cross in Fig. 3B, which are actually white).

There have been many experimental studies of such subjective figures. One enduring issue has been whether they are best thought of as reflecting the operation of relatively ‘low-level’ processes (e.g. blurring at low-spatial frequencies), or of much ‘higher-level’ processes (e.g. the inference of a superimposed figure). Space constraints preclude a full review of this debate. Suffice to say that there is now substantial evidence that low-level properties of the stimulus can strongly affect the subjective strength of the subjective figure, but also that seemingly higher-level factors, such as set or knowledge, can affect the percept to some extent (see Leshner, 1995; Petry & Meyer, 1987, for reviews). In other words, several different levels of processing may contribute to the final subjective percept. Below, we consider the possible influence of attentional factors on subjective figures, and vice-versa.

8. Phenomenology and its limits once again

The overwhelming majority of previous studies on subjective figures have relied

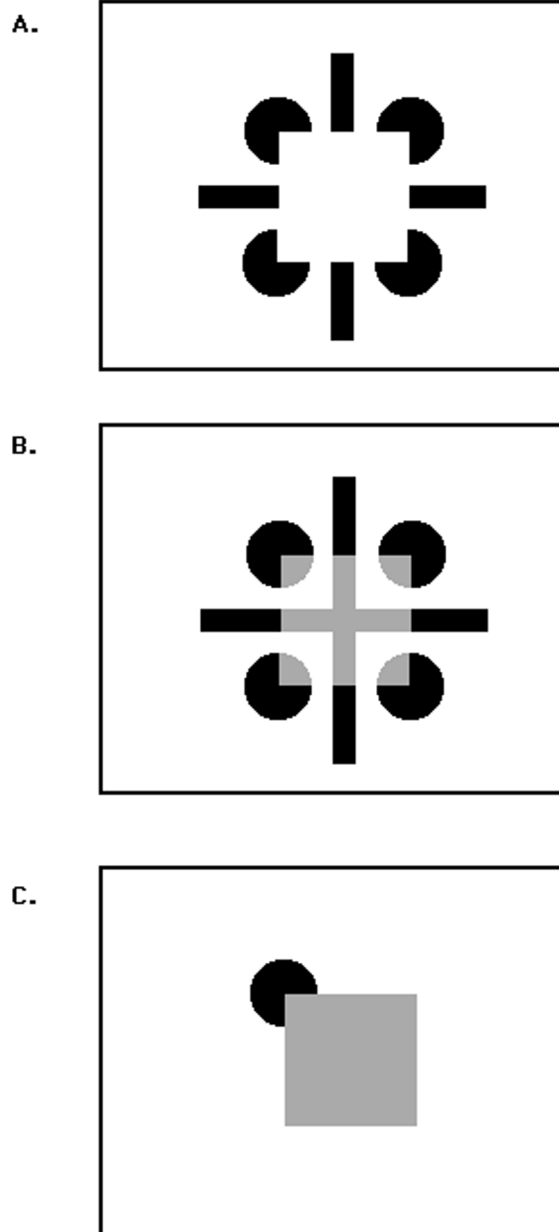


Fig. 3. (A) Example of a Kanizsa subjective square. (B) Example of an apparently transparent Kanizsa subjective square. The central region between inducing circles and the cross appears infused with transparent grey, although it is actually white. This is one example of 'modal' completion. (C) Illustration of 'amodal' completion of a shape which appears to be partly occluded, in this case the circle with the quadrant that appears occluded by the square. Note that no illusory properties are seen in the completed region of the circle.

on phenomenal measures (e.g. rating the strength of the subjective figure, as in Warm, Dember, Padich, Beckner, & Jones, 1987; though see Dresp & Bonnet, 1995; Rubin, Nakayama, & Shapley, 1996, for more objective measures). Moreover, most studies have taken such phenomenal measures while presenting just a single subjective figure display to the observer, often for unspeeded visual inspection and self-rating. Such methods are clearly unsuitable for addressing the many attentional questions one can ask about image segmentation in general, and about subjective figures in particular. For instance, how could an observer possibly ignore a subjective figure while rating its supposedly ‘unattended’ phenomenal strength? This seems a clear case of the spotted-pink-elephant problem. Such unspeeded phenomenal methods will presumably also increase the likelihood of many different levels of processing contributing to the final report.

The restrictions that arise when relying too much upon phenomenology may apply to other recent work on grouping processes also. Just as it is tempting to ask whether subjective figures reflect relatively ‘early’ or ‘late’ processes in vision, one may ask whether segmentation processes in general arise early (as traditionally thought), or might also involve higher-level processes. Rock and Palmer have raised this general question for Gestalt grouping displays in a series of studies with their collaborators (e.g. Palmer, Neff, & Beck, 1997; Palmer & Nelson, in press; Palmer & Rock, 1994; Rock, Nijhawan et al., 1992). More specifically, they asked whether grouping might operate before or after various perceptual constancy mechanisms and various completion mechanisms. A typical experiment used grids of elements, rather like the Wertheimer (1923) grouping demonstrations. The elements on one side of the grid would be of a uniform type, and the elements on the other side would be of a different uniform type. The central column comprised elements of the critical experimental type; the question was whether these would group phenomenally with the type of elements on the left or on the right. Observers judged this explicitly (i.e. making leftwards versus rightwards grouping judgements) in an unspeeded manner with free vision.

Using such a method, Rock, Nijhawan et al. (1992) found that grouping by similarity in achromatic colour apparently takes place only *after* the operation of lightness constancy (they exploited cast shadows or translucent overlays on the central column to decouple retinal luminance and perceived lightness). In forced-choice decisions, observers reported that the central column appeared to belong with other elements that matched in phenomenal lightness, rather than in retinal luminance. Palmer et al. (1997) used a similar method to argue that grouping by similarity in shape may take place after amodal completion of apparently occluded shapes has taken place. More recent studies by Palmer and Nelson (in press) have extended this to the case of modal completion for subjective figures.

Palmer and Rock have taken such results to indicate that ‘grouping is later than you think’, being based on quite sophisticated visual properties, which are closer to the phenomenal percept than to the retinal image. However, one can query the method used, and wonder whether instead grouping might be ‘later only *when* you think’! Given the unspeeded viewing conditions, subjects had ample time to inspect (and even think about) each element of the stimulus in turn. Moreover, given

the forced-choice task (i.e. does the central column group with the left or right?) there may even have been no true grouping at all, with observers simply using their response to indicate which stimuli looked most similar phenomenally. We already know which stimulus pairs these would be: perceived brightness for individually inspected stimuli depends on lightness constancy, partly-occluded shapes usually get amodally completed, and the appropriate inducing stimuli will form subjective figures.

In other words, while Palmer, Rock and their colleagues have asked an extremely pertinent question about the level(s) at which grouping processes can arise, their actual method may only tell us what we already know from introspection (i.e. what the particular stimuli look like, phenomenally, with unsped viewing). Methods from the attention literature such as those reviewed here may provide better handles on the level(s) at which particular segmentation processes arise, and on whether the observer can modulate these at will. Below, we address modal and amodal completion from this perspective.

9. Visual search studies of modal and amodal completion

Several recent studies have examined the roles of amodal completion for partly-occluded objects (e.g. Davis & Driver, 1998; He & Nakayama, 1992; Rensink & Enns, 1998) and of modal completion for subjective figures (e.g. Davis & Driver, 1998) in visual search tasks. In such tasks, the observer must typically determine the presence or absence of a prespecified target among a varied number of non-targets as fast as possible. This task has traditionally been used with the aim of distinguishing those visual properties extracted ‘preattentively’ (i.e. efficiently in parallel, regardless of the number of items in the display) from those extracted only ‘attentively’ (classically, these were thought to require strictly serial processing, leading to linear increases of search time with the number of items to be searched; e.g. Treisman, 1986; Treisman & Gelade, 1980). It is now well known that many caveats must be placed on such interpretations of search data (e.g. see Wolfe, 1998, for review), just as many caveats must be placed on any simplistic preattentive/attentive dichotomy. Nevertheless, the visual search task still remains very useful as one way of addressing attentional issues for specific processes. Moreover, with appropriate tweaking, it can get around the spotted-pink-elephant methodological problem that was described earlier.

As one example of this, Rensink and Enns (1998) examined the search for notched target shapes among complete (i.e. un-notched) non-targets in 2D displays (see Fig. 4 for schematic examples that have been adapted for our present purposes). In control conditions this was an easy task, yielding a ‘parallel’ search outcome (i.e. little or no increase in search time with set-size). In the critical experimental conditions, the same notched target (a notched circle in our example; see Fig. 4A) was placed so that pictorially it appeared to be occluded by an abutting shape in the region corresponding to the notch. Hence, with unrestricted viewing, it would tend to be amodally completed as a partly-occluded version of the same complete shape

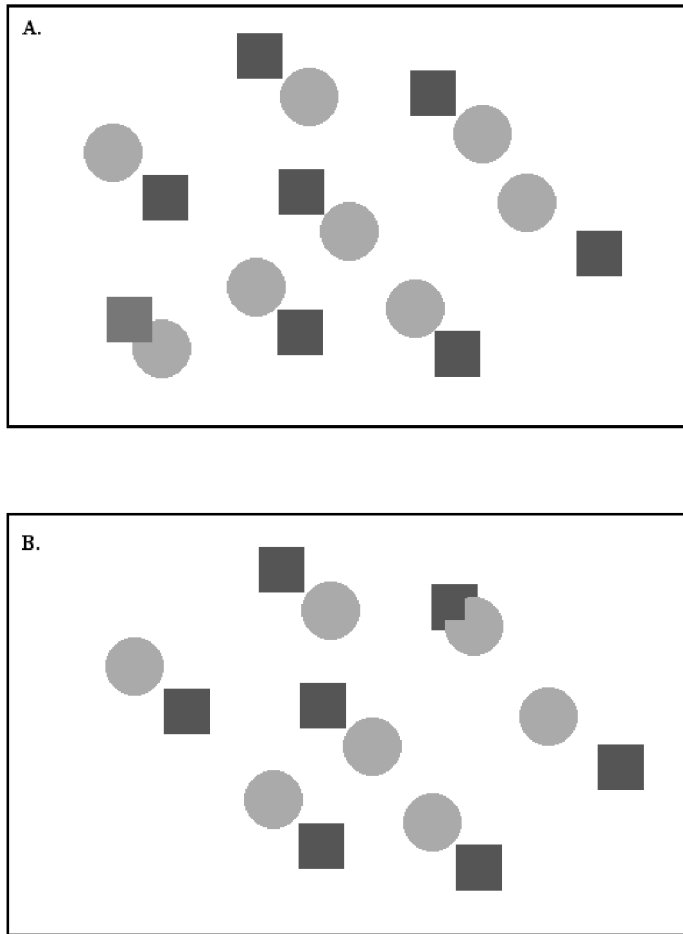


Fig. 4. Search displays analogous to some of those used by Rensink and Enns (1998). The task for these examples would be to determine the presence or absence of a notched circle. Note that this seems much harder in (A) (where the notched circle appears to be partly occluded in the notched region by the abutting square) than in (B).

as the (circular) non-targets. The critical question was how such an apparently occluded target would behave in the search task. Would it act like the notched target (whose shape it still had in the retinal ‘mosaic’) that was previously so easy to find, or like a completed shape, in which case it should presumably now be hard to find (given the presence of many physically complete versions of this shape among the non-targets)? The results clearly showed that the search became very hard, now showing a classic ‘serial’ pattern, with a steep increase in search time against the number of items in the display. He and Nakayama (1992) had previously found a very similar result in a study which used stereoscopic depth to determine whether amodal completion should take place.

These results have several implications. They imply that amodal completion may arise in so-called ‘parallel’ vision (i.e. efficiently for all items at once, without serial inspection of each item in turn). They also imply that amodal completion may still arise even when counter to the observer’s current intentions. Although amodal completion of partly-occluded objects may be very useful for many real-world visual tasks (see Nakayama et al., 1989), it could only be detrimental to performance in the visual search task, as shown by the impaired search performance when such completion arose. This implies that amodal completion (or at least, assignment of the notched edges to the apparent occluder, rather than to the target shape) may arise even against the observer’s will, despite many hundreds of trials attempting to see the notched target in terms of just its retinal shape. Note that in reaching this conclusion, the spotted-pink-elephant problem was neatly side-stepped. The observers never had to be asked directly about whether they saw the target as amodally completed or not; its coding by their visual system was simply inferred from their objective search performance. The only downside of this is that the results tell us little, directly, about phenomenological visual appearance. On the other hand, they do tell us something quite specific about the nature of amodal completion processes, which are presumed to have some ultimate influence on phenomenal vision.

We (Davis & Driver, 1998) recently adapted the basic methods of He and Nakayama (1992) and Rensink and Enns (1998) to examine whether *modal* completion (i.e. as in Kanizsa subjective figures) might similarly arise within ‘parallel’ vision in search tasks, even when this could only be detrimental to objective performance. Essentially, we asked whether a subjective figure might act as an occluder to induce unwanted amodal completion of an abutting notched target, just as the real squares do in Fig. 4. Similarly to the search displays in Fig. 4, observers again searched for a notched circle among complete circles. Smaller ‘pacmen’ stimuli were now arranged so as to produce possible subjective squares, which might serve as occluders in the critical conditions (see Fig. 5). The search for the notched circle among full circles was very slow and inefficient when these appeared (see Fig. 5A) stereoscopically *behind* the potential subjective surfaces (yielding a classic ‘serial’ search outcome), but was fast and efficiently parallel when they appeared in front (Fig. 5B). Various control conditions confirmed that the search difficulty was specific to apparent occlusion of the notched target by a subjective surface, rather than simply to searching in the back plane. Moreover, these control conditions also ruled out trivial accounts of the effect from the subjective surface in terms of low spatial frequency blurring, or the mere alignment of inducing edges with the notched circle target.

As with the results of He and Nakayama (1992) and Rensink and Enns (1998), these results imply that amodal completion (of the notched target) may arise in ‘parallel’ vision, and may do so even when counter to the observer’s intentions. The data of Davis and Driver (1998) further indicate that both these assertions also apply to the formation of subjective surfaces via modal completion. Such surfaces are evidently treated in parallel vision as occluding regions which cannot be seen through, even though they are entirely ‘subjective’ creations of the visual system, as no physical occluder is actually present.

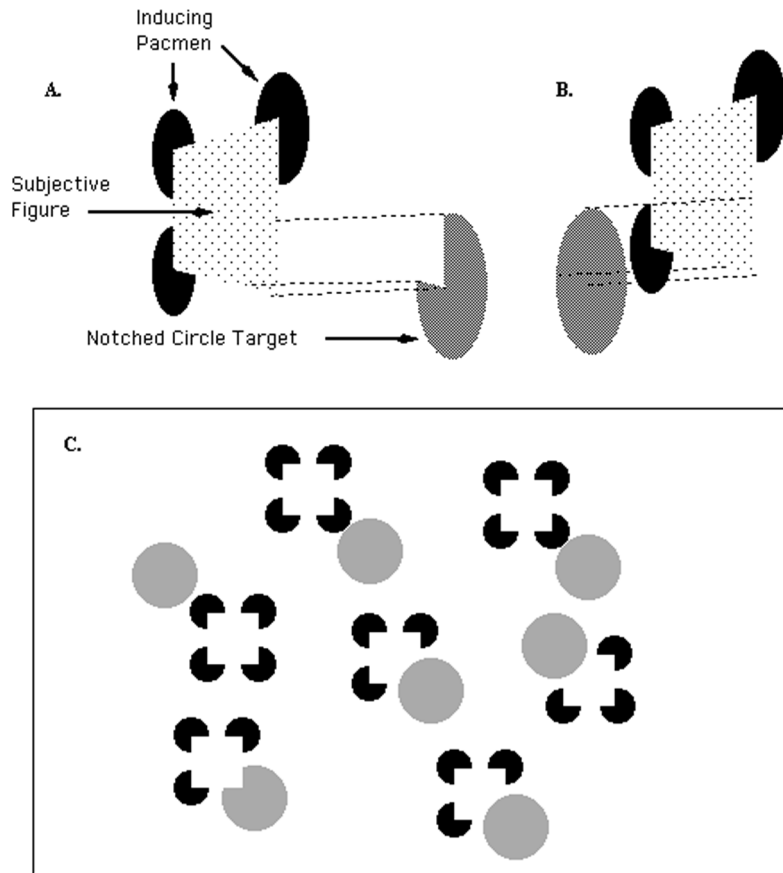


Fig. 5. Schematic depiction of target stimuli (A,B) or a typical search display (C) from Davis and Driver (1998). For a search display like (C), the task would be to determine the presence or absence of a large notched grey circle. Note that this appears difficult, due to apparent occlusion (and associated amodal completion) by the Kanizsa subjective square. In the actual experiment, the large pacmen and small pacmen were in different stereoscopic depth planes (see cartoons of this for one cluster of stimuli in A and B). When the small inducers of the subjective square lay in the front plane (as in A), the search for the large notched circle was very difficult. But when the latter target lay in front instead (see B), so that the subjective square could no longer appear to occlude it, the search became efficient.

10. Distinguishing modal versus amodal completion in visual search: the Metelli rule

The recent evidence from visual search reviewed above suggests that modal and amodal completion may operate similarly in visual search tasks, both arising in parallel, and in an apparently obligatory manner. This can be added to a long list of similarities between the two forms of completion, as previously articulated by Kellman and Shipley (1991). Indeed, Ramachandran (1995) has gone so far as to suggest that modal and amodal completion may not be differentiated until very late

in the visual system. However, despite their commonalities, modal and amodal completion do differ in a crucial phenomenological respect. As noted earlier, in the case of subjective figures, completion is ‘modal’ in the sense that illusions specific to the visual modality (i.e. filling-in of illusory colour and/or brightness) take place within the completed region (e.g. see Fig. 3B). No such illusions of colour and brightness arise for amodally-completed regions, which are seen as lying behind apparently occluding surfaces (e.g. Fig. 3C). We (e.g. Davis & Driver, 1997, 1998) have examined whether this filling-in of illusory colour and/or brightness, for the case of modal completion only, might be of functional as well as phenomenal significance, as described below. In more recent studies, we have also begun to address whether modal and amodal completion may behave differently within visual search tasks. Our search experiments took advantage of one well established subjective difference between modal and amodal completion, which relates to the perception of transparency.

Kanizsa (1979) and Van Tuijl and de Weert (1979) demonstrated that for simple patterns comprising three luminance levels, modal completion of transparent subjective figures is constrained by the luminance profile of the inducing stimuli in a way that amodal completion is not (see also Nakayama & Shimojo, 1992). We will refer to this constraint as the ‘Metelli rule’, since it stems from the description by Metelli (1974) of the luminance conditions constraining transparency perception. In order for the elements of an inducing stimulus to form part of a transparent, modally-completed surface, the luminance of these inducing elements must fall between the luminance of the surround and the luminance of the inner region over which the subjective figure is formed. For example, the stimulus in Fig. 6A satisfies these constraints. The luminance of the four small grey sectors on the otherwise black circles falls between that of the white inner region and the black surround of the circles. These grey sectors can therefore be seen to form part of a transparent, modally-completed surface. However, the stimulus in Fig. 6B fails to satisfy the Metelli rule prescription for the luminance relations. The luminance of the small dark sectors on the otherwise grey circles does *not* fall between the luminance level of the grey surrounding regions and the white inner region, being less than both. Thus, the dark sectors cannot form parts of a modally-completed surface.

This dependence of modal completion upon the luminance profile of the stimulus makes functional sense when one considers the optics of transparent surfaces (see Metelli, 1974), and also that modal completion processes may serve to encode partly-camouflaged surfaces (e.g. Davis & Driver, 1994; Ramachandran, 1987). If the grey regions of Fig. 6A did indeed belong to a uniform transparent surface, that surface must be rather light, since it apparently causes the black circles to appear lighter where it overlaps with them. Such a light surface could conceivably become invisible in the current retinal image at regions where it overlapped with a lighter background. This luminance configuration is therefore at least consistent with the presence of a partly-camouflaged transparent surface. It might therefore be adaptive for modal completion mechanisms to code a partly-camouflaged but complete transparent surface under these circumstances.

By contrast, if the small dark sectors in Fig. 6B actually formed part of a larger

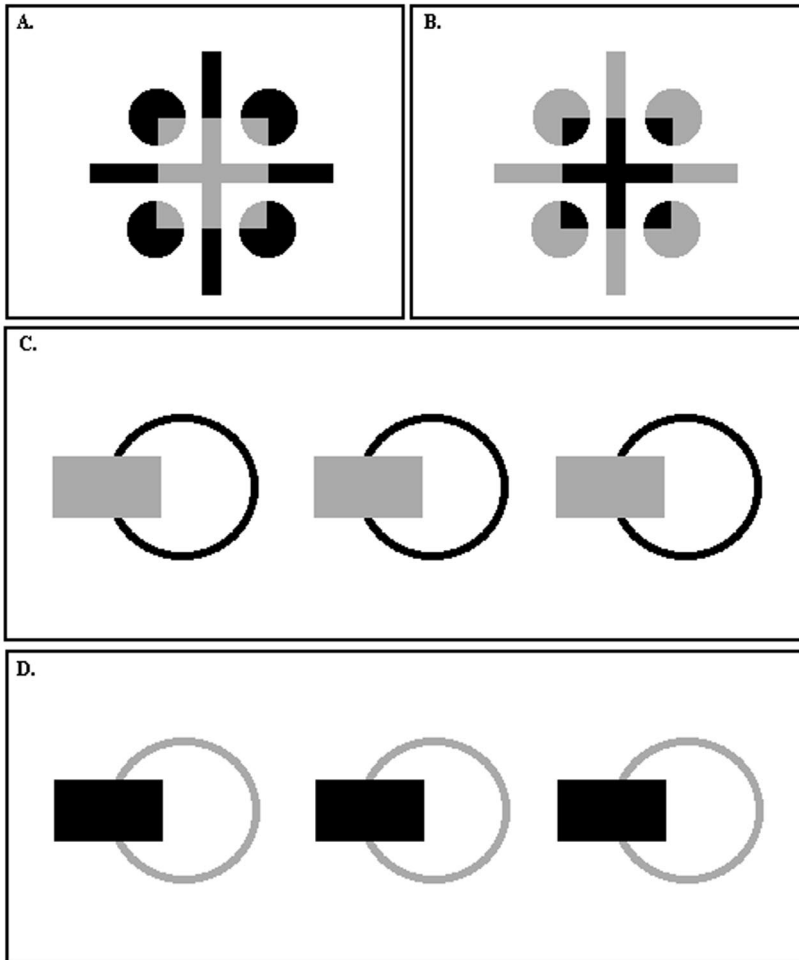


Fig. 6. (A) Transparent Kanizsa square. Note that the grey of the inducing quarter-segments on the four circles and of the central region on the cross has a luminance which is intermediate between the surrounding black and the white over which modal completion takes place. This luminance profile is consistent with the Metelli rule. (B) The same display layout, but now with luminance levels that do *not* fit the Metelli rule. Note that now no subjective figure is formed. (C,D) Example displays that with free fusion illustrate the change from modal to amodal completion with a reversal in depth, and the dependence of completion on the Metelli rule only for the case of modal completion (i.e. only when the circular element appears in front of the abutting rectangle, as with uncrossed free fusion of left and centre columns; uncrossed fusion of right and centre columns should place the circle at the back plane). (C) is inconsistent with the Metelli rule and (D) is consistent.

uniform transparent surface, this would have to be a dark surface, since it causes the otherwise grey circles to appear darker where it overlaps in front of them. It is unlikely that such a dark surface could be invisible where it passes in front of a

white background. Such a luminance configuration is therefore inconsistent with the presence of a partly-camouflaged, transparent surface, and hence modal completion mechanisms should not operate. Of course, we would not suggest that the visual system literally undertakes the kind of inferential procedure we have outlined here, in explaining why the Metelli rule makes reasonable sense. Nakayama and Shimojo (1992) provide one biologically plausible account of how the visual system could come to obey such constraints.

The dependence of modal completion upon the luminance configuration of the inducing stimuli can be experienced by free-fusing the *left* and centre patterns in Fig. 6C with uncrossed fusion. This procedure should yield the impression of a notched ring element at a nearer plane than the rectangle it abutts. Note that the notched ring is *not* modally completed to form a complete, transparent ring, because its luminance does not fall between that of the rectangle and the background (i.e. the luminance profile is inconsistent with modal completion in terms of the Metelli rule). However, when the luminance profile of the elements is reversed, as in Fig. 6D, it now becomes consistent with modal completion of a transparent surface in terms of the Metelli rule. When the *left* and centre patterns in this figure are free-fused (uncrossed), the grey notched ring element should now appear modally complete, as a transparent subjective surface (a complete ring) lying in front of the abutting black rectangle.

The Metelli transparency constraints apply to modal completion, because modally-completed surfaces appear to pass *in front of* the surfaces they overlap with (and transparent surfaces lying in front of other surfaces can alter the amount of light reaching the retina from the latter surfaces). However, this does not apply to *occluded* regions of surfaces, such as those generated by amodal completion. Occluded regions generally cannot prevent light from other objects reaching the observer. Hence, the luminance constraints that apply to modal completion do not apply to amodal completion. Indeed, several studies have found no modulation of amodal completion by the luminance profile of the inducing stimuli (e.g. Kanizsa, 1979; Nakayama & Shimojo, 1992). This can be experienced by free-fusing the centre and *right* patterns of Fig. 6C,D. The stimulus should now appear as a ring that is further away than the abutting rectangle. Note that the notched ring becomes amodally completed to form an apparently complete but partly-occluded ring in both cases (i.e. even when its luminance does not fall between that of the rectangle and the background (Fig. 6C), because the Metelli constraint does not apply to cases of partial occlusion).

We (Davis, Russell, & Driver, unpublished data) recently tested whether these constraints apply within the ‘parallel vision’ tapped by appropriate search tasks, and whether they do so even when any completion can only be detrimental to performance. Subjects searched for notched rings (i.e. like those in Fig. 6C,D) among physically complete rings in four different search conditions. The luminance profile of the displays was either consistent or inconsistent with the Metelli rule; orthogonally to this, stereo disparity was arranged so that all the rings either lay in the back plane (behind the rectangles) or in the front plane (as when viewing Fig. 6C,D with the appropriate free fusion). The results showed efficient parallel search only when

the rings lay in the front plane, with a luminance profile that was *inconsistent* with the Metelli rule. In all the other three conditions, a classic ‘serial’ search pattern was found. The implication is once again that in these three conditions, completion of the notched target arose in parallel vision (thus making it hard to distinguish from the physically complete non-targets) even though this was highly detrimental to performance, and thus counter to the observer’s intentions. Amodal completion (of notched circles apparently lying behind the abutting rectangles) was unaffected by the Metelli constraints. By contrast, modal completion (i.e. of notched circles lying in front of the abutting rectangles) was highly dependent on the Metelli rule, taking place only when the luminance profile was consistent with the constraints on perception of a subjective transparent surface. When the Metelli rule was broken, no modal completion took place, and so the target ‘popped out’ as the only incomplete circle in the display. Note that these results cannot be explained by the luminance relations alone, because the effect of these depended on depth relations, in a manner consistent with the Metelli rule applying to modal but not amodal completion.

This shows that the Metelli rule for the perception of transparency, which relates to the real-world optics of transparent surfaces, can operate within parallel vision as tapped by appropriate visual search tasks. Moreover, it also suggests that cases of modal and amodal completion are treated differentially within such parallel vision, in a manner consistent with their very different phenomenology, and with the differential impact of the Metelli rule that is apparent when each item is inspected with free vision (see Fig. 6).

11. Differential effects of modal versus amodal completion upon attentional processes

The conventional interpretation of the above search results would be that both modal and amodal completion can arise, and be distinguished, within ‘preattentive’ vision. In further research, we have gone on to examine whether modal and amodal completion may have different effects upon attentional processes (Davis & Driver, 1997, 1998). These studies used displays like Fig. 7A, comprising two segments of a grey ellipse separated by a vertical white bar. When stereo disparity placed the bar in front of the ellipse segments, the latter were amodally completed, to give the appearance of a continuous ellipse occluded centrally by the white bar (much as it appears in Fig. 7A). But when stereo disparity placed the bar behind the grey segments, these were now modally completed as a continuous transparent grey surface in front of the bar, with illusory grey ‘filling-in’ of the central region (much like the cartoon in Fig. 7B; note that the luminance profile of the inducing stimulus was consistent with the Metelli rule, as in Fig. 7A).

In two series of studies, we examined whether modal versus amodal completion might have differential effects on the ‘spreading of attention’ (as shown by distractor or cueing effects), consistent with the differential phenomenology for the central region (which is filled-in with illusory colour for the modal case only). In one series (Davis & Driver, 1997), we placed a target character on the central bar, and distrac-

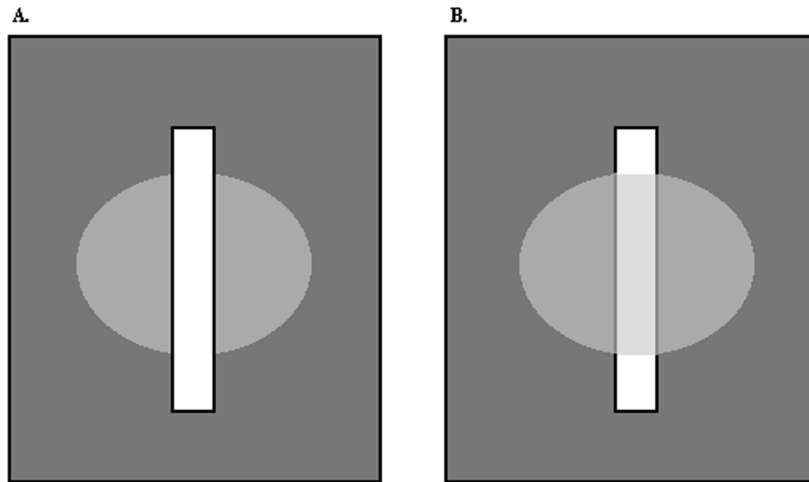


Fig. 7. (A) Example of the displays used by Davis and Driver (1997, 1998). In some conditions the central white bar was placed stereoscopically in front of the flanking grey regions; this led to the percept of an amodally-complete grey ellipse passing behind the white bar. In other conditions, the white bar was placed stereoscopically behind instead. This led to the percept of a modally-complete, transparent grey ellipse passing in front of the central bar, with illusory grey being ‘filled-in’ where it appeared to overlap the white bar, as cartooned in (B).

tor characters on the outer segments of the ellipse (or vice-versa). The efficiency of distractor rejection was measured (via compatibility effects on RT to the target). We found that distractors produced more interference in the modal completion condition than in the amodal completion condition. We took this to show that when observers attempted to select just one region of the display (e.g. just a central region, or just one outer region of the ellipse), they tended instead to select the entire modally-completed object in the modal completion situation. Thus, their ‘attention’ (i.e. their selective intake of information) tended to spread from modally-completed regions to include their inducers, and vice-versa, in a manner that did not apply for amodally-completed regions and their inducers. This objective performance result is, of course, consistent with the phenomenology here. The central region of the display (see Fig. 7) appears to include a visible surface that ‘belongs’ to the outer inducers of the ellipse only in the modal completion condition (i.e. where a subjectively complete ellipse becomes filled-in with illusory grey at the centre). This does not apply to the amodal completion condition (where although a complete ellipse is perceived, it appears to be occluded by the bar in the central region).

In another set of studies (Davis & Driver, 1998), we made a similar point but with a different attentional measure. We ‘cued’ the bar, or the outer segments of the ellipse, by transiently altering their vertical extent. We then measured the effect of this cueing on RT to subsequent target characters, which could be presented either on the bar or on the outer (physically present) segments of the ellipse. In the modal completion condition, cueing the outer segments of the ellipse appeared to attract

attention to the central region as well as to the outer regions, whereas in the amodal completion this did not arise.

Again, this objective performance result seems consistent with the phenomenology of the displays, as the central region of the displays appears to include a visible surface that ‘belongs’ to the ellipse only in the modal completion condition, where it becomes ‘filled-in’ with illusory transparent grey. Indeed, we went so far as to suggest that such filling-in may take place precisely in order to indicate a potentially visible surface, which may then lead to appropriate attraction of attention (or spreading of attention across the entire surface), just as for physically present surfaces. In other words, ‘filling-in’ may serve functional as well as phenomenal purposes, and indeed these may be closely related. These studies appear to provide examples of ‘object-based’ visual attention, in which the critical object depends on a subjective illusion. Of course, in all the cases we have discussed (i.e. even for the experiments which did not directly concern subjective figures), the effective groups or objects are similarly the products of segmentation processes in the visual system, given the stimulus input.

The studies described so far seem broadly consistent with traditional views (e.g. Neisser, 1967; Duncan, 1984; Driver & Baylis, 1989, 1998), within the simplifying preattentive/attentive dichotomy, that much image segmentation can precede and constrain attentional selectivity. We have suggested here that Gestalt grouping can still arise under conditions of ‘inattention’, that undetected background changes in the flicker paradigm may still be extracted by the visual system, and can even influence the phenomenology of foreground elements, and that modal and amodal completion may arise and be differentiated by parallel vision within visual search, incorporating relatively sophisticated constraints such as the Metelli rule. Finally, we have suggested that modal and amodal completion may impose differential constraints on attentional processes, in accord with their very different phenomenology.

All this seems consistent with the traditional view of segmentation processes largely preceding attentional processes. However, we think that under appropriate conditions, the influence can also work in the other direction (see Farah, Wallace, & Vercera, 1993; Humphreys, Olson, Romani, & Riddoch, 1996), with attentional processes modulating segmentation processes to some extent. In our final experimental section, we give a recent example of this possibility from our own work.

12. Possible attentional modulation of very early visual segmentation

Here we exploited a particularly well understood form of primitive image segmentation, which is thought to reflect interactions between interconnected neurons in primary visual cortex (Polat & Sagi, 1993, 1994). Contrast thresholds for detecting a central Gabor patch stimulus (i.e. a cosine grating convolved with a Gaussian filter; see Fig. 8) can be systematically affected by flanking Gabor patches. Such Gabor stimuli are experimentally useful, not only because of their precise

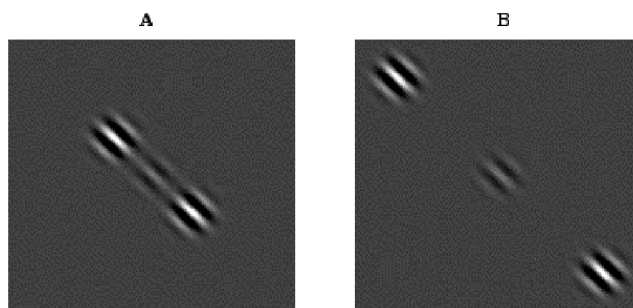


Fig. 8. (A) Central Gabor patch with high-contrast and collinear flankers separated from the central target by only three Gabor wavelengths. In such conditions, the flankers lower contrast thresholds for detecting the central target. This facilitatory effect falls off with increasing separation, being reduced at the eight-wavelength separation shown in (B).

physical characteristics, but also because they are thought to stimulate highly specific populations of cells in primary visual cortex.

It is now well established that thresholds for a central target patch are affected by the presence and nature of flanking patches (e.g. Polat & Sagi, 1993, 1994), reflecting highly systematic lateral interactions. Target detection is maximally facilitated for a central patch that is iso-oriented and aligned with flanking patches of the same spatial frequency at separations of around three Gabor wavelengths (see Fig. 8A). Suppression is found at lower separations, while the facilitation gradually declines at higher separations (e.g. Fig. 8B). Global configuration is critical; non-collinear arrangements of iso-oriented patches produce suppression instead of facilitation (Zenger & Sagi, 1996). Further properties of the lateral interactions between target and flankers are also well characterized; they are invariant over magnification and independent of relative phase above critical separations, and flankers with orthogonal local orientation to the target have no influence (e.g. Field, Hayes, & Hess, 1993; Polat & Sagi, 1993).

These lateral interactions between Gabor patches have been modelled by an array of detectors akin to V1 neurons (e.g. Polat, 1999; Usher, Bonneh, Sagi, & Herrmann, 1999), with neighbouring units reinforcing or suppressing each other via lateral connections, similar to those actually observed in primary visual cortex (e.g. Spillmann & Werner, 1996). More importantly for our current topic, these lateral interactions may reflect a very early form of segmentation involved in contour integration and texture segmentation (e.g. Polat, 1999). For example, note that when viewing Fig. 8A, a virtual contour may appear to be formed, linking the central patch and the aligned flankers, while this applies less at the greater separation in Fig. 8B. Given the suspected substrate in primary visual cortex, lateral interactions between flanking Gabor patches may tap into the very first cortical stage of image segmentation.

We wondered whether these lateral interactions and the associated formation of virtual contours might depend on the flankers being attended. To address this, one of us (E.F.) instigated a novel variant of the lateral interactions paradigm (see Fig. 9),

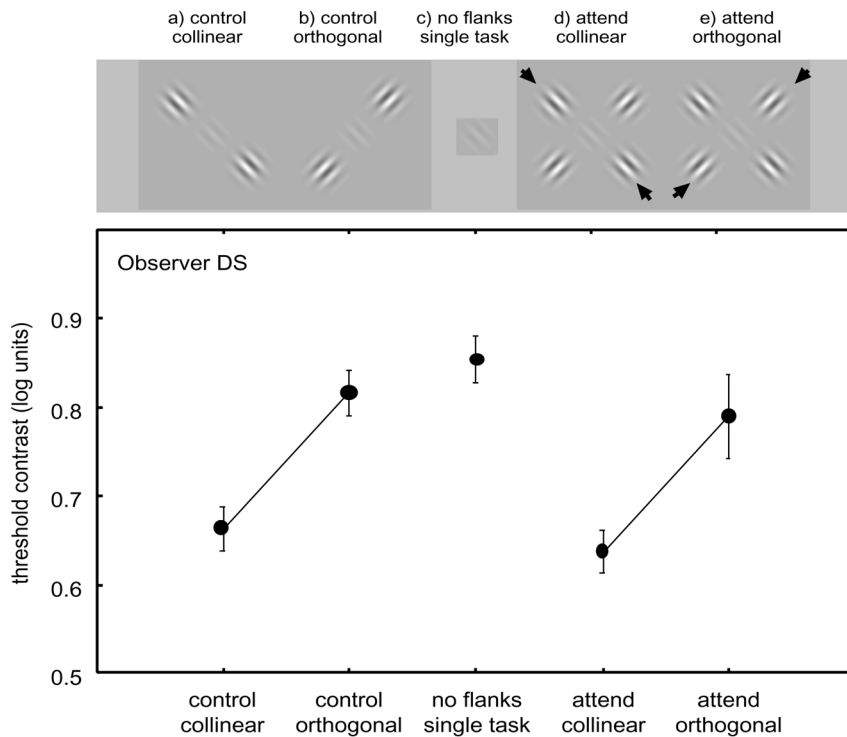


Fig. 9. The upper panel shows example displays from the preliminary study by Freeman, Sagi, and Driver, as described in the text (the arrows in D and E were not physically present, and are shown only to indicate which flankers were attended for the secondary task). The lower panel plots contrast thresholds for detecting the central target in each of the conditions shown above. Data from one experienced observer (D.S.) with standard error are shown.

which we have since pursued in collaboration with Dov Sagi (Freeman, Sagi, & Driver, unpublished data). In this new version, contrast thresholds for detecting a central target Gabor patch are measured not only in the context of one pair of flankers (as is conventional; see Fig. 9A,B), but also with *two* pairs of flankers, which have orthogonal global and local orientations (see Fig. 9D,E). Thus, in any given display of four flankers, two of them will be collinear with the target (those along one diagonal), while the other two will not (those on the other diagonal). We manipulate which pair of flankers (i.e. which diagonal) is attended for a secondary task performed in addition to central target detection.

In the study illustrated by Fig. 9, this secondary task was vernier acuity (i.e. judging the direction of a small offset for the two flankers along one diagonal). On each trial, there were two successive displays. Observers then made unsped, independent decisions concerning which of the two intervals contained (a) a central target, and (b) a particular direction of vernier offset for the relevant two flankers. In some blocks of trials (see Fig. 9A,B) only two flankers were present, which could be

collinear or non-collinear with the target. Observers performed both the central detection task and the vernier task for the flankers. In other blocks (see Fig. 9D,E) four flankers were presented, with observers performing the secondary vernier task for only one relevant pair (i.e. those indicated by arrows in Fig. 9, although no arrows were actually present in the study). Finally, baseline blocks (Fig. 9C) presented no flankers, with observers performing only the central target detection task. Any flankers were separated from the central target by four Gabor wavelengths, a separation where facilitation of target detection thresholds by collinear flankers is usually found, while orthogonal flankers have no effect.

Contrast thresholds for one experienced psychophysical observer (an initially sceptical Dov Sagi!) are shown in the lower panel of Fig. 9 underneath the corresponding display type from the upper panel. A second experienced observer, E.F., produced a very similar pattern. With a single pair of flankers (leftmost two points in the graph), the usual lowering of target threshold was found when the flankers were collinear rather than orthogonal. With two pairs of flankers (rightmost two points in the graph), central target detection depended heavily on which pair was attended for the secondary vernier task, with lower central thresholds when the pair that was collinear with the target was attended. Indeed, with four flankers present, the central thresholds behaved as if the currently ignored pair of high-contrast flankers was simply absent (compare leftmost two datapoints with rightmost two datapoints in the graph of Fig. 9).

These initial results provide preliminary evidence for attentional modulation of lateral interactions between Gabor patches, which had previously been attributed to hard-wired lateral connections within primary visual cortex (e.g. Polat & Sagi, 1993; Usher et al., 1999), that may subservise contour integration and texture segmentation. We shall exploit this apparent attentional modulation in further psychophysical studies to determine exactly how the lateral interactions are influenced by attentional set in terms of existing V1 models (e.g. Li, 1998; Zenger & Sagi, 1996). These models suggest various possibilities. For instance, attending to specific flankers for the secondary task might simply boost activity in cells responding to those flankers, alter their dynamic range, or affect the strength of connections between cells responding to the flankers, and those responding to the target. Models of primary visual cortex are now sufficiently advanced that these different possibilities for the underlying mechanism of the attentional modulation can lead to different psychophysical predictions.

This study provides an apparent case of attentional factors modulating a very early and primitive form of image segmentation, namely lateral interactions between detectors for specific spatial frequencies and orientations, which relate to contour integration. In this example, attention seems able to influence segmentation processes, rather than the influence being only in the reverse direction. Nevertheless, various further questions from the literature on ‘object-based’ visual attention still arise, even for this situation. For instance, is the effect of attending to one pair of flankers rather than the other (upon central target thresholds) only found when the secondary task for these relevant flankers concerns their *spatial* alignment (as in the vernier task, which may have encouraged a specifically spatial strategy of judging

the virtual contour formed by central target and collinear flankers together; see Fig. 9)? Or is it still found even when observers judge entirely non-spatial properties for the relevant flankers? The latter outcome might be predicted by some ‘object-based’ accounts of attention (e.g. Duncan, 1984), which propose that whenever one property of an item is judged, all of its other properties (e.g. spatial frequency and orientation) may also be extracted.

13. Mutual constraints between segmentation and attention

We hope the studies we have described illustrate that attentional issues can provoke new questions about visual segmentation, and vice-versa. Our description may also have illustrated the intimate but vexing relation with phenomenology which characterizes much research on perception and attention. Phenomenology provides a useful starting point for many questions, and is also part of what must be explained, yet it is clearly inadequate for resolving many of the questions that an attentional perspective raises. This is not only because of the spotted-pink-elephant methodological problem, but also because our phenomenal visual experience represents only some aspects of the underlying visual processing. Indeed, the study of attention is largely motivated by the selective nature of perceptual awareness.

As noted earlier, much of the evidence we have described (e.g. on grouping under inattention, visual search in relation to modal versus amodal completion, and the differential effects of these forms of completion on distractor interference, or on cueing effects) is broadly consistent with the traditional story that many segmentation processes constrain attentional selectivity (e.g. Driver & Baylis, 1998; Duncan, 1984; Neisser, 1967). Our results suggest that such segmentation processes can include cases as sophisticated as modal or amodal completion, and incorporate constraints as subtle as the Metelli rule, concerning the optics of transparent surfaces. On the other hand, the final study described here (on detection of Gabor patches) indicates that some aspects of attention (here, the task-relevance of one pair of flankers versus the other) can modulate very early segmentation processes (perhaps even the very first cortical stage of image segmentation).

If approached within the conventional preattentive/attentive dichotomy, this combination of results (i.e. cases where segmentation influences attention, but also cases of the reverse) would seem paradoxical. However, this seeming paradox only arises because the conventional preattentive/attentive view assumes a strictly serial succession of (just two!) distinct stages. It is now well known that human vision incorporates many more than just two levels of processing, and that back projections are at least as numerous as forward projections. Moreover, as we have emphasized, attention is best thought of as an umbrella-term for several selective processes, not as a single process with a single modulatory influence. In any case, ‘interactive activation’ models, incorporating back projections, have long been able to resolve apparent paradoxes of the kind raised here, whereby seemingly ‘later’ processes can influence apparently ‘early’ processes (e.g. McClelland & Rumelhart,

1986). We would suggest that the relation between various forms of segmentation, and of attention, has this mutually constraining nature.

The phenomenology of many grouping displays is also consistent with such mutual influences, whereby segmentation can constrain selective attention, but the reverse influence is also possible. While grouping can certainly restrict which units you can pick out subjectively, you can also modulate subjective grouping at will to some extent in many grouping displays, especially when the stimulus factors are relatively subtle. Finally, the burgeoning neuroscience evidence also suggests that many forms of attentional selectivity and grouping are implemented in the brain by interactions between multiple levels of processing (e.g. Desimone & Duncan, 1995; Driver & Frith, *in press*; Kastner et al., 1999), and not in the strictly feed-forward manner envisaged by the traditional preattentive-then-attentive dichotomy.

We consider such an interactive architecture to be a plausible explanation for the mutual constraints between segmentation and attention. However, some of the apparent conflicts in attention research may arise for other reasons. For instance, it has been proposed (Lavie, 1995) that one critical factor may be attentional load, which often varies in an uncontrolled manner across different experiments, or is not manipulated across the entire possible range (i.e. may not include situations where attention reaches ‘absolute zero’; Nakayama & Josephs, 1998). We are sympathetic to this general point, and concede that it may apply to some of the observations we have discussed. For instance, the evidence for implicit grouping of unattended displays, found with the method shown in Fig. 1, might subsequently be shown to vary with the load of the central task. However, on its own, the absolute-zero point does not satisfactorily resolve why in some cases segmentation processes can be modulated by attention, whereas in other cases segmentation clearly constrains attentional selectivity. The absolute-zero approach essentially claims only that many apparently ‘preattentive’ processes might in fact prove subject to some attentional limitations if the latter were pushed far enough (e.g. Joseph et al., 1996). This emphasizes only one direction of influence (i.e. attention affecting segmentation), saying nothing about the many cases where segmentation constrains selectivity.

14. Relation to other ‘object-based’ notions in cognitive science

Throughout this paper, we have rarely used the term ‘object’, despite the stated theme of this volume. Typically, we have referred to ‘segmentation’ or ‘grouping’ instead, quite intentionally. We did so because we consider ‘object’ to be an unfortunately vague and intuitive term in our particular context, which might nevertheless be mistaken to have a highly specific meaning by some readers. In our view, most (perhaps all) of the literature on ‘object-based’ attention is in fact concerned with how segmentation processes constrain attentional processes, and vice-versa (see also Driver, 1999, for a similar account of the extensive neuropsychological literature on so-called ‘object-centred neglect’). We did not have space to describe all the variations on the object-based attention theme here (see Scholl, 2001, for review). For instance, some of this literature is concerned with how segmentation into distinct

feature-bundles may unfold or be maintained over time in dynamic displays (e.g. Kahneman, Treisman, & Gibbs, 1992; Pylyshyn & Storm, 1988; Scholl, Pylyshyn, & Feldman, 2001). Other papers are concerned with how different dimensions (e.g. colour and shape) may be bound together at a particular location to construct multi-dimensional objects (e.g. Luck & Vogel, 1997; Treisman & Gelade, 1980). But in our view, these are all variations on a more general question concerning the relation between attention and segmentation (which pieces of information should be bound together, be this over space or over time).

Despite our present reluctance to use the term ‘object’, it seems clear that a major role of segmentation processes is to distinguish between those parts of the sensory input that are heuristically likely, on the basis of segmentation factors, to correspond to distinct objects (or separate image sources) in the real world, thus constituting ‘proto-objects’. As described above (and more extensively by Driver & Baylis, 1998), image segmentation often constrains our ability to attend selectively, so that we may tend to select all or most of one ‘proto-object’, even when we try to select just part of it (as in our study (Davis & Driver, 1997) of distraction in the context of modal versus amodal completion; see Fig. 7). Similarly, image segmentation can constrain what we may efficiently search for. For example, the search studies of modal and amodal completion suggest that we cannot obtain direct access to the retinal shape of stimuli, as completion processes operate to some extent even against our will, yielding proto-objects which go beyond the various shapes in the retinal mosaic.

Such proto-objects produced by the various image segmentation processes may often be more primitive than the conceptual objects, or real-world objects, which we discuss verbally in daily life. As the term ‘mid-level’ vision implies, they typically reflect bundles of visual information which are packaged on the basis of properties that go beyond raw image statistics (or primitive ‘features’), yet which fall short of conceptually recognized entities. Nevertheless, they may bear a fairly close correspondence to the distinction between separate visual ‘objects’ which we experience phenomenally (as for the modally-completed ellipse with its filled-in colour that is cartooned in Fig. 7B versus the white bar which it appears to occlude). Moreover, the proto-objects produced by image segmentation processes, and the constraints they impose on various attentional processes, may provide an essential precursor for learning more conceptual distinctions between separate types of objects, as in the developing child (e.g. see Carey & Xu, 2001).

Despite these possible links to other object-based concepts in cognitive science, we still prefer to refer to the processes we have studied in terms of image segmentation rather than ‘objects’, as we think this encourages the right way of thinking about them. It leads to questions about the various visual processes which construct the effective ‘proto-objects’, rather than simply assuming separate objects as givens. In this respect, we think that the various manifestos proclaiming visual attention to be ‘object-based’ (see Scholl, 2001, for review) may now have served their useful purpose. They have focused research on the perceptual units which attentional processes operate upon, and thus moved attention research beyond the simplistic spatial spotlight metaphors which dominated much previous work, to trigger new

questions about segmentation and binding. We think that further progress in visual attention research will now stem from examining specific visual processes, such as particular forms of segmentation, in more detail. Many such segmentation processes have already been well characterized in previous vision research (as for lateral interactions between Gabor patches, or modal versus amodal completion), but with little previous consideration of attentional issues. We hope to have illustrated here that perception research and attention research can be complementary on such issues, with the two previously separate research traditions being strengthened when combined.

Acknowledgements

The experimental work described here was supported by project grants from the Economic and Social Research Council (UK), and the Biology and Biotechnology Research Council (UK) to J.D. Our thanks to Gordon Baylis, Diane Beck, Henry Howlett, Brian Scholl, and the late Irvin Rock.

References

- Baylis, G. C., & Driver, J. (1992). Visual parsing and response competition: the effect of grouping factors. *Perception & Psychophysics*, *51* (2), 145–162.
- Broadbent, D. E. (1958). *Perception and communication*. Oxford: Oxford University Press.
- Carey, S., & Xu, F. (2001). Infant knowledge of objects: beyond object files and object tracking. *Cognition*, this issue, *80*, 179–213.
- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, *25*, 975–979.
- Davis, G., & Driver, J. (1994). Parallel detection of Kanizsa subjective figures in the human visual system. *Nature*, *371* (6500), 791–793.
- Davis, G., & Driver, J. (1997). A functional role for illusory colour spreading in the control of focused visual attention. *Perception*, *26* (11), 1397–1411.
- Davis, G., & Driver, J. (1998). Kanizsa subjective figures can act as occluding surfaces at parallel stages of visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *24* (1), 169–184.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.
- Dresp, B., & Bonnet, C. (1995). Subthreshold summation with illusory contours. *Vision Research*, *35* (8), 1071–1078.
- Driver, J. (1999). Egocentric and object-based visual neglect. In N. Burgess, K. J. Jeffery, & J. O’Keefe (Eds.), *The hippocampal and parietal foundations of spatial cognition* (pp. 67–89). Oxford: Oxford University Press.
- Driver, J. (in press). A selective review of selective attention research from the past century. *British Journal of Psychology*.
- Driver, J., & Baylis, G. C. (1989). Movement and visual attention: the spotlight metaphor breaks down. *Journal of Experimental Psychology: Human Perception and Performance*, *15* (3), 448–456.
- Driver, J., & Baylis, G. C. (1998). Attention and visual object segmentation. In R. Parasuraman (Ed.), *The attentive brain* (pp. 299–325). Cambridge, MA: MIT Press.
- Driver, J., & Frith, C. (in press). Shifting baselines in attention research. *Nature Reviews Neuroscience*.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, *113* (4), 501–517.

- Fahle, M., & Koch, C. (1995). Spatial displacement, but not temporal asynchrony, destroys figural binding. *Vision Research*, 35 (4), 491–494.
- Farah, M. J., Wallace, M. A., & Vercera, S. P. (1993). 'What' and 'where' in visual attention: evidence from the neglect syndrome. In I. H. Robertson, & J. C. Marshall (Eds.), *Unilateral neglect: clinical and experimental studies* (pp. 123–138). Hillsdale, NJ: Erlbaum.
- Fernandez-Duque, D., & Thornton, I. M. (2000). Change detection without awareness: do explicit reports underestimate the detection of change in the visual system? *Visual Cognition*, 7, 323–344.
- Field, D. J., Hayes, S., & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local "association field". *Vision Research*, 33 (2), 173–193.
- He, Z. J., & Nakayama, K. (1992). Surfaces versus features in visual search. *Nature*, 359 (6392), 231–233.
- Holender, D. (1986). Semantic activation without conscious activation in dichotic listening, parafoveal vision, and visual masking: a survey and appraisal. *Behavioural and Brain Sciences*, 9, 1–23.
- Humphreys, G. W., Olson, A., Romani, C., & Riddoch, M. J. (1996). Competitive mechanisms of selection by space and object: a neuropsychological approach. In A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), *Converging operations in the study of visual selective attention* (pp. 365–393). Washington, DC: APA Press.
- Joseph, J. S., Chun, M. M., & Nakayama, K. (1996). Attentional requirements in a "preattentive" feature search task. *Nature*, 379, 805–807.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: object-specific integration of information. *Cognitive Psychology*, 24 (2), 175–219.
- Kanizsa, G. (1979). *Organization in vision: essays on Gestalt perception*. New York: Praeger.
- Kanwisher, N., & Driver, J. (1992). Objects, attributes and visual attention: which, what and where. *Current Directions in Psychological Science*, 1 (1), 26–31.
- Kastner, S., Pinsk, M. A., De-Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, 22 (4), 751–761.
- Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, 23 (2), 141–221.
- Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 451–468.
- Leshner, G. W. (1995). Illusory contours: toward a neurally based perceptual theory. *Psychonomic Bulletin & Review*, 2 (3), 279–321.
- Lewis, J. (1970). Semantic processing of unattended messages using dichotic listening. *Journal of Experimental Psychology*, 85, 225–228.
- Li, Z. (1998). A neural model of contour integration in the primary visual cortex. *Neural Computation*, 10 (4), 903–940.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390 (6657), 279–281.
- Mack, A., & Rock, I. (1998). *Inattention blindness*. Cambridge, MA: MIT Press.
- Mack, A., Tang, B., Tuma, R., Kahn, S., & Rock, I. (1992). Perceptual organization and attention. *Cognitive Psychology*, 24 (4), 475–501.
- Marr, D. (1982). *Vision: a computational investigation into the human representation and processing of visual information*. San Francisco, CA: W.H. Freeman.
- McClelland, J. L., & Rumelhart, D. E. (1986). *Parallel distributed processing: explorations in the microstructure of cognition*. Cambridge, MA: MIT Press.
- Metelli, F. (1974). The perception of transparency. *Scientific American*, 230 (4), 90–98.
- Moore, C. M., & Egeth, H. (1997). Perception without attention: evidence of grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception and Performance*, 23 (2), 339–352.
- Nakayama, K., He, Z., & 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. An invitation to cognitive science* (2nd ed., Vol. 2, pp. 1–70). Cambridge, MA: MIT Press.
- Nakayama, K., & Josephs, J. S. (1998). Attention, pattern recognition, and pop-out in visual search. In R. Parasuraman (Ed.), *The attentive brain* (pp. 279–298). Cambridge, MA: MIT Press.

- Nakayama, K., & Shimojo, S. (1992). Experiencing and perceiving visual surfaces. *Science*, 257 (5075), 1357–1363.
- Nakayama, K., Shimojo, S., & Silverman, G. H. (1989). Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception*, 18 (1), 55–68.
- Neisser, U. (1967). *Cognitive psychology*. New York: Appleton.
- Neisser, U., & Becklen, R. (1975). Selective looking: attending to visually specified events. *Cognitive Psychology*, 7, 480–494.
- O'Regan, J. K., Rensink, R. A., & Clark, J. J. (1999). Change-blindness as a result of 'mudsplashes'. *Nature*, 398 (6722), 34.
- Palmer, S. E., Neff, J., & Beck, D. (1997). Grouping and amodal completion. In I. Rock (Ed.), *Indirect perception* (pp. 63–75). Cambridge, MA: MIT Press.
- Palmer, S. E., & Nelson, R. (in press). Late influences on perceptual grouping: illusory contours. *Perception & Psychophysics*.
- Palmer, S., & Rock, I. (1994). Rethinking perceptual organization: the role of uniform connectedness. *Psychonomic Bulletin & Review*, 1 (1), 29–55.
- Pashler, H. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.
- Peterson, M. A. (1994). Object recognition processes can and do operate before figure-ground organization. *Current Directions in Psychological Science*, 3, 105–111.
- Petry, S., & Meyer, G. E. (1987). *The perception of illusory contours*. New York: Springer-Verlag.
- Polat, U. (1999). Functional architecture of long-range perceptual interactions. *Spatial Vision*, 12 (2), 143–162.
- Polat, U., & Sagi, D. (1993). Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiments. *Vision Research*, 33 (7), 993–999.
- Polat, U., & Sagi, D. (1994). The architecture of perceptual spatial interactions. *Vision Research*, 34 (1), 73–78.
- Pomerantz, J. R., & Kubovy, M. (1986). Theoretical approaches to perceptual organization: simplicity and likelihood principles. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Cognitive processes and performance. Handbook of perception and human performance* (Vol. 2, pp. 1–46). New York: Wiley.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: evidence for a parallel tracking mechanism. *Spatial Vision*, 3, 179–197.
- Ramachandran, I. (1987). Visual perception of surfaces: a biological theory. In S. Petry, & S. E. Meyer (Eds.), *The perception of illusory contours*. New York: Springer-Verlag.
- Ramachandran, V. S. (1995). Filling in the gaps in logic: reply to Durgin et al. *Perception*, 24, 841–843.
- Rensink, R. A. (2000). Seeing, sensing, and scrutinizing. *Vision Research*, 40, 1469–1487.
- Rensink, R. A., & Enns, J. T. (1998). Early completion of occluded objects. *Vision Research*, 38, 2489–2505.
- Rensink, R. A., O'Regan, J. K., & Clark, J. (1997). To see or not to see: the need for attention to perceive changes in scenes. *Psychological Science*, 8 (5), 368–373.
- Rock, I. (1975). *An introduction to perception*. New York: Macmillan.
- Rock, I. (1983). *The logic of perception*. Cambridge, MA: MIT Press.
- Rock, I., & Gutman, D. (1981). The effect of inattention on form perception. *Journal of Experimental Psychology: Human Perception and Performance*, 7 (2), 275–285.
- Rock, I., Linnett, C. M., Grant, P., & Mack, A. (1992). Perception without attention: results of a new method. *Cognitive Psychology*, 24, 502–534.
- Rock, I., Nijhawan, R., Palmer, S., & Tudor, L. (1992). Grouping based on phenomenal similarity of achromatic colour. *Perception*, 21 (6), 779–789.
- Rubin, E. (1921). *Visuell Wahrgenommene Figuren*. Kobenhaven: Glydenalske Boghandel.
- Rubin, N., Nakayama, K., & Shapley, R. (1996). Enhanced perception of illusory contours in the lower versus upper visual hemifields. *Science*, 271 (5249), 651–653.
- Scholl, B. J. (2001). Objects and attention: the state of the art. *Cognition*, this issue, 80, 1–46.
- Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multi-element tracking. *Cognition*, this issue, 80, 159–177.
- Simons, D. J. (2000). Current approaches to change blindness. *Visual Cognition*, 7, 1–16.
- Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, 1 (7), 261–267.

- Smilek, D., Eastwood, J. D., & Merikle, P. M. (2000). Does unattended information facilitate change detection? *Journal of Experimental Psychology: Human Perception and Performance*, 26, 480–487.
- Somers, D. C., Dale, A. M., Seiffert, A. E., & Tootell, R. B. (1999). Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proceedings of the National Academy of Sciences USA*, 96 (4), 1663–1668.
- Spillmann, L., & Werner, J. S. (1996). Long-range interactions in visual perception. *Trends in Neuroscience*, 19 (10), 428–434.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662.
- Treisman, A. (1986). Features and objects in visual processing. *Scientific American*, 255 (5), 114B–125B.
- Treisman, A. M., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, 12 (1), 97–136.
- Usher, M., Bonneh, Y., Sagi, D., & Herrmann, M. (1999). Mechanisms for spatial integration in visual detection: a model based on lateral interactions. *Spatial Vision*, 12 (2), 187–209.
- Van Tuijl, H. F., & de Weert, C. M. (1979). Sensory conditions for the occurrence of the neon spreading illusion. *Perception*, 8 (2), 211–215.
- Warm, J. S., Dember, W. N., Padich, R. A., Beckner, W. N., & Jones, S. (1987). The role of illumination level in the strength of subjective contours. In S. Petry, & G. E. Meyer (Eds.), *The perception of illusory contours* (New York: Springer-Verlag).
- Wertheimer, M. (1923). Untersuchungen zur Lehre von Gestalt. *Psychologische Forschung*, 4, 301–350.
- Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *The psychology of attention*. Hove: Psychology Press.
- Wolfe, J. M. (1999). Inattentional amnesia. In V. Coltheart, *Fleeting memories: cognition of brief visual stimuli* (pp. 71–94). Cambridge, MA: MIT Press.
- Zenger, B., & Sagi, D. (1996). Isolating excitatory and inhibitory nonlinear spatial interactions involved in contrast detection. *Vision Research*, 36 (16), 2497–2513.