Endogenous and Exogenous Attention Recover Target Identity During Crowding

by Jeffrey Nador

B.A. Psychology, Concordia University, Montreal QC
M.A. Psychology Northeastern University

A dissertation submitted to

The Faculty of
the College of Science of
Northeastern University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

August 11th, 2015

Dissertation Directed by

Adam Reeves
Professor of Psychology
Abstract of Dissertation

The environment constantly bombards us with information, but with limited processing capacity, we can only select a fraction of it for further processing. Broadly defined, this selection forms the basis of visual attention, and it can be further divided into endogenous and exogenous subtypes. Exogenous attention involves selecting information based on the salience of external stimuli, while endogenous attention involves selecting information based on the one’s internal state.

Attentional selection is of paramount importance in peripheral vision, wherein target objects become unidentifiable if too densely cluttered by other flanker objects. This phenomenon, called visual crowding, spreads over larger areas of the visual field as target eccentricity increases. The minimum spacing at which flankers leave target identity intact is called critical spacing, extending over distances up to half the target’s eccentricity (Bouma, 1970).

Previous crowding research suggests that the greater the resemblance between target and flanker features (e.g. orientation, or spatial frequency), the harder target identification becomes (Kooi et al., 1994). But, while flankers obscure target identity, salient feature differences between them remain perceptible (Petrov & Popple, 2007). This has important implications for attentional selection: before the visual system can recognize a flanked target, it must be able to covertly select target-relevant information. Plus, it must do so without knowing what features pertain to the target a priori.

Jonides & Yantis (1988) showed that targets revealed by abrupt onsets capture exogenous attention, and are thus processed more faithfully than those revealed by removing surrounding features. This processing advantage has been extended to include
crowded targets (Harrison & Bex, 2014) such that when a target appears after its flankers, its isolated onset attracts exogenous attention to its features. Yet, when it appears simultaneously with the flankers, they also capture exogenous attention, thus leaving only endogenous attention with the capacity to preselect the target’s features.

Although Harrison and Bex (2014), as well as Greenwood, Sayim and Cavanagh (2014) have both demonstrated that exogenous attentional cues can reduce crowding, neither controlled observers’ endogenous attentional states. Thus, it remains an open empirical question whether transient onsets can relieve crowding in the absence of endogenous attentional selection of the target.

As such, the present series of experiments addresses the role of exogenous attention in reducing crowding during endogenous attentional cueing. We first hypothesize that in the absence of any other salient target features, exogenous attentional cueing alone is insufficient to constrain endogenous attention to the target. Second, we hypothesize that target-flanker feature contrast reduces crowding when the target is exogenously cued and, conversely, when exogenous spatial attention is directed to the flankers, crowding should be exacerbated more at low levels of target-flanker feature contrast.

On each trial, observers first completed an attentional task that endogenously cued them to shift their attention to one of two flanked targets, whose orientation they would then have to identify. At random stimulus onset asynchronies (SOA), we produced abrupt onsets by briefly blinking off the targets, flankers, or both. Trials where blinks occurred prior to the first task’s cue (negative SOAs) allowed us measure their effect on target identification, absent endogenous attention to the target. Trials with positive SOAs
allowed us to measure the duration of any combined effects of endogenous and exogenous attentional cues on target identification. And finally, trials where no blink was applied allowed us to measure the isolated effect of endogenous attention to the target.

At SOAs < 0, we find no effect of abrupt onsets in any blink condition, suggesting that exogenous attention to the target alone cannot modulate crowding. Additionally, blinking the targets from 0 to +100 ms SOA improves target identification, suggesting that the combination of endogenous and exogenous attention to the target can reduce crowding. Moreover, we find that the magnitude of this effect depends on feature similarity along the attended feature (orientation), as it reduces the benefit conferred by blinking the target. We also find that blinking the flankers at positive SOAs hinders identification only when the flankers are similar. Taken together, these findings suggest that a combination of endogenous and exogenous attentional selection of the attended target feature is necessary to recovering target identity.

Comparison of naïve and experienced observers reveals that the effects of endogenous and exogenous attention depend strongly on learning. Experienced observers can use unattended features (spatial frequency) to obtain generalized gains in performance, whereas naïve observers cannot, implying that pre-attentive feature selection during crowding can be learned. Plus, critical spacing seems to increase under attentional load following substantial practice, suggesting that practice may change the strategies that experienced observers use to identify peripheral objects.
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract of Dissertation</td>
<td>i</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>iv</td>
</tr>
<tr>
<td>List of Figures</td>
<td>v</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Chapter 1: Experiment 1 – High vs. Low Attended Feature Contrast</td>
<td>18</td>
</tr>
<tr>
<td>Chapter 2: Experiment 2 – Attended vs. Unattended Feature Contrast</td>
<td>41</td>
</tr>
<tr>
<td>Chapter 3: Experiment 3 – Attentional effects on critical spacing</td>
<td>46</td>
</tr>
<tr>
<td>Chapter 4: Experiment 4 – Replication with naïve observers</td>
<td>56</td>
</tr>
<tr>
<td>Chapter 5: Experiment 5 – Replication of Critical Spacing</td>
<td>63</td>
</tr>
<tr>
<td>Chapter 6: Conclusions and Recommendations</td>
<td>67</td>
</tr>
<tr>
<td>References</td>
<td>73</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1 – Crowding Calibration Stimulus 23
Figure 2 – Letter Identification Calibration Stimulus 35
Figure 3 – Calibration Experiment Results 26
Figure 4 – Experiment 1 Dual Task Stimulus 30
Figure 5 – Experiment 1 Individual Observer Performance Across SOAs 33
Figure 6 – Experiment 2 Averaged Performance Across SOAs 34
Figure 7 – Experiment 1 Attention Operating Characteristic 37
Figure 8 – Experiment 2 Averaged Performance Across SOAs 44
Figure 9 – Experiment 3 Individual Observer Performance Across SOAs 52
Figure 10 – Experiment 4 Calibration Experiment Averaged Results 58
Figure 11 – Experiment 4 Averaged Results Across SOAs in All Conditions 59
Figure 12 – Experiment 4 Averaged Attention Operating Characteristic 60
Properties of Visual Crowding

The visual crowding effect generally refers to reduced discriminability of a target object resulting from the presence of nearby distractor objects (flankers). This effect was first characterized by Korte in describing his perception of words presented in his visual periphery:

“It is as if there is a pressure on both sides of the word that tends to compress it. Then the stronger, i.e. the more salient or dominant letters, are preserved and they “squash” the weaker, i.e., the less salient letters, between them” (Korte, 1923, translated by Uta Wolfe).

Notably, Korte’s description suggests that even the less salient letters are still detected under such circumstances, but somehow compressed or muddled among more salient ones. (Had such objects gone undetected, a report of their being “squashed” would be impossible). Although this characterization does not in itself provide any information about what factors might increase or decrease the saliency or discriminability of such visual objects, it nevertheless implies that their spatial positioning is one such determining factor. i.e. The less salient letters are those muddled between the salient ones. This might suggest that, at least for words, the outermost letters tend to crowd the innermost letters.

Indeed, later research by Bouma (1970, 1973, 1976) showed that Korte’s characterization was particular to words, in that crowding itself operates independently of whether or not the letter stimuli composed words or even pronounceable syllables. He
compared identification performance on the initial and final letters of words and unpronounceable letter sequences. His findings showed that crowding occurred even in non-word sequences, and that the persistent positional effect was not of the target letter’s specific position at the beginning or end of the sequence, but was instead based on its position relative to the other letters and to fixation.

For the same letter sequence, Bouma (1970) found that its initial letter could be more accurately identified to the left of fixation than to the right, whereas the inverse was true of the final letter. This implied that the outermost letters (those furthest from fixation) in a given sequence were more accurately identified than the innermost letters (those closest to fixation). This constituted the first indication that crowding operated over a broader range of stimuli than words, and that it was a visual rather than a lexical effect.

Since this finding, other research has found similar crowding effects for a wide variety of stimuli beyond words, or letter sequences: line segments (Andriessen, J. J. & Bouma, 1976), Verniers (Levi, Klein & Aitsebaomo, 1980; Sayim, Westheimer & Herzog, 2011), Gabors (Petrov & Meleshkevich, 2011(a & b); Petrov, Popple & McKee, 2007), simple luminance- or hue-defined defined disks (van den Berg, Roerdink & Cornelissen, 2007), and oriented bars (Pöder, 2006), among others. Consistently, across experiments, outward flankers produce greater detriments to object recognition than their inward counterparts (Levi, 2008). This inward-outward asymmetry has become something of a litmus test for distinguishing it from other forms of peripheral visual masking.
The spatial extent of crowding (i.e. the target-flanker spacing that produces the effect) is typically measured as the minimum center-to-center separation between target and flanker that produces no threshold elevation (TE) (Levi, 2008). This ‘zone of interaction’ scales linearly with eccentricity, such that flankers positioned within a region centered on the target, whose radius is approximately half the target’s eccentricity, will elevate target discrimination thresholds (Bouma, 1970; Pelli, Palomares & Majaj, 2004).

The spatial extent of crowding is also biased towards the periphery, such that the critical spacing is larger for outward flankers than their otherwise identical inward counterparts (Bouma, 1973; Petrov & Popple, 2007; Petrov & Meleshkevich, 2011b). This inward-outward asymmetry has been explained both by the cortical magnification factor (CMF) in retinotopic visual cortex (Motter, 2006; Motter & Simoni, 2007; Pelli, 2008), and limitations on exogenous spatial attentional resolution (Intrilligator & Cavanagh, 2001; Petrov & Melshkevich, 2011a).

According to the former explanation, inner and outer flankers might be physically equidistant from the target, while their cortical projection shifts the outer flanker inwards, closer to the target (Pelli, 2008; Motter and Simoni 2007). According to the latter, spatial attentional resolution is insufficient to constrain the attentionally selected region of the visual field to the target (Motter, 2006; Cavanagh & Intrilligator, 2001), and is biased outwards (i.e. more peripherally than the target) (Petrov & Meleshkevitch, 2011a).

In addition to inward-outward asymmetry, the zone of interaction is anisotropic (Levi, 2008). Flankers positioned radially (along a radius centered at the fovea, extending outwards through the target) hinder discrimination at larger separations than tangential
ones (along the perimeter of the circle centered on the fovea, whose radius equals the target’s eccentricity).

This anisotropy is unlikely to be a direct result of inward-outward asymmetry, since crowding scales linearly with eccentricity. That is, if no such asymmetry existed, two flankers with eccentricities equal to the target’s (arranged tangentially) would produce an equivalent amount of interference to two flankers with equal separations, but with one shifted inward and the other outward of the target (arranged radially); however, this is not so. Critical spacing is reduced as much as five-fold for tangential flankers (Toet & Levi, 1992), but less than two-fold for inward flankers (Petrov & Meleshkevich, 2011b). Thus, this latter property of crowding is likely distinct from the inward-outward anisotropy.

Such large spatial extents are atypical of other types of masking occurring in peripheral vision (Pelli, Palomares & Majaj, 2004; Petrov, Popple & McKee, 2007). As well, masking generally affects target detection, that is, the amount of target stimulus energy needed to reach threshold. In contrast, crowding has been shown to affect discriminability, the ability to distinguish a number of potential target stimuli from one another, to the exclusion of detectability (Pelli, Palomares & Majaj, 2004; Levi, 2008), at least for 4 or fewer flankers (Wilkinson, Wilson & Elleberg, 1997; Poder, 2008). This suggests that the target-related signal is not completely lost, as is typically the case for masking. Rather, it seems to remain detectable in crowded displays, but obscured or ambiguated by flankers situated within the zone of interaction (Pelli, Palomares & Majaj, 2004) as Korte (1923) had originally proposed.
Petrov and Popple (2007), for example, show that orientation contrast between the target and flankers survives crowding, even though the individual stimulus elements’ orientations become ambiguous. In their experiments, observers matched one of several triads of Gabors, simultaneously presented in the periphery, to a previewed comparator triad. Confusion matrices of observers’ errors reveal that observers tended to confuse triads with equal numbers of sites where orientation changed from one element to the next. In other words, information about the orientation differences between stimulus elements remained salient despite crowding, even though the individual elements’ orientations did not. Thus, it seems that the presence of flankers does not entirely suppress the target signal. Rather, it seems that uncertainty about the individual items’ features (i.e. when there are fewer feature contrasts) renders the signal ambiguous. Crowding again differs from traditional forms of masking in this respect, in that masking entirely suppresses featural information about the target or any contrast between it and the mask.

Crowding seems to operate quite similarly across a wide range of features (van den Berg, Roerdink & Cornelissen, 2007). Generally, it varies with target-flanker feature similarity (Kooi, Toet, Tripathy & Levi, 1994; Pöder, 2007; van den Berg, Roerdink & Cornelissen, 2007; Sayim, Westheimer & Herzog, 2011), such that target-flanker combinations with similar features consistently produce larger TEs – that is, more crowding – than dissimilar combinations (Levi, 2008).

Feature similarity between two stimuli depends on the concordance between their defining dimensions, while differences between their features along those dimensions results in feature contrast. On one hand, when two stimuli share the same dimensions, as
would a pair of Gabors, then similarity between them defaults to feature contrast between their spatial frequencies, amplitudes, envelopes, and phases. And on the other, when stimuli differ in dimensionality, as would a circle and a square, then only some shared features (e.g. area) can vary in similarity, whereas others (e.g. orientation) cannot (Garner, 1974).

While the width of similarity tuning varies substantially across a variety of stimulus dimensions, the maximal TE that crowding produces remain relatively stable. For instance, Levi and Carney (2009) find that orientation and spatial frequency produce similar maximal TE, even though the target Gabors used in their experiments were flanked by ring-shaped sectors (with clearly defined edges) filled-in by sine wave grating patterns. Equally, critical spacing measured within a given dimension increases as a function of eccentricity for orientation, size, saturation and hue with similar slopes – slightly less than 0.5 times the target’s eccentricity (van den Berg, Roerdink & Cornelissen, 2007). However, the functional relationship across dimensions is not known.

Likewise, the magnitude of crowding varies with the number of flankers, and their spatial configuration. Adding more similar flankers generally increases TE, independent of critical spacing, while adding concentric flanking elements to an already-crowded display can reduce the crowding effect when the flankers form a perceptual group that excludes the target (Pöder, 2006; Manassi, Sayim & Herzog, 2011). However, grouping in the latter studies covaried with similarity, in that flankers were substantially different from their targets along the reported target feature, and often along other task-irrelevant features (e.g. color and spatial frequency content). Hence, the notion that crowding operates similarly along numerous feature dimensions (van den Berg, Roerdink &
Cornelissen, 2007) warrants a distinction between target-flanker similarity effects along attended (or task-relevant) and unattended (or task-irrelevant) features.

**Properties of Visual Attention**

The environment constantly bombards the visual system with a plethora of potentially useful information. But, with a limited capacity to process all of it from moment to moment, the visual system must be selective about which available information from among such a vast array it will process further. This allows it to efficiently and accurately create meaningful representations of the physical stimuli in the environment. As a broad and cursory definition, attention can be thought of as the deployment of cognitive processes or resources that select relevant information from the vast array of sensory inputs that the environment provides.

More specifically, attention can be thought of as the selection of information in service of recognizing objects in the visual scene (Yantis, 1998). Before the visual system can recognize an object, it must select information potentially relevant to an object’s identity. Yet, the visual system cannot know *a priori* what the properties of any single object will be, when more than one potential object could occur, and thus cannot select for object-properties in advance.

In order to resolve this dilemma, the process of attentional selection can be subdivided into what William James (1890) called active and passive modes of selection. Active selection occurs when an observer’s internal state orients them towards some subset of the information available in the visual scene. Passive selection, by contrast, refers to the capturing of observers’ attention by some salient aspect of the information in the visual scene that is unrelated to their internal states.
More recently, these two types of attentional selection have been referred to as endogenous attention (generated internally in accordance with the observer’s goals or expectations) and exogenous (generated externally by stimulus salience) attention, respectively. Thus, endogenous attention resolves features based on internal states and exogenous attention passively encodes the most salient features in the environment.

To further clarify this distinction, Yantis (1998) notes the simple example of a shopper in the cereal aisle at the supermarket. In an aisle full of many different colorful boxes, no single one is likely to be particularly salient to the shopper at first blush. Yet it is possible for one or several of the boxes to pop-out from among its neighbors, even in the absence of strong feature contrast. For instance, if the hypothetical shopper happened to be searching for a particular brand of cereal in a green box, and thus endogenously attended to the color green.

On the other hand, attention could be drawn to one of the boxes’ features exogenously even if the observer had no designs on any particular brand of cereal, say, if it happened to be the only yellow box in an aisle full of green boxes. In this case, the salient feature contrast between the green and yellow boxes could draw the shopper’s attention to the markedly different item, regardless of any pre-existing endogenous attentional state.

Jonides and Yantis (1988) showed that offset stimuli, those which are revealed by deleting segments (think of converting an ‘8’ into a ‘5’ by removing the upper right and lower left segments), do not attract attention exogenously, in contrast to the same stimuli presented as onsets, by turning on segments (a ‘5’ from a grey field). Stimuli used in crowding experiments have all been presented as onsets; curiously, it is not yet known
whether crowding can occur with offset stimuli. Though following onset presentation, it seems certain that exogenous attention would be captured equally by both targets and flankers, and so the ability to select the target from among them would require endogenous attention by definition. Nevertheless, some role for exogenous attention likely does exist in determining at least the magnitude of crowding. As an example, feature contrast between flankers and targets could increase target salience and thereby reduce crowding, since feature contrast is preserved during crowding (Petrov & Popple, 2007).

Exogenous spatially localized cues such as an underscore or momentary blink may also act to increase target saliency; again, crowding may be somewhat ameliorated. Harrison and Bex (2014) and Greenwood, Sayim and Cavanagh (2014) used similar manipulations to reduce the deleterious effect of flankers on discriminability of a nearby target. When rapidly (e.g. 50ms) blinking off some subset of stimulus elements in the target-flanker array, Greenwood’s study found improvements in orientation discrimination performance for the target. Similarly, predominantly valid exogenous precues (Scolari, Kohnen, Barton & Awh, 2007; Yeshurun & Rashal) and flanker preview followed by abrupt target onset (Harrison & Bex, 2014) also recover its discriminability.

Endogenous attention to objects in the visual periphery can summon eye movements, typically saccades, so that the object can be foveated. An abrupt presentation of a salient object may attract attention exogenously, and in turn, elicit such a saccade. Similarly, an endogenous attention shift to a target in visual search may also elicit a foveating saccade. Foveal presentation will typically break crowding, since critical
spacing is so much smaller at the fovea (~20 arcmin), and permit the subject to identify the once-crowded target (Levi, Hariharan & Klein, 2002). Thus visual information, if sufficient to indicate that a relevant stimulus might be present, can be useful in the visual periphery, even if the information is crowded. As such, there is some interest in discovering how much information is actually available in the period before a foveating saccade can occur.

Studies of crowding can limit saccades by presenting stimuli briefly (saccades take 200 ms or more) or by controlling fixation in some manner. Until the saccade occurs, attention is said to be ‘covertly’ oriented to the target, that is, not yet overtly drawn to the target as would be demonstrated by a saccade. Since crowding studies rely on the target and flankers being peripheral, any relevant visual attention, endogenous or exogenous, is covert. The properties of covert attention as they apply to the visual periphery are therefore critical in understanding crowding. First, covert attention takes approximately 160 ms time to shift from central vision (Reeves & Sperling, 1986). And second, the spatial resolution of covert attention is limited to 2 cycles per degree or less (Blaser, Sperling & Lu, 1999). The inability of covert attention to fully resolve the target in space, and perhaps in time if the stimulus is brief, is therefore a possible reason for crowding.

Many authors follow Duncan (1984) in distinguishing between feature-based attention and object-based attention. In Driver’s (2001) experiment, line segments belonged to the same or different rectangles; attention was slightly, but significantly, faster to shift within an object (rectangle) than across objects. It therefore becomes of interest to enquire whether any attentional effect on crowding is feature or object based.
Experimentally this distinction is muddied when visual processing involves feedback interactions at multiple levels (Ahissar and Hochstein, 2002), rather than a simple progression from features to objects as once thought, since a disruption at either level can then affect processing at the other level. However, in this dissertation, the orientation contrast between like objects (Gabor of the same spatial frequency) was varied in order to test the role, if any, of feature-based exogenous attention. Note that the participant’s internal state can be set to endogenously attend either features or objects, while stimulus differences at either level would vary salience and attract attention exogenously, so in principle these two dimensions of attention are orthogonal.

**Attentional Theories of Visual Crowding**

Across studies, crowding has been alternately explained as arising from some combination of limited spatial resolution of attention (Intrilligator & Cavanagh, 2001), compulsory feature integration (Parkes, Lund, Angelucci, Solomon & Morgan, 2001; Chakravarthi & Cavanagh, 2009b; Greenwood, Bex & Dakin, 2009) and object substitution (Greenwood, Bex & Dakin, 2009), or grouping of target and flanker elements (Herzog & Fahle, 2002; Sayim, Westheimer & Herzog, 2011; Manassi, Sayim & Herzog, 2012).

Some research suggests that insufficiency of endogenous attentional spatial resolution accounts for the magnitude (He, Cavanagh & Intrilligator, 1996) and spatial extent (Intrilligator & Cavanagh, 2001) of crowding. When covertly (i.e. without eye movements) tracking multiple objects attentively, observers’ performance decays towards chance at greater target-flanker separation than would be predicted from the limits of visual acuity alone. Furthermore, when tracking overtly (with voluntary, endogenously
initiated eye-movements), targets of saccadic eye movements are briefly uncrowded. When these same saccades are prevented, this effect is reduced or destroyed (Harrison, Mattingley & Remington, 2013). Thus, it seems that the endogenous orienting of attention in preparation for eye movements. Intrilligator & Cavanagh (2001) suggest that crowding occurs when flankers fall within an area surrounding the target that is smaller than covert endogenous spatial attention is capable of resolving. Thus, they propose that crowding occurs when several similar objects occupy such an area; when the objects can no longer be tracked, individuation of any single one becomes impossible.

There is also evidence for a bilateral field advantage in crowding, such that when two exogenously precued locations (each containing a target surrounded by flankers) are situated in opposite hemifields, target identification is better than when they are situated in the same hemifield (Chakravarthi & Cavanagh, 2009b). Notably, this advantage disappears in the absence of flankers. This implies that crowding is exacerbated by the impairment of attentional selection in the presence of flankers that fall within the attentionally selected region of the visual field.

Petrov and Meleshkevich (2011a) propose that crowding’s inward-outward asymmetry depends on the spatial allocation of attention. They find that when target and flanker stimuli are displayed in precisely the same location on every trial, inward and outward flankers induce similar TEs. However, the addition of spatial uncertainty to the target’s location (holding flanker locations constant) produces the typical asymmetry. This suggests that spatial uncertainty about the attended location of the target contributes to the crowding effect. Inward-outward anisotropy only occurs when the target’s location
is ambiguous; when the target’s location is constant trial-to-trial, observers can constrain the attentionally selected region to exclude the flankers.

Explanations of crowding in terms of attention focus mainly on its spatial distribution, yet there is some evidence for feature-based influences on spatial deployment of attention. For instance, spatial selection of an endogenously cued target (by a series of tones) improves for large feature contrasts (color or luminance) between the cued target and other uncued stimulus elements (Moore, Lanagan-Leitzel, Chen & Halterman, 2007). Equally, the magnitude of crowding is reduced when the target and flankers reverse contrast in counter-phase, but only within the temporal limits of phase selection (Chakravarthi & Cavanagh, 2007). When the flicker becomes fast enough that selection of phase becomes impossible (despite phase differences in the distal stimulus), target-flanker contrast polarity differences no longer reduce crowding. So, in addition to endogenous spatial attention, it seems that endogenous feature selection can also influence the distribution of exogenous spatial attention by constraining it to the target’s salient features.

Several studies have found that flankers’ effects on target identification depend on global stimulus configuration – specifically on grouping of target and flanker elements (Pöder, 2006; Livne & Sagi, 2007; Sayim, Westheimer & Herzog, 2011; Manassi, Sayim & Herzog, 2012). A given target remains unidentifiable when its flankers group together with it; yet, when flankers group separately from it, its identity is recovered (Manassi, Sayim & Herzog, 2012). This pattern of results is stable across a number of Gestalt principles of grouping, including similarity (Manassi, Sayim & Herzog, 2012), closure
(Mathes & Fahle, 2007), continuity (Livne & Sagi, 2007), and common fate (Chakravarthi & Cavanagh, 2007; Levinthal & Franconeri, 2011).

This finding has also been extended to figure ground segmentation (Põder, 2006) and texture perception (Parkes et al., 2001). When a feature-singleton is surrounded by flankers, increasing the number of concentric flankers decreases the crowding effect. Assuming that spatial attention processes all salient feature singletons (i.e. elements differing from those surrounding them), then exogenous feature-based spatial attentional selection of the target persists despite low spatial attentional resolution (Põder, 2006). If instead a target is crowded within an array of similar flankers (with no singleton defining the target), observers’ assessments of the array’s average orientation would include the target’s orientation (Parkes et al., 2001). This suggests that distinctive target features are not “lost” during crowding, but that they might instead be pooled together with surrounding flankers’ features to form a texture that makes up the background. The mechanism for target identification among clutter thus likely includes a combination of attentional and pooling mechanisms, which could also be involved in grouping, texture perception and figure ground segmentation.

There is some evidence from imaging studies to suggest that exogenous spatial attention can influence target-related neural responses. Electro-encephalography (EEG) studies in human observers find that steady state visually evoked potentials (SSVEP) are greater for attended stimulus locations than unattended ones, even when they are task irrelevant (Hillyard, et al., 1996). Similar results exist for oddball feature searches within the same spatial locations, in an eccentricity dependent manner (Ding, Sperling &
Srinivasan, 2006). These results support the combination of spatial and feature attention in enhancing both target- and flanker-related SSVEPs.

Harrison & Bex (2014) and Greenwood, Sayim and Cavanagh (2014) both demonstrated psychophysically that onset transients applied only to the target in a crowded display recover target discriminability. Onset transients are also strongly implicated in directing exogenous spatial attention during visual search (Yantis & Jonides, 1974) and change detection (Phillips & Singer, 1974) tasks. However, prior research examining the effects of such transient onsets on the spatial extent of crowding is split. Along with Greenwood and colleagues, Scolari, Kohnen, Barton & Awh (2007) report that exogenous spatial cueing does not affect the spatial extent of crowding (i.e. the critical spacing), whereas Yeshurun and Rashal (2010) and Harrison and Bex (2014) both find reductions in critical spacing due to exogenous attentional cueing.

This discrepancy could be the consequence of a lack of control or manipulation of endogenous attention during exogenous spatial cueing. Neumann (1996) posits that endogenous and exogenous modes of attention are not mutually exclusive to one another. The end goal of any attentional process is to select information for further processing, and in this sense both endogenous and exogenous attention must work in tandem towards this common goal. For instance, the two studies by Yeshurun and Rashal (2010) and Scolari and colleagues (2007) report using exogenous pre-cues with variable validity. However, in both cases observers were instructed beforehand as to the cue’s predictive validity – providing them with an endogenous motivation to use the cue in their selection processes during the experiment. In the absence of such instruction, it is unclear whether observers would endogenously attend these cues in quite the same way. Harrison and Bex
and Greenwood and colleagues employed abrupt onset transients as exogenous attentional cues. However, given the task demand of directing one’s attention to the target stimuli, their observers would presumably be motivated to do so endogenously, even in the absence of any exogenous cueing manipulation. It is unclear whether similar results would have been obtained if endogenous attention were directed away from the transient at the time of its application. Thus, the lack of manipulation or control of endogenous attention across studies could in principle account for the inconsistent effects of exogenous spatial cues on critical spacing.

However, in the case of Greenwood and colleagues (2014), the targets’ and flankers’ orientations were randomized across trials using the method of constant stimuli, and therefore do not explicitly control for target-flanker similarity. Rather, under such variable conditions, the target and flankers were quite likely to be substantially different from one another in orientation. It is therefore possible that with more similar flankers, the improvements due to transients would have been reduced or completely expunged, as would be predicted by explanations of crowding based on grouping.

Taken together, these results and those from the aforementioned EEG studies could imply that crowding occurs when neither the target’s onset nor its features sufficiently distinguishes it from its flankers, thereby obscuring its identity. On the other hand, if the target were to differ from its flankers, exogenous attention to the transient onset or to the feature pop-out could in principle supplement endogenous spatial attention in order to resolve the target’s identity. We propose such an attentional mechanism for un-crowding, wherein the attentionally selected region is constrained by task-relevant feature contrast between the target and its flankers.
General Hypotheses

In a previous study employing EEG (Nador, 2012), we found that the amplitude of flanker-related steady-state visually evoked potentials (SSVEP) was reduced when target and flankers had orthogonal orientations, as compared to parallel. We observed correlated decreases in flanker-related SSVEP when observers were instructed to attend away from the peripheral stimulus array, ignoring the target and flankers altogether. These results can be explained by a model wherein endogenous control over restriction of the attentionally selected region only isolates the target when it has sufficient feature contrast with respect to its flankers.

We suspect that the correlated decreases in flanker-related SSVEPs we observed previously between un-crowded and un-attended flanked targets might arise from a common mechanism: directing endogenous spatial attention away from the flankers. Both endogenous spatial (Hillyard et al., 1996) and (presumably also endogenous) feature (Ding, Sperling & Srinivasan, 2006) attention increase steady-state visual evoked potentials (SSVEP). Hence, whether attention is directed purely to the target, or directed away from the entire stimulus array, the resulting flanker-related SSVEP is comparably reduced.

Previous research suggests that crowding is exacerbated when the attentionally selected region cannot be restricted to the target (Intrilligator & Cavanagh, 2001). We propose that if the target differs from its flankers on some dimension, feature attention can in principle supplement spatial attention to restrict the attentionally selected region to
areas of high feature saliency. Thus, in our previous experiment, just as directing spatial attention away from the crowded display reduced flanker-related SSVEPs, providing a salient target feature decreased the amount of spatial attention devoted to the surrounding flankers. Similarly, Greenwood et al (2014) recover target identification by providing onset transients at the target’s location that serve to capture spatial attention, though they did not explicitly control the feature contrast between their targets and flankers.

The aim of the following studies is to corroborate our previous EEG results psychophysically, as well as to determine the limiting conditions within which focusing attention on the target might reduce crowding. First, we hypothesize that in the absence of endogenous attention, exogenous attention is insufficient to constrain the attentionally selected region to the target. Second, given endogenous attention, target-flanker feature contrast reduces crowding when exogenous spatial attention is directed toward the target, and conversely when exogenous spatial attention is directed to the flankers, crowding should be exacerbated more at low than high levels of target-flanker feature contrast.

Chapter 1

Experiment 1- High vs. Low Attended Feature Contrast

This first experiment aims to test whether endogenously generated shifts of attention, exogenous spatial cueing, or some combination thereof can effectively disambiguate the identity of an otherwise crowded target. Based on our general hypotheses, we predict that when targets and flankers are similar, the absence of feature contrast between them will prevent spatial attention from isolating the target, leaving its
orientation crowded. However, since exogenous spatial cues can induce transient responses facilitating deployment of spatial attention, we hypothesize that application of a brief blink to endogenously attended targets could briefly reduce the effects of flankers on identification.

In a similar set of experiments, Greenwood, Sayim and Cavanagh (2014) found that applying an onset transient to a crowded target by briefly blinking it off and then on again at a random time following initial stimulus onset could partially recover observers’ ability to identify it as compared to an equivalent unflanked target. Meanwhile, they found that when these same transient blinks are applied to the flankers instead, this manipulation yields no improvement in target identification. Finally, they also found no improvement in target identification when the transients are applied to all stimulus elements (targets and flankers) simultaneously.

Since onset transients attract exogenous spatial attention, it is possible that only blinks constrained to the target (that isolate it among flankers) provide any benefit to target identification. Thus, even though blinking the flankers in fact provides a brief physical isolation of the target, the perceptual consequence is that attention is diffused over the area within which the transients occur, thereby biasing attention away from the target, towards the flankers.

In addition to replicating the findings of Greenwood and colleagues (2014), we here assess the time-course of this effect. First, by manipulating the initial state of endogenous attention’s focus, we can determine the stimulus onset asynchronies (SOA) that yield improvements in identification. Second, by comparing positive (precued) and negative (post-cued) SOAs, we can assess whether processing of the transient requires
endogenous attention to the flanked target as a precondition for generating such improvements.

If application of transients to the targets alleviates crowding regardless of the presence of endogenous spatial attention, then we should find an increase in performance when transients are applied only to the targets. Conversely, if endogenous spatial attention is a necessary precursor to generating the benefits observed by application of onset transients, then only simultaneously or pre-cued targets should acquire the benefit conferred by the transients.

Method – Preliminary Experiments

Participants

Five experienced psychophysical observers with normal or corrected normal vision (20/20) participated in the present experiment. One observer is an author; the other four are undergraduate students all with previous coursework and lab experience in Psychology. Informed consent was obtained verbally and in writing from the four undergraduate observers; they received a full explanation of each experiment prior to confirming their participation. Each one additionally completed a Snellen chart test of visual acuity prior to their participation, in order to ensure that they had normal (20/20) visual acuity. All experiments conducted in the present series of studies were approved by the Northeastern University IRB.

Procedure

Before observers completed each block of the main experiment, which consisted of a dual-task paradigm, the two component tasks (an orientation discrimination task and a letter identification task) were assessed independently in order to equilibrate
performance across observers during blocks of the dual-task experiment. Following this procedure, baseline performance thresholds (75% correct letter and tilt discrimination) for each observer could be estimated independently for each task and for each block of trials a priori.

In addition, adjustment of stimulus parameters between blocks following each thresholding experimental session prevented any block-to-block learning effects from confounding the resulting performance in the main experiment. As observers became more proficient at completing the tasks, we could recalibrate each of them to match their performance gains prior to the run’s beginning, to ensure that task difficulty remained relatively constant with respect to the observer’s ability.

*Orientation Discrimination Task.* In the orientation discrimination task, observers were asked to maintain their fixation on a small centrally presented cross. They were further instructed to use their peripheral vision to discriminate the tilt of a flanked target Gabor from vertical as quickly and accurately as possible, by pressing the mouse button corresponding to the observed direction of the target’s tilt. They pressed the right mouse button to indicate clockwise (CW) tilts and the left mouse button to indicate counterclockwise (CCW) tilts. Each observer also ran one uncrowded block of trials in the main experiment to ensure that their performance indeed reflected their ability to identify the target despite some intermediate level of crowding.

The target’s and the flankers’ tilts were systematically varied from trial to trial following the method of constant stimuli. On any given trial, all stimulus elements were given an equal magnitude of tilt, but the directions in which they were tilted (CW vs. CCW) were fully randomized. This way there was no explicit relationship between the
target and flanker tilts across trials (besides equal magnitude on any single trial), and so observers could not use the flanks as reliable comparators in evaluating the target’s tilt. Furthermore, any potential extraneous effects of target-flanker contrast or grouping by similarity across trials would be mitigated by this randomization.

Tilts were applied to the target following the method of constant stimuli, and all trials were randomized within each block. At the end of each block, the level of tilt corresponding to the 75% correct threshold performance was evaluated for the observer, and fixed at that level in the main experiment. Once again, this was done to equilibrate performance across observers in the main experiment.

**Letter Identification Task.** In the letter identification task, observers were asked to signal the presence of one of two possible target letters, L or R, within a RSVP sequence of Ls and Rs. On each trial, letters in the sequence were initially presented in white and, after a random number of letters had passed, the remainder of the sequence was presented in black. The target letter was always the last letter presented in white. Thus, the first black letter was essentially a temporal cue that notified observers of the target letter’s passing. This forced observers to carefully attend to each of the white letters presented in the sequence.

It is important to note that at the time of each white letter’s presentation, the observer could not have known whether it would be designated as the target until the following letter had appeared. Therefore, the observers’ ability to complete the task was dependent on their ability to maintain their attention to each white letter lest they risk missing its identity at the time of the first black letter’s appearance. In other words, all white letters had to be attended carefully, since any one of them could potentially have
been signaled as the target letter by the subsequent letter’s appearance in black. Therefore, failure to attend to each white letter would increase the probability of missing the target.

The speed of letter presentation was varied systematically from trial to trial, following the method of constant stimuli. As in the tilt discrimination experiment, the presentation speed corresponding to the 75% performance threshold was noted, and later entered as a fixed parameter for that aboserver in the next block of the main experiment.

**Stimuli**

In both the tilt discrimination experiment and the letter identification experiment, all stimuli were presented against a neutral gray background at 42 cd/m² on a ViewSonic CRT computer monitor at 1280 x 1024 pixel resolution with 60 Hz vertical refresh rate (i.e. 60 video frames per second). Stimulus display synchronization was accomplished by measurement of the monitor’s vertical raster sweep to ensure accurate display timing. Each trial began with the appearance of a fixation cross subtending 0.3 deg of visual angle, displayed at the center of the monitor.

![Figure 1](image.png)

**Figure 1.** Schematic of Orientation Calibration stimuli. Fixation cross at center; target Gabor at 8° eccentricity, with flankers positioned at center-to-center separation from the target of 2.5°. Stimuli not drawn to scale.
After a 1000 ms delay, the stimuli would appear, and each stimulus presentation lasted 1500 ms. Note that this stimulus duration in principle permits a saccade – an overt shift of attention – to the cued side of the visual field. However, subjects were asked not to shift their eyes, and, critically, the stimulus manipulations of interest all occurred during the period of a covert attention shift, before any saccade could be executed, as explained below.

**Orientation Discrimination Task.** Target and flanker Gabors always appeared to the right of fixation on the horizontal meridian. The target was at 8 deg eccentricity, flanked inwards and outwards, at center-to-center spacing of 2 deg (or a factor of 0.25 target eccentricity). Each Gabor element was tilted either left or right, randomly, by an angle between 1 and 6 degrees, in steps of 1 degree (See Figure 1 for an exemplar of the stimulus). Tilt magnitude was randomized across trials; all were of equal magnitude within trials, varying only in sign (CW vs. CCW).

**Letter Identification Task.** Following the fixation cross’ disappearance, letters were displayed at the center of the screen, where the cross had been, in a RSVP sequence. Initially, a random number of letters were displayed in white, and afterwards, the remaining letters were displayed in black. The inter-stimulut-interval (ISI) between any 2 letters (black and white) varied from trial to trial, between 50 ms and 250 ms in steps of 50 ms, following the method of constant stimuli. On any given trial, all ISIs were identical (See Figure 2 for a schematic of a sample trial’s progression).
Results

Psychometric function fitting

The proportion of correct trials was tabulated separately for observer, at each angle of tilt. Psychometric functions were then fitted to the data by estimating the least-squared error about the slope at the function’s inflection point, and the 75% correct threshold.

For all but one observer (MM), the values returned by the CDF allowed for linear estimation of the 80% criterion performance that was later entered as a fixed parameter in the main experiment (Figure 3). The estimated value was always rounded to the nearest integer because, in the letter identification task, the estimated threshold was measured in video frames. Therefore, it would have been impossible for the computer to display any non-integer number of frames.

As well, for all observers except MM (whose performance was relatively stable at 72%), the linear fits accounted for more than 60% proportions of the variation in performance (Figure 3). At least for single-task performance, this suggests that the actual
performance values for each subject correspond well to the linear estimates of slope, meaning that their performance varies predictably as a function of the stimulus parameter, and their psychometric curves were not at asymptote over the measured range.

Figure 3. Linear fits of Letter Identification (left column) and Orientation Discrimination (right column) calibration data. Each row corresponds to one observer’s data. Note: every 3 video frames corresponds to 50 ms
Method – Experiment 1

Procedure

Experiment 1 used a dual-task paradigm on each trial of which observers would first need to identify a target letter (either L or R) by clicking the corresponding mouse button, and then identify the orientation of a peripherally presented flanked Gabor target from vertical by pressing the left mouse button for CCW tilts and the right one for CW tilts.

In order to ensure comparable performance across observers and minimize practice effects, the two tasks (L/R letter identification and Gabor orientation discrimination) were calibrated separately for each observer prior to completion of each block of trials in the experiment. The experiment was run in four blocks, each one consisting of the two single-task calibrations followed by one block of the dual-task experiment.

In the letter identification calibration, observers completed the 1-back RSVP task alone, in order to determine the inter stimulus interval (ISI) yielding 80% correct performance (Figure 3). This was estimated by systematically varying the RSVP’s speed according to the method of constant stimuli, between 50 and 250 ms per letter. The obtained threshold was then entered as a fixed parameter in the corresponding block of the dual-task experiment.

Similarly, in the orientation discrimination calibration, target and flanker Gabors’ tilts were systematically varied between 1° and 6° according to the method of constant stimuli to yield the threshold orientation contrast between target and flankers producing 80% correct performance. The direction of tilt applied to the flankers was randomized.
with respect to the target’s, such that equal and opposite flanker tilts were equally likely. However, the magnitude of tilt applied to either flanker was always equal to that applied to the target.

In the dual-task experiment, the central letter targets served as endogenous spatial-attentional cues, signaling observers to shift their attention to the left (if “L” was perceived) or right (if “R” was perceived) of fixation as quickly as possible. Meanwhile, the two Gabor targets were presented throughout each trial, with one on either side of fixation. Observers were instructed to report the orientation (CW vs CCW) of the target to the side of fixation corresponding with the perceived target letter (left side for L, and right for R). That is, if they chose L as the perceived target letter, they were to shift their attention to the left of fixation and report the orientation of the left side Gabor target and vice-versa (See Figure 4 for schematic of stimuli). They were also explicitly instructed to maintain fixation at the center of the screen throughout each trial, even after they had identified the target letter.

**Stimuli**

In the dual-task experiment, each trial began with the presentation of a fixation cross at the center of the screen on a neutral gray background for 750 ms. Immediately following its disappearance, a RSVP of letters L and R, in random order, would appear for 17 ms each, followed by an ISI determined as described in the Calibration Experiments’ Procedures. The letters initially appeared in white and, after a random delay (250 to 1250 ms), the remaining letters would appear in black. The target letter was always the last one presented in white, and the cue letter was the first one presented in black. Each letter
subtended 0.3° visual angle in diameter or less. Each trial lasted a total of 1500 ms, followed by a 1000 ms ITI during which the central fixation cross reappeared.

Meanwhile, two target Gabors were presented at 8° eccentricity throughout each trial; one was presented in each hemi-field, both on the horizontal meridian, flanked by 2 Gabors (all $\lambda = 0.67^\circ$, $\sigma = \frac{\lambda}{2}$) at 2.5° visual angle of center-to-center separation from the target, with one flanker inwards of the target and one outwards, as described in Figure 1.

Since observers reported the identity of the last letter displayed in white (Figure 2), as quickly and accurately as possible, this ensured that they carefully attended to each white letter presented, since every white letter could potentially be followed by the first black letter. Randomization of the RSVP cue’s appearance would prevent fixating either peripheral target. Since the RSVP task requires a 1-back judgment and the cue’s onset (letter contrast reversed from white to black) is randomized within blocks, participants fixating elsewhere would need to guess the identity of the previous letter even if they could accurately detect the contrast reversal peripherally.

On each trial, a subset of the Gabors on each side of fixation could ‘blink’ off for 50ms and reappear (Figure 4), with a systematically varied stimulus onset asynchrony (SOA) relative to the RSVP cue. Blink SOAs were randomized from trial-to-trial within blocks, in 5 steps of 100ms, from -200 to 200 ms. Negative SOAs indicate the cue letter’s arrival before the blink. SOA randomization ensured that observers could not develop any expectancy for blink occurrence, preventing their use as endogenous attentional cues to the onset of the RSVP cue.
Figure 4. Diagrammed exemplar of a trial in Experiment 1. Here, the letter target is an R, and the Gabor target blinked for a duration of 50ms, at +200ms SOA with respect to the cue letter. The target letter was always the last letter presented in white and the cue letter was always the first letter presented in black. Note that the sequence of Ls and Rs on each trial was completely random.

Additionally, we randomized flanker similarity across blocks. Within any single block, observers would see either only flanker Gabors with similar (nearly parallel) or
dissimilar (nearly orthogonal) orientations to the target Gabor at all 5 SOAs tested. The similar flanker Gabors’ orientations were drawn from the set of flanker tilts matching the 80% threshold in the Orientation Calibration experiment, while orthogonal flankers were drawn from this same set and rotated by 90°.

Results

Observers’ performance on the two tasks was not independent, in that truly obtaining the correct answer to the orientation discrimination task on any given trial required that they first correctly identify the target letter in order to shift attention to the correct hemifield. Therefore, for the analyses of the dual experiment, we measured the outcome variable as the conditional probability of observers giving the correct responses to both tasks (each with a 50% chance guessing rate), given that they had correctly responded to the first task. This is calculated as the frequency of correct responses to both tasks divided by the frequency of responses to the first task:

\[
P(C_1 \& C_2) = P(C_2 | C_1) = P \left( \frac{C_1 \cap C_2}{C_1} \right)
\]

This reflects correct answers given on the second task only when observers endogenously shifted their attention to the appropriate side of fixation. We calculated this probability independently for each observer, at every combination of SOA and blink event. Thus, for each blink condition this yields 5 data points (targets blink, flankers blink, all blink). However, in the no-blink condition, since no SOAs exist, data are averaged across the 5 time points to form a hypothetical flat line (blue curves in both panels of Figure 5 & 6), indicating performance in the absence of any transient blink application. The shaded region about the line represents the 95% confidence interval.
about the no-blink data. Points lying beyond this region are significantly different from no-blink performance at the $\alpha = 0.05$ level.

In the similar-flankers blocks, target blinks improved observers’ performance by an average of 6%, as compared to the no-blink condition, but only at blink-cue simultaneity (where SOA = 0) and at 100 ms SOA. Similarly, blinking both the targets and flankers (all-blink condition) resulted in a 5% improvement in performance at 100 ms SOA (Figure 6). Conversely, observers’ performance was hindered in the flankers-blink condition at simultaneity and got progressively worse at larger SOAs. Flanker blinks had no effect on performance at any negative SOAs. Note that performance at negative SOAs is roughly equal to no-blink performance across all conditions (Figure 5, left column).

We obtained similar results in the dissimilar-flankers blocks (Figure 5, right column); once again, performance improved when the targets blinked at simultaneous and 100ms SOAs. And, as before, no improvements in performance were seen as a result of blinking the targets or flankers at negative SOAs. Performance did increase at -100ms SOA in the all-blink condition, though this effect is driven entirely by observer MM.

Equally, there was a notable improvement in performance induced by target blinks at positive SOAs. Particularly by comparison with the similar-flankers blocks, the observed change in performance increased to 13% –more than twofold that observed with similar flankers. Also, dissimilar flankers produced no decrement in performance at any SOA.
It should be noted, however, that while the averaged data indicate that both simultaneous and 100ms cue-blank SOAs produce increased performance (Figure 5), this

Figure 5. Performance in Experiments 1 (left column) and 2 (right column), with parallel and orthogonal flankers, respectively. Negative SOAs indicate blinks preceding cue letters and vice-versa. Proportion correct on y axes refers to $P(C_1|C_2)$. Shaded region represents 95% CI about the no-blink baseline condition.
is not necessarily typical of any single observer. Rather, peaks in target-blink performance were somewhat unstable among observers. Even within observers, peak timing was inconsistent; ABU and MM, for instance, both had narrowly tuned peaks for similar flanker blocks (Figure 5 left panels), and relatively wide peaks for dissimilar flanker blocks (Figure 5, right panels). Peak timing was also inconsistent across blocks, shifting by ±100ms for 4 of the 5 observers, and completely disappearing for the 5th in the dissimilar flanker blocks (Figure 5, observer MG).

**Figure 5.** Performance in Experiments 1 (parallel flankers) and 2 (orthogonal flankers) averaged across observers. Proportion correct refers to the $P(C_1|C_2)$. Shaded regions correspond to the 95% CI based on between-observer variability in no-blink data.

**AOC Analyses.**

Although the instructions given to the observers required serially completing the letter identification and then the orientation discrimination task, it is possible that they could have devoted varying amounts of endogenous spatial attention to each of them, without regard for any stimulus manipulations. In order to test for the possibility that observers were trading-off performance on one task to achieve gains on the other, we constructed an attention operating characteristic curve (AOC) for each observer.
Previous research has shown that when observers are forced to complete two tasks in quick succession, they distribute their attentional resources between them in accordance with task demands (Hoffman & Nelson, 1981; Sperling & Melchner, 1978). Consequently, performance increases on one task come at the cost of performance on the other. Such trade-offs can be characterized by AOCs. These plot observers’ rates of correct responses on each task against one another, and compare this with performance on the same tasks when completed alone.

If observers reach levels of performance on the dual task that are commensurate with their ability to complete each task alone, the two tasks are inferred to be independent of one another. In other words, at independence the two tasks no longer interfere with one another, presumably because they require fewer resources than the total capacity of endogenous attention.

In our experiments, we measured single-task performance for orientation discrimination and letter identification during calibration (see Figure 3). Therefore, we can compare performance on each of the two tasks separately to joint performance in the main experiment. If the observed benefit of target blinks truly uses exogenous transient spatial attention, this would be reflected by gains in performance on the orientation task not met by reductions to letter identification ability. Conversely, if any benefit to one task is instead met by reduction in the other, this would constitute evidence of a trade-off between tasks. A paired t-test found no difference between letter identification performance in positive versus negative SOAs ($t(4) = 0.716, \rho > 0.05$) across blink conditions, suggesting the former.
In particular, since the target letter in the first task was followed by a nearly unmistakable temporal cue (the remaining letters being presented in black) we have assumed that the experienced observers were certain of the time at which they should have shifted their attention, but had to ensure that their behavioral responses included only those trials where they were certain of the side to which they should shift their attention. Therefore, in Figure 7, plotting letter identification and orientation discrimination task performance against one another as an AOC, we corrected the conditional probabilities of correct responses on the orientation discrimination task, \( P(C_1|C_2) \), for ‘lucky guesses’ (hits due to guessing), rendering this measure independent of false positive letter identifications. Observers guessed the wrong letter on \( 1-P(C_1) \) of trials, so we can infer that an equal number guesses (given 50% guessing rate) exist where the correct response was given. The conditioned probability of correctly discriminating the target’s orientation includes those correct guesses. Consequently, we removed them by assuming that half of the correct guesses of letter identity were accompanied by correct guesses of target tilt (since this task also has 50% guess rate).

As such, the orientation discrimination axes on all AOC plots denote the conditioned probability given by equation (1), net of guesses. This way, they truly include only trials on which attention was endogenously shifted in response to a correct letter identification response, and do not depend on guess rates for the first task.

As it turns out, the observers either reached the point of task independence (performance equivalent to single task), or exceeded their calibration performance (Figure 7). This suggests that the two tasks are no longer in competition with one another for endogenous attentional resources. Note that orientation discrimination performance in
the target-blink condition consistently lies within ±SEM of the corresponding point in the no-blink condition, suggesting no generalized processing advantage for the targets at positive SOAs.

In contrast, flanker-blink discrimination performance is consistently lower than no-blink performance, indicating that the transients did hinder orientation discrimination.
at positive SOAs when just the flankers blinked. Furthermore, this is not met with any consistent decline in letter identification performance. Taken together these results suggest that blinking the flankers hindered performance on the crowding task independently of any trade-off effects (i.e. a true decrement in target discriminability).

**Discussion**

Previous research by Harrison & Bex (2014) and Greenwood et al. (2014), has found improvements in orientation discrimination due to transient onset isolation of the target, particularly when only target onset is rendered salient. However, these experiments only manipulated exogenous attention through the application of transients to the targets and flankers (as we have done in the current experiment); they did not control for any co-occurrence of endogenous covert shifts of spatial attention to the target. Rather, their observers’ endogenous attention would be relatively fixed on the target throughout each trial. Thus, their results do not fully account for the combined effects of exogenous spatial attention in the absence of endogenous attention. Given our results, their exogenous attentional manipulations most likely operated in concert with covert endogenous shifts of attention to produce the orientation discrimination benefits they observed, rather than being solely a result of transient exogenous attentional capture.

Since we manipulated the shifting of endogenous spatial attention to the flanked targets by the addition of a challenging letter identification task, we could empirically test whether application of transient blinks improves performance by reducing crowding in both in the presence and absence of endogenous spatial attention. We find that when onset transients are applied to the targets prior to observers’ shifting endogenous attention
from one task to the other (i.e. at negative SOAs), no benefit in target orientation discrimination occurs.

The addition of the letter identification task potentially allowed us the ability to determine whether any contribution of the transient blinks to orientation discrimination performance occurred as a result of an interaction between the endogenous and exogenous components of spatial attention. Our observers’ performance indicates no trade-off between tasks, suggesting that the two tasks are processed independently of one another in the dual-task experiment. Thus, changes in performance across SOAs were not the results of a trade-off, but of the aforementioned interaction.

In this experiment, we observed improvements in Gabor orientation discrimination only at SOAs of 0 and 100ms, which strongly suggests that the onset transient blinks only improve performance in concert with endogenous spatial attention to the blinked stimulus. Potentially, this effect was more narrowly temporally tuned than 100 ms intervals of SOA could capture, so averaging positive and negative SOAs separately further diminished its magnitude.

Previous research suggests that orientation discrimination performance is significantly impaired by the addition of an attentionally demanding RSVP task (Joseph, Chun & Nakayama, 1997), similar to the letter identification task employed here. Instructing observers to identify a white letter among black letters in an RSVP stream reduced their observers’ ability to subsequently identify a peripheral orientation odd-ball Gabor target, even at lags up to 700ms after a RSVP target’s occurrence.

Therefore, our dual-task paradigm most likely impaired naïve observers’ ability to complete both tasks simultaneously and thereby forced them to process the letter and
Gabor items serially. We surmise that, in addition to the benefits gained from target-blinks at positive SOAs, observers likely improved upon their single-task calibration performances over time based on prolonged exposure to similar stimuli during the main experiment.

Improvement in orientation discrimination performance due to blinks did not persist over 200 ms SOA. Indeed, previous research suggests that spatial cuing even as much as 118 ms prior to a flanked target’s appearance can still improve discrimination thresholds during crowding (Scolari et al., 2007). Furthermore, and Sperling (1986) find similar time courses for shifting attention to the visual periphery.

Therefore, the results of this experiment suggest that increased discriminability induced by the blinks is fleeting (lasting only up to 200ms following the cue letter). If endogenous attention shifts were indeed much slower than 100ms, then the blinks could only be serving to distinguish the targets and flankers from one another pre-attentively. Yet no benefit in discriminability is seen at any negative SOAs tested. Hence, in the case of slow attention shifts, the onset transient blinks would have had to act on endogenous spatial attention, as pre-cues to the to-be-attended target location. In this experiment, the blinks were always applied on both sides of fixation, preventing them from providing any information about the true target’s location (left or right of fixation). As a result, the only informative cue as to the correct hemifield location was the endogenous letter cue (the last white letter in the RSVP stream).

Previous research indicates that attention is ballistic; once an endogenous spatial attentional shift is instantiated, it is difficult to alter without reducing resulting task performance (Posner, 1980). In his experiments, Posner (1980; Posner, Snyder &
Davidson, 1980) demonstrated that exogenous spatial cues only improve performance when they predominantly inform about the target’s true location. If the cue is predominantly invalid or neutral, it is generally ineffective or debilitating.

Thus, in our experiment, where blinks were always applied on both sides of fixation, they would have little or no contribution to performance without information carried by the endogenous attentional cue. Therefore, accurate covert shifts of endogenous spatial attention prior to the arrival of the target letter (i.e. the last letter in white) could not be made at better than chance levels of performance. As such, shifts of endogenous attention would have had to be initiated after the blink in order to receive a benefit from exogenous spatial cues. Since the transient-induced effect of blinking the target occurs within a limited range of 200 ms, this suggests that the advantage might be in focusing endogenous spatial attention on the target.

Taken together, our results exclude the possibility of an explanation purely in terms of endogenous spatial attentional cueing. Rather, the benefit of the transient blink applied to the target increases its saliency either during or just following the shift of endogenous spatial attention. The observed improvement in performance suggests that exogenous properly localized onset transients facilitate spatial attentional selection.

Chapter 2

Experiment 2 – Comparison of attended and un-attended features

Our previous experiment found that the combination of endogenous and exogenous attentional capture could un-crowd under conditions of low and high feature
contrast between targets and flankers when. However, it remains possible that restriction of spatial attention to the target (i.e. by onset transients) was sufficient to un-crowd whenever feature differences were large enough to induce target pop-out, regardless of whether that feature was actually attended. Since only the attended feature (orientation) was varied to create feature contrast between target and flankers, we intend to compare the effects of attended versus unattended feature contrasts to determine whether attentional selection of the target occurs simply spatially, or is limited to the attended feature.

Here, we will control for feature attention using a similar paradigm to experiment 1, in order to determine whether feature contrasts produce uncrowding pre-attentively in the feature domain, as well as the spatial domain. Whereas only the orientations of the flanker Gabors varied in the previous experiment, here we vary their carriers’ spatial frequencies instead. Following previous research by Chung, Legge and Tjan (2001), discriminability is independent of spatial frequency up to approximately 5 cpd. All targets in the previous experiment were fixed within this range, and therefore we assume that their judgments were relatively independent of their SF content at suprathreshold contrast levels. Also, the orientation discrimination task will be kept independent of spatial frequency content, so that observers should only need to attend to the target’s orientation to provide optimally informed responses.

If the flankers retain similar orientations to the target, and crowding is determined by the combination of spatial and feature attention, we should expect that changing unattended task-irrelevant flanker features (such as spatial frequency) would produce no un-crowding. Conversely, if crowding occurs purely on the basis of spatial proximity and
feature contrast (regardless of feature attention), un-crowding should occur for any sufficiently large feature contrast.

Method

Stimuli

Here we replicated all procedures from Experiment 1, with an important variation: rather than manipulating the target-flanker orientation contrast between blocks, we instead varied their spatial frequencies. It should be noted that these blocks were interleaved with those run in experiment 1, using the same observers. In a given block, either all parameters were identical to those described in Experiment 1, or the flanker Gabors’ spatial frequencies would be doubled. The observers’ tasks would remain same, regardless of the flanker feature manipulation: identification of the target letter, followed by orientation discrimination of a flanked Gabor target.

Results

Panel (b) in Figure 8 shows performance on the task in blocks where the flankers’ spatial frequencies were doubled. A two-tailed repeated measures t-test determined that performance increased sharply between similarity conditions \( t(5) = 3.357, \rho = 0.028 \). No significant differences between blink conditions were found in panel (b).
Figure 6. Plots of performance against SOA averaged across observers for (a) similar flanker orientations (replicated from Figure 5) and (b) flankers with doubled carrier spatial frequency. The shaded region represents the 95% CI about the mean no-blink performance.

We did not create AOC plots as in experiment 1, since calibration data were collected using parallel flankers.

Comparison of the flanker-blink conditions in panels (a) and (b) in Figure 8 reveals that the deleterious effects of blinking the flankers noted in Experiment 1 disappears; relative to the no-blink condition, performance in the flanker-blink condition is stable and comparable to performance in the no-blink condition.

Discussion

In the absence of endogenous spatial attention to the target (i.e. at negative SOAs) we find improved performance as a result of increasing flanker Gabor carrier spatial frequency, relative to performance when flankers were similar. Much like the implication of Treisman’s (1982) finding that feature singletons can be processed pre-attentively, our results seem to indicate that spatial frequency contrast feature singleton targets (i.e. spatial frequency) might be processed pre-attentively.

To add to this point, even in the no blink condition (i.e. in the absence of any exogenous transient attentional cue), we still find improved performance when the
flankers’ spatial frequencies are doubled. But more importantly, the magnitude of the improvement in target discriminability conferred by target transients noted in Experiment 1 is greatly reduced. Whatever the pre-attentive processing advantage conferred by spatial frequency differences between targets and flankers, there may still be some leftover benefit to target discriminability based on the combination of endogenous and exogenous attentional capture.

In the crowding literature, Sayim, Herzog and Westheimer (2011) have shown that target conspicuity not only reduces crowding, but that it also yields lower reaction times on visual search tasks. They demonstrated that target pop-out reduced crowding when flankers’ lengths or colors were different from a Vernier offset target’s. Importantly, neither of these features was relevant to the offset discrimination task required for correctly identifying the targets. Furthermore, the same kinds of target-flanker feature contrast also produced shorter search times for those same targets that had been shown to pop-out in the crowding task. Taken together, their results suggest that target-flanker feature contrast induces pop-out when the flankers are grouped separately from the target, and potentially alleviate crowding.

Feature-singleton searches are thought to occur pre-attentively (Treisman, 1982), though the advantage for processing targets in this experiment (Figure 8, panel b) is not explicable in terms of target pop-out alone. Such an account falls short of explaining the results of Experiment 1 (Figures 5 and 6, right panels), wherein orientation contrast did not produce pre-attentive feature pop-out. This account would have predicted the results of Experiment 1 to resemble those seen in panel (b) of Figure 8. Instead, we find that
only targets with spatial frequencies different from the flankers’ (but not orientations) become uncrowded at negative SOAs.

Our results therefore suggest that task-relevant orientation feature contrast confers a selective benefit to target processing when endogenous spatial attention is captured by the target (i.e. when it blinks at 0-100ms SOA). When feature contrast occurs along the task-irrelevant spatial frequency dimension, the interaction of endogenous spatial attention and exogenous transient spatial attention continue to provide this advantage. However, the persistent benefit in discriminability obtained by increasing task-irrelevant spatial frequency contrast occurs regardless of the observer’s endogenous attentional state, or any exogenous attentional capture due to transient blinks.

Chapter 3

Experiment 3 – Spatial distribution of un-crowding

The previous experiments have thus far tested the magnitude of transient endogenous and exogenous spatial and feature attention’s ability to alleviate crowding of endogenously and exogenously attended targets. However, crowding is largely characterized by its spatial extent (critical spacing). Therefore, the present experiment seeks to determine whether crowding and attention share a spatial profile.

It is possible that, even though spatial and feature attention can improve identification of a flanked target, this is not achieved by reducing crowding itself, per se. Rather, it could be that attention simply increases target salience by inducing pop-out, thus increasing attentional capture of the blinked elements and reducing the likelihood of
compulsory averaging with un-blinked elements. In other words, even an elevated level of crowding could leave an intrinsically more discriminable target relatively identifiable, but would not change the area over which flankers interfere with its identification, however slightly they might do so.

Pelli, Palomares & Majaj (2004) propose that temporal manipulations affecting discriminability are not necessarily reflective of the same underlying process that produces crowding. Particularly, they posit that such ‘temporal crowding’ could arise from overloading attention during discrimination tasks, wherein the observer has insufficient time to adequately process target and distractor signals. Under these circumstances, flanked discriminability can be impaired by comparison with isolated targets, even over spatial extents much greater than 0.5 times eccentricity. Presumably, this arises from inadequate processing time for the target and flankers and not from feature pooling or grouping (Pelli, Palomres & Majaj, 2004).

Yet, our stimuli were available for at least 1300 ms on every trial in Experiments 1 & 2 – more than enough time for observers to process them in parallel. Therefore, it is unlikely that they included too much information. And, given that our previously observed results (Experiment 1) were indicative of independent processing as opposed to trade-offs (Figure 7), it is also unlikely that interference between tasks or the number of elements to be processed over time in the RSVP stream would have induced temporal crowding overall. In sum, while temporal crowding would have a larger spatial profile than Bouma’s bound of half-eccentricity, it still cannot adequately explain the effect of the transient blinks on orientation discrimination performance.
Therefore, in the present experiment, we turn to an examination of the spatial extent of crowding induced by our stimuli. If critical spacing results from insufficient endogenous spatial attentional resolution, as Intrilligator and Cavanagh (2001) have claimed, then the transient-induced orientation discrimination advantage should be confined to Bouma’s bound (beyond which there would be no spatial crowding to alleviate). Harrison & Bex (2014) find that critical spacing decreases by a factor of 2, following 100 ms target-flanker onset asynchrony (leading flankers), beyond which it remains stably reduced. Therefore, at the SOAs at which we have tested so far, we would expect that increasing the spacing between targets and flankers ought to benefit overall performance, and the blinks would have a reduced effect at wider spacings, once the flankers are already beyond Bouma’s bound.

Consequently, this experiment distinguishes two competing alternative hypotheses. The first is that the combination of endogenous and exogenous spatial attention to the target determines the zone of interaction, such that transients applied to the targets will produce no processing advantage at target-flanker spacings greater than half the target’s eccentricity. Under this hypothesis, performance is expected to improve across all SOAs when target spacing exceeds Bouma’s bound. Equally, flankers presented within the zone of interaction would produce the same selective advantage for the target transients from 0-100 ms SOA found in Experiments 1 & 2..

The second alternative is that the performance advantage conferred by endogenously attended transient blinks operates along a separate mechanism from crowding. This hypothesis predicts improved orientation discrimination at eccentricities where target-flanker separation is greater than the usually observed Bouma’s bound of
half-eccentricity, but only at positive SOAs. This would be evidence of a separate mechanism for un-crowding specifically owing to the combination of endogenous spatial attention and exogenous attentional capture induced by the target transient, rather than an implication that attention is causally involved in the initial instantiation of crowding. So, according to this hypothesis, endogenously attended target transients improve orientation discrimination by increasing orientation discriminability at the target’s spatial location, independent of crowding.

Indeed, there is previous research to suggest that crowding results from compulsory averaging of target and flanker features within the zone of interaction (Parkes, et al., 2001). It therefore follows that, when flankers are located beyond Bouma’s bound, local estimates of the target’s orientation are sufficiently isolated to render the flankers’ contributions to the average negligible.

Other research also suggests that the effects of feature averaging and endogenous spatial attention on flanked target discriminability are separable. Dakin, Bex, Cass and Watt (2009) manipulated crowding and attentional load by adding an attentional load, flankers, or both while observers reported the average orientation of an array of peripheral oriented noise targets. They tested whether crowding and attention affected either the local precision of orientation estimates during feature averaging, or global orientation sampling efficiency. Adding attentional load (i.e. with dual task present) shifted the entire distribution of orientation discrimination thresholds upwards, indicating a general loss of global sampling efficiency. In contrast, adding flankers surrounding the target array caused the distribution to widen. Furthermore, these two effects were found to be additive, with separable effects on orientation discrimination thresholds.
Based on these accounts, we hypothesize that when the flankers are closely spaced, performance should improve when target-blinks occur at 0-100ms SOAs, as in Experiments 1 & 2. Conversely, when the precision of local orientation estimates is maximized by way of target-flanker separations exceeding Bouma’s bound, we expect no enhancement to orientation discrimination performance to arise from blinking the targets. Once target-flanker spacing becomes sufficiently wide, minimal endogenous attentional resources should be necessary for processing the targets’ orientations separately from the flankers’, thus minimizing the benefits of interaction with exogenous attention.

Method

Observers

We used the same observers as in Experiments 1 and 2, and they were therefore very well practiced.

Stimuli

All stimuli are generally the same as those used in Experiment 1. However, in this experiment we test only for the effectiveness of the target transient in reducing crowding. As such, we present the same Gabor targets, but only blink either the target Gabor element, or none of them. Another crucial difference in the present experiment is that we test the effects of endogenous and exogenous attention on critical spacing by either presenting the flanker Gabors at the same 2.5° center-to-center spacings as in Experiment 1 (Figure 1), or doubling them to 5°, which should lie beyond Bouma’s bound of half eccentricity with the targets displayed at 8°.
Procedure

Our manipulations yield a $2 \times 2 \times 5$ experimental design: target-blink and no-blink conditions crossed with near and far spacings, at all 5 SOAs (from -200 ms to +200 ms) tested previously. Each observer completed 4 blocks of 16 trials per condition for a total of 320 trials in each block, and 1280 trials total.

Prior to completing each block in the main experiment, we measured observers’ orientation discrimination thresholds for flanked Gabor targets to determine their 80% performance thresholds at 8º eccentricity, as in the previous experiments. Similarly, we calibrated their letter identification ability using the RSVP stream of white and black letters described in Experiment 1.

In the main experiment, the orientations of the targets and flankers were fixed according to the orientation discrimination thresholds collected from each observer as described previously. We did not manipulate any other target or flanker features here, except that flanker Gabors were still tilted left or right by the same magnitude as the target Gabors’ tilts.

Results

Here again, we expected to find a processing advantage for the targets when they were blinked at 0 or +100 SOA. Furthermore, we expected a general advantage in target processing at wider target-flanker spacings, similar to the effect seen when varying flanker spatial frequency (Figure 9). Assuming a critical spacing near 0.5 times the target’s eccentricity following Bouma’s (1970; 1973) bound, any processing advantage should disappear beyond this point, so a spacing of 2.5º would be within Bouma’s bound at 8º eccentricity, but beyond it at spacings of 5º.
Instead, comparison of the target transient and no-blink conditions at near spacings reveal no benefit of the transients (solid red curves in Figure 9) for any of the observers tested. Clear between-subjects differences exist, such that 2 of the 5 observers

![Graphs showing orientation discrimination against SOA for different observers.](image)

**Figure 9.** Orientation discrimination plotted against SOA. Blinked elements are noted in the legend. Error bars denote mean ±SEM for the target blink conditions, and the 95% CI for the no-blink conditions.
accrue slight benefits from the blinks (though statistically null) whereas 3 of the 5 observers see minor hindrances instead (also statistically insignificant). In the far-spacing condition, however, the blink significantly improved performance across observers, though not significantly within any of the individual observers (Figure 9, dashed lines).

Additionally, no consistent pattern of performance is notable between spacings, neither in the target-blink or no-blink conditions, nor when averaging across blink conditions. In order to ensure that failure of the spacing manipulation was not the result of weak baseline crowding, our subjects ran an uncrowded version of the same experiment (at all five SOAs tested, but with the flanker Gabors absent), and all reached ceiling (100% correct) or near-ceiling performance on the orientation discrimination task, suggesting that the stimuli used were indeed crowded. TE plots are therefore not displayed, as they are uninformative at ceiling levels of performance.

Discussion

Clearly, our results show that the advantage to orientation discrimination performance conferred by transient target blinks does not share the spatial profile that would be predicted by Bouma’s bound. Therefore, while transient spatial attentional capture can improve identification of a flanked target, this is not due to a reduction in crowding’s spatial extent. It seems instead that the transient blinks simply increase target salience, thus reducing the likelihood of compulsory averaging within the target’s vicinity.

These results do agree with previous studies of the effects of spatial attentional cueing on critical spacing. Scolari and colleagues (2007) found that even though predominantly valid exogenous spatial transient pre-cues produced a reduction in the
magnitude of crowding, they did not affect its spatial distribution. This dissociation between the spatial extent and magnitude of crowding further implies that crowding and exogenous transient attentional capture operate differently to affect target discriminability. Even a salient target can be crowded; increasing its salience may boost its identifiability, but that does not necessarily mean that it has become uncrowded. It could simply mean that the target has become more detectable at the same level of crowding.

Similarly, Dakin and colleagues (2009) found that sustained endogenous feature attention’s effect on crowding is simply to enhance the precision of global feature salience across spatial locations during target identification. Thus, whether the target appears with nearby or distant flankers only changes the variability of local feature estimates, presumably by adding more elements. So, whereas transient spatial attention serves to increase the target signal, crowding’s effect is to increase featural noise at the target’s location. These two effects might act in opposition to determine the overall discriminability of the target’s features, but by no means do they reflect the same underlying process. Truly uncrowding a target object should be taken to imply reducing the local noise around its spatial location.

In this sense, our results could be explained by learning effects since the observers all ran this experiment last. By the time they participated in this final experiment, they were already familiar with the targets’ positions, and had reached the point of independent processing for the letter identification and orientation discrimination tasks (Figure 7). Thus, even in the absence of a transient target blink, observers had an endogenous distribution of spatial attention that favored the two targets’ locations even
before the cue letter had appeared. As such, over the course of relatively long (1500 ms) trials, observers had ample time to extract the target signal as it was amplified by endogenous spatial attention.

If ‘temporal crowding’ were the process underlying our results, we would expect transients applied to the target to improve discrimination even when spacing exceeded 0.5 times eccentricity. However, we saw no consistent improvements, suggesting that our results were not easily explicable in these terms.

Finally, we surmise that our results can be explained by our failure to counterbalance the order of experimental block completion within observers. Since all observers ran these blocks last, they had probably developed strategies for maximizing their ability to endogenously attend to the targets’ locations. As targets always occurred in the same spatial locations to either side of fixation, observers may have learned to endogenously shift their attention to the spatial location covered by the targets and flankers in Experiments 1 and 2.

Despite giving observers substantial practice at the task during Experiments 1 and 2 at narrow spacings, when we enlarged the spatial extent of the area over which they had become accustomed to spreading their attention, they managed to attend widely to the targets and flankers. Thus, we find consistently improved target-blink performance for wide spacings, but inconsistent results at near spacings.
Chapter 4

Experiment 4 – Replication with Naïve Observers

In order to determine whether the results of Study 1 were brought about by learning effects, we replicated its experiments using naïve observers, and added a control to check that our stimuli indeed follow Bouma’s bound in the absence of the attentional load imposed by the dual-task paradigm. Previous research suggests that learning through practice can reduce crowding (Sun, Chung & Tjan, 2010). Importantly, this research also contends that observers improved as a result of endogenously calibrating the spatial distribution of attention – or “perceptual window”, to match the learned stimuli. Likewise, Hussain, Webb, Astle and McGraw (2012) find that observers experience reduced crowding in a letter identification task at post-test intervals ranging up to 8 months. Indeed, having each run 8 sessions for a total of 6400 trials, our observers had run as many trials as those in Sun and colleagues’ (2010) and may have thereby reduced the extent to which they experienced crowding on the stimuli used, despite our repeatedly calibrating then for optimally challenging task difficulty just prior to each experimental session.

Method

Participants

14 naïve psychophysical observers were recruited from the Northeastern University Psylink pool for participation, for partial credit towards completion of their Foundations of Psychology course. As part of obtaining informed consent, observers were informed that their data would not be retained unless they had normal or corrected-
normal vision (20/20 as measured by Snellen chart test of visual acuity). Observers were informed of the general nature of the study before consenting to participate. They were told that the study would measure their ability to discriminate crowded objects under varied conditions of attention, by having them try to complete two tasks in sequence with varying delays.

Procedure

Calibration experiments. The orientation discrimination calibration experiment was run exactly as detailed in Experiment 1, except that the naïve observers only completed one block of it. While this provides a somewhat noisier estimate of the 80% criterion performance we sought, it addresses the potential problem of long-term pre-exposure to the stimuli used in the main experiment. Presumably this should reduce any perceptual learning that experienced observers might have accrued during their calibration blocks.

The letter identification discrimination task was modified to be made easier for these observers. Pilot testing revealed that their thresholds were prohibitively high and would have limited the number of testable SOAs without substantially increasing trial durations in the main experiment. In the modified task, observers attended to a random sequence of letters displayed in white against a neutral gray background. On every trial, one of the letters in the sequence would be an R or an L, but all remaining items could be any of the other 24 letters. Participants provided their responses by pressing the left or right mouse buttons to indicate their decisions (left mouse button for L and right mouse button for R).
Results

Calibration Experiments

Across the 14 observers, only 2 had poorly fitted calibration data, although in both of these cases their performance was stable at 80% across the range of parameter values tested. We therefore retained their data by fixing the offending parameter in the main experiment at the lowest value for which their actual data indicated they had achieved that mark. Otherwise, all $r^2$ values were greater than 0.4 (See Figure 10, Panel c) and individuals’ 80% correct thresholds were estimated by interpolation with the linear fit using equation (1).

Figure 10. Linear fit to observers’ averaged performance data in the (a) orientation discrimination and (b) letter identification tasks. Note: 3 video frames = 50ms. Panel (c) box and whisker plots of the distribution of $R^2$ fit parameters.

Main Experiment

Following equation (1), we once again evaluated the conditional probability of correctly discriminating the true Gabor target’s tilt given that the correct cue letter was perceived. This was done separately for each observer, and the resulting probabilities were averaged within blink $\times$ SOA $\times$ tilt condition. Additionally, we took the arbitrarily generated SOAs for each observer’s no blink condition as a measure of intrinsic
variability, and used the variability between these randomly sampled data points to construct the 95% CI about the mean no-blink performance across observers. As such, means of repeated no-blink samples falling beyond the shaded region in Figure 11 would only occur with 5% likelihood over repeated samples. Thus, assuming a null hypothesis of no blink effect, data that lie outside the shaded region in Figure 11 are significantly different from performance in the no-blink condition at the $\alpha = 0.05$ level.

![Figure 11](image-url)

**Figure 11.** Target identification conditioned on letter identification, plotted against SOAs for (a) parallel flankers, (b) orthogonal flankers, doubled SF flankers (c), and Near vs. Far flanker spacing (d). Shaded regions denote 95% CI of the no-blink mean, based on random resampling of no-blink data. Note: the shaded regions represent 95% CI about no blink data for near spacings in all panels.

**Attention Operating Characteristic (AOC) Analyses**

Since the two tasks were fixed at the level determined in calibration on each run, the performance on each task in isolation of the other is known, and any change in performance could arise from one of two sources: (i) a change in the attentive processing...
state of the observer (i.e. shifting the focus attention from one task to the other), or (ii) a change in signal discriminability induced by the various conditions. Meanwhile, the baseline metric for performance of both tasks simultaneously can be taken as the no-blink condition since both tasks are relatively comparable to their single-task calibration counterparts, with no additional manipulations (such as blinks) added. Here we again correct for ‘lucky guesses’ as we did for the experienced observers in experiment 1.

Since the naïve observers completed both tasks on each trial, and our outcome measure is based on the joint probability of correct responses (on both the letter identification and orientation discrimination tasks), it is possible that changes in dual-task performance might be induced by a trade-off. At simultaneous and positive SOAs, observers should have been devoting more attention to the orientation discrimination task, and less to the letter identification task, resulting in better performance as measured by the conditional probability $P(C_2 | C_1)$, particularly since they were instructed to

![Attention Operating Characteristics averaged across observers in similar flankers conditions at (a) negative SOAs and (b) positive SOAs. Black points on axes represent single-task performance calibrations for orientation discrimination (x axis) and letter identification (y axis). Results corrected for guessing, as described in the Method section of Experiment 1. Error bars represent 95% CI about the means for all points.](image-url)
shift their attention from task to task in congruence with their letter identification responses.

For the single task calibrations, observers’ performance is commensurate with the 80% threshold estimates. The dashed line connecting these points represents the linear trade-off in performance between tasks, meaning that as performance on either task shifts, an equal and opposite shift in performance would occur in the other task. For all observers, performance in the no-blink condition lies below the independence point \((x,y) = (0.8, 0.8)\), indicating that these two tasks do indeed trade-off against one another (Figure 12).

Performance in the target-blink condition specifically is shifted positively with respect to the no-blink condition along both axes, suggesting that the observed improvement in performance due to the transients is due to a benefit in target discriminability rather than a trade-off (Figure 12). Moreover, since performance on the RSVP task is concomitantly improved, this result is unlikely to be due to a reduction in the number of trials preserved in the initial analysis of \(P(C_2 | C_1)\).

Discussion

Orientation discrimination performance in every blink condition falls between the trade-off line and the independence point \((x,y) = (0.8,0.8)\) for all \(\text{SOA} < 0\) (Figure 12). Comparison of these data (Figure 12) with those of the experienced observers, whose data fell at or above their single-task calibration points (Figure 7), suggests that performance differences between them cannot fully be accounted for by shifts in the distribution of endogenous attentional resources between the two tasks (i.e. trade-offs).
Rather, they arose from the difference in overall processing capacity for the stimuli due to learning.

Assuming a fixed capacity for sustained endogenous spatial attention across the visual field, naïve observers were still engaged in the letter identification and thus forced to attend carefully for the cue letter. Unlike the experienced observers, this left them with a smaller remainder of their endogenous attentional resources to spare for the second task until such time as the cue letter appeared.

As we saw with the experienced observers in Experiment 1 (Figures 5 and 6, left hand panels), the naïve observers in the present experiment obtain benefits in target discriminability when only the target blinks at 0 and +100ms SOAs (Figure 11, panel a), and only cases where flankers had either orthogonal orientations, or doubled spatial frequency carriers with respect to their targets. On one hand, this could be taken to suggest that the naïve observers receive comparable benefits from feature contrasts along both task-relevant and irrelevant features. On the other, it might imply that their discrimination ability rests more on object-level attention to the targets than feature-based endogenous attention. Yet, transients applied to the flankers did not hinder observers’ orientation discrimination for doubled spatial frequency flankers, whereas it does so marginally for orthogonal flankers. This could imply that exogenous attentional cues might operate on the features independently, since only in the case of the endogenously attended feature (orientation) did drawing attention to the flankers exogenously hinder performance at all.

As well, it should be noted that the observed benefit of blinking the target cannot be explained simply based on the exogenous transient blinks, regardless of the types of
flankers used, since these benefits only occurred at positive SOAs for the parallel, orthogonal and doubled spatial frequency. This strongly implicates endogenous attentional capture of the target in the gains produced by addition of transient blinks to the targets in these conditions.

However, when target-flanker spacing was varied from trial to trial, it seems that the naïve observers received benefits in performance only at negative SOAs at spacings beyond Bouma’s bound. This suggests that once Bouma’s bound is exceeded, endogenous attention the target may no longer be necessary to producing a benefit from exogenous transient-induced benefits in orientation discrimination performance. We do not, however, see any generalized improvement from near to far spacings between the no-blink conditions (Figure 11, panel d), suggesting that even for the naïve observers, Bouma’s bound is increased to well beyond the normally observed half-eccentricity. And, we see no improvements in performance due to target-transient blinks at any positive SOAs. As a whole, our results suggest that Bouma’s bound is extended when endogenous attention to the target is not supplemented by exogenous transient attentional capture.

Chapter 5

Experiment 5 – Replication of Critical Spacing

Experiments 3 and 4 failed to obey Bouma’s bound of half-eccentricity on spatial crowding, Given the number of replications that the eccentricity scaling of crowding has enjoyed, it has come to be regarded as a hallmark of crowding. Therefore, we cannot draw any conclusion about our study’s ability to necessarily undo the deleterious effects
of spatial crowding without first demonstrating that some level of crowding was indeed present in the stimulus displays we used. Therefore, the present experiment tests whether Bouma’s bound is obeyed in the absence of the high endogenous attentional load induced by the letter identification tasks in all our previous experiments.

In particular, both the naïve and experienced subjects receive no benefit of wider spacings (0.625 * eccentricity) in experiments 3 and 4. This leads to one of two possibilities: either our stimuli induced no crowding, or endogenous spatial attentional capture breaks Bouma’s bound despite crowding. In order to test these possibilities, in this experiment we will vary the center to center spacing of target and flankers under conditions of minimal endogenous attentional load.

Method

Participants

Ten naïve observers were recruited for participation from the Northeastern University undergraduate participant pool, as partial credit towards completion of their Foundations of Psychology course. All were tested with a Snellen letter chart to ensure normal or corrected-normal (20/20) visual acuity, as in the previous experiments.

Stimuli

All target and flanker Gabor stimuli are identical to those used in the previous experiments, as explained in the Orientation Calibration section of Experiment 1 (Figures 2 and 3). Here, however, targets could appear to the left or right of fixation at either 4 or 8 eccentricity, and were either flanked by nearly parallel Gabors at 0.25, 0.5, 0.75 times eccentricity, or displayed in isolation (unflanked). Additionally, 750 ms before the start of the trial, a small arrow (0.3° visual angle) replaced the fixation cross.
Procedure

Before completing the main experiment, each observer completed one full orientation calibration session and the tilt discrimination task was then entered as a fixed parameter in the main experiment, as described in our previous experiments. In the main experiment, a 750 ms arrow pre-cued the target location with 100% validity. On each trial, a flanked Gabor target appeared on the cued side of fixation (left or right) along with one inward flanker and one outward, tilted in like fashion to those in the orientation calibration experiments. The observers’ task was to discriminate the central target Gabor’s tilt from vertical by pressing the corresponding mouse button (left for CCW tilts and right for CW tilts).

Results

Since we employed a 2 alternative discrimination task, the maximum possible TE given chance performance of 50% is 2. However, in the present experimental design performance was fixed at 80% correct for the narrowest spacing tested (i.e. the calibration condition was equivalent to the smallest spacing of 0.25 in the main experiment). This sets a different theoretical bound on TE as expected from calibrated

![Figure 13](image-url). Threshold elevation plotted against target-flanker spacing at 4° (near) and 8° (far) target eccentricity. Error bars represent mean ± SEM. Dashed line represents the 95% confidence interval about a TE of 1.
performance: \( P_{\text{max}}/P_{\text{calibrated}} \), yielding a value of 1.25. Our data suggest that observers reached half of this maximum at spacings of 0.25 times target eccentricity (Figure 12).

**Discussion**

Given our previous results showing benefits of endogenous attentional capture between 0 and 100 ms of SOA, observers should have ample time to properly attend to the target location in advance of each trial, and experience minimal cognitive load due to any endogenous shift of spatial or featural attention. If anything, the pre-cueing used here should reduce positional uncertainty about the target’s future location, and allow pre-allocation of endogenous spatial attention to that location in anticipation of the target’s appearance.

As such, endogenous spatial attention should be generally directed at least to the correct visual hemifield at the time of target onset. So, at that time, if exogenous and endogenous transient spatial attentional capture indeed cooperate to alleviate crowding, then we ought to expect similar results here (Figure 13) to those of the dual-task spacing experiment (Figure 11, panel d) that also used naïve observers. Instead, no effect of spacing is present in that instance, whereas removal of the relatively demanding letter identification task in this experiment results in a typical spatial profile of crowding, precisely following Bouma’s limit of half eccentricity.

This suggests that Bouma’s bound widens as a result of diverting endogenous spatial attention away from the target. First, across all the flanker-blink conditions tested in which similar flankers blinked at 0, 100 and 200 ms SOAs, we can find performance decrements in target orientation discrimination. Potentially, this occurs as a result of endogenous attention being misdirected towards the flankers’ features.
Meanwhile, the results of Experiment 1 show that target blinks induce improvements in performance, possibly via a similar mechanism: transient onsets strongly capture exogenous spatial attention. Moreover, this effect is generally larger when the targets and flankers differed along the endogenously attended feature of orientation.

In sum, the present experiment raises the possibility that the deleterious effect of flankers may persist in the absence of endogenous spatial or featural attention, but does not follow Bouma’s bound. At least under conditions of high attentional load, it seems that this effect could be akin to Pelli, Palomares & Majaj’s (2004) notion of temporal crowding.

**Chapter 6**

**General Discussion**

Across our experiments, we generally find our initial hypotheses confirmed. First, our results support the possibility that transient spatial attentional capture alone is insufficient to constrain the attentionally selected region to the target. While we observed several instances of improved performance resulting from application of onset transients to a target item, none of these occurred in the absence of endogenous spatial attention to the target’s location. Rather, in almost all instances, benefits in performance were confined to transients applied simultaneously with or just after an endogenous spatial cue.

One notable exception to this, however, was the generalized improvement in performance across blink conditions and SOAs for the expert observers, when the
flankers’ spatial frequencies were doubled with respect to the targets’ (Figure 8). However, the experts had potentially reached the performance benchmark of task independence. So, when we attempted to replicate this effect using naïve observers (Figure 11, panel c), we found no such benefit. This suggests that the results obtained for the experienced subjects were not the result of a pre-attentive process without capacity limits, as others have regarded feature singleton target visual search under conditions of exogenous attentional capture by feature pop-out (Treisman, 1982).

Rather, when the experienced subjects reached the point of task independence, they were capable of devoting spare endogenous spatial attentional resources to the targets’ locations even prior to the arrival of the letter cue. Since crowding preserves feature contrast (Petrov & Popple, 2007), this could be the result of prolonged exposures to our targets’ and flankers’ narrow band of orientations to which the dual task experiments were calibrated. This might have allowed observers to learn the feature contrast associated with left and right tilts (Sun et al., 2007). Thus, the narrowed target orientation bandwidth to which experts had become accustomed would be more highly weighted than flanker orientations or spatial frequencies to which the expert observers had not been exposed in other blocks.

The naïve observers, even with a simpler letter identification task to complete, were instead forced by the demands of the letter identification task to spend the majority of their attentional resources foveally before they could complete both tasks. It is likely the case that they simply could not devote any further endogenous spatial or featural attentional resources to the peripheral targets during the earlier portions of each trial. Consequently, they would have spent less time directing their attention to the relevant
feature contrasts, potentially missing the opportunity to reduce the flankers’ effects. Taken together, these results suggest that, at least under circumstances where covert endogenous spatial attention is directed away from the target, exogenous transient attentional capture fails to recover target discriminability.

Second, our results confirm that target-flanker feature contrast improves target discriminability when exogenous spatial attention is directed toward the target by transient attentional capture. However, given that our target-flanker separation manipulations under high attentional load did not produce typically observed spatial extents of the zone of interaction (i.e. critical spacing) in crowding, we cannot necessarily conclude that this effect is reducing crowding. Unsurprisingly, others have found similar results manipulating only endogenous attentional capture both spatially (Dakin, et al., 2009; Greenwood et. al., 2014) and temporally (Scolari et. al., 2007, but see Yeshurun and Rashal, 2010).

Across our experiments, we find consistent benefits to target discriminability at 0 and 100 ms SOAs, at least for narrowly spaced targets and flankers. At wider spacings, however, it seems that the effect of onset transients varies from observer to observer. For some, it produced the same benefits we observed in Experiments 1 and 4, but for others it produced no effect or even a negative effect (hinderance) on target discriminability. Such large inter-observer differences in peak latency suggest that attention shifting has variable time-course and spatial distribution across observers, and even from block to block within our experienced observers. Potentially, addition of reaction time measures to the current paradigm in future research could account for some of the variability found here in peak latency.
However, our results consistently indicate that in the absence of a transient blink, directing endogenously directing attention towards the peripheral target breaks Bouma’s bound. None of our observers, either naïve or experienced, saw significant increases in performance from near to far spacings (0.25 and 0.625 times target eccentricity) in the no-blink conditions. And yet, when the endogenous spatial attentional load was removed (as in Experiment 5) we obtained threshold elevations at spacings quite commensurate with Bouma’s bound of half eccentricity.

These findings raise the possibility that Bouma’s bound depends on the distribution of endogenous spatial attention. As discussed earlier, previous research by Dakin and colleagues (2009) suggests that the effects of attention and crowding on acquisition of average orientation information are independent of one another. According to their findings, loading endogenous spatial and feature attention with a demanding foveal task shifts thresholds upwards uniformly (i.e. a constant added to the distribution), indicating reduced capacity for global orientation processing. Addition of flankers instead widens the response distribution, indicating reduced capacity for fine local feature discriminations.

However, our manipulations of endogenous attentional load are accompanied by an additional covert shift to a peripheral target, and subsequent attentional capture. Presumably, under these specific circumstances we obtain an interaction in the sense that flanker spacing no longer obeys Bouma’s bound under our load conditions. Since the general scheme of our task is rather similar to Dakin and colleagues’ (2009), with only the attentional shift and capture as added parts, we suspect that one of these two elements might be at the root of the observed interaction.
And finally, *when exogenous spatial attention is directed to the flankers, we expected crowding would be exacerbated at greater feature contrasts than when it was directed to the target*. Although the results of Experiments 1 and 4 (panel (a) in both of Figures 6 and 11) strongly suggest that directing attention to the flankers increases their deleterious effect on performance, we find that in fact they only reduce performance when they are relatively similar to the targets, and only at positive SOAs. Conversely, when the flankers differed either in spatial frequency content or orientation, no additional degradation of performance was observed at any SOAs.

Taken together, these results suggest that target-flanker feature contrast is important in determining the flanker’s effect on target discriminability, but in the opposite sense that our original hypothesis predicts. Initially, we had assumed that biasing attention to the flankers would hinder discrimination simply by misdirecting spatial attention away from the target. However, it seems that this only happened in cases where the flankers were similar to the target (Figure 6, panel a).

Thus, it could be that transients applied to the flankers widened the attentionally selected region (rather than shifted it) to weight them more highly in feature averaging with the target. When flankers were sufficiently different, their features could have been processed in separate channels (as may have been the case for spatial frequency), or it could be that feature contrast, saliency or pop-out simply take precedence over transient exogenous attentional capture over time. Either way, this is not a case of the transient system pre-attentively encoding orientation or spatial frequency with relatively limitless capacity, otherwise this effect would have persisted at negative SOAs.
Finally, across experiments, we find that transients applied at short positive SOAs improve discriminability of a flanked target, but only do so consistently (though briefly) when applied to an endogenously attended target. In addition, this effect is more pronounced when the target and flankers are orientated orthogonally (Figure 1, panels a and c). This supports our general hypothesis that the combination of endogenous and exogenous spatial attention to the target can alleviate crowding, at least when the relevant feature contrasts occur along the endogenously attended feature (orientation). Whereas, in the absence of endogenous spatial attention, we find no evidence to suggest that transient blinks can alleviate spatial crowding.
References


