Object Perception, Attention, and Memory (OPAM)
2008 Conference Report
16th Annual Meeting, Chicago, IL, USA
Organized by Kim Curby, Sarah Shomstein, Joseph Brooks, and Artem Bepolsky

1093 Perceptual interactions between facial properties
Rama Amisha and Ruth Kimchi

1097 Delayed transfer to visual short-term memory follows express attentional shifts after an invalid predictive spatial cue: Evidence from the N2pc and SPCN
Benoit Brisson and Pierre Joliceur

1101 The influence of strategy on working memory guidance of attention
Nancy Carlisle, Leanne Boucher, and Geoffrey F. Woodman

1105 Object-based attention is “turned off” by top-down control
Louis K. H. Chan and William G. Hayward

1109 Auditory object-based attention
Lauren N. Hecht, Brandon Abbs, and Shaun P. Vecera

1115 Individual differences in perceptual sensitivity to emotional human movement
Martha D. Kaiser and Maggie Shiffrar

1119 Context-free inhibition: Attentional biases transfer strongly across temporal and spatial search tasks
Brian R. Levinthal and Alejandro Lleras

1123 Objects in scenes: Is one system enough?
Michael L. Mack and Thomas J. Palmeri

1127 Face recognition with contrast chimeras
Ming Meng, Sharon Gilad, and Pawan Sinha

1131 Defining perceptual load: The role of local competitive interactions in visual cortex
Ana Torralbo and Diane Beck

1135 Perceptual organization across hemispheres
M. van Assche, P. Gos, and A. Giersch

1140 Multisensory synchrony guides attention in dynamic cluttered environments
Erik van der Burg, Christian N. L. Olivets, Adelbert W. Bronkhorst, Dark Talsma, and Jan Theeuwes

1144 Selective attention of working memory contents by 11-month-old infants
Mariko Yamaguchi and Lisa Feigenson
Perceptual interactions between facial properties

Rama Amishav and Ruth Kimchi

University of Haifa, Israel

The consensus view among students of face perception is that faces are processed holistically/configurally. However, the exact nature of these processes is still under dispute. The holistic approach (e.g., Tanaka & Farah, 2003) suggests that faces are represented as unified Gestalts, in which facial components and the spatial relations between them are fused together into an undifferentiated template. The configural approach (e.g., Diamond & Carey, 1986; Rhodes, Brake, & Atkinson, 1993), and in particular the dual-mode hypothesis (e.g., Searcy & Bartlett, 1996), assumes that face perception is supported by two sources of information that are processed independently—componential (e.g., eyes) and configural (e.g., inter-eyes distance). The processing of upright faces is dominated by configural properties, whereas components dominate the processing of inverted faces.

We report a novel attempt to understand the relations between configural and component properties in face perception by using Garner's speeded classification paradigm (Garner, 1974). This paradigm provides a rigorous test of perceptual interaction between stimulus dimensions. Participants classify stimuli (e.g., faces) on a relevant dimension (e.g., components) while ignoring variation on an irrelevant dimension (e.g., configural properties), in two conditions. In the control condition only the relevant dimension varies while the irrelevant dimension is held constant at each of its two values. In the filtering condition, both the relevant and the irrelevant dimensions vary independently. Equal performance in the control and filtering conditions indicates that participants are able to focus exclusively on the relevant dimension, and the dimensions are considered perceptually independent. Poorer performance in the filtering condition than the control condition—Garner interference—indicates that it is not possible to selectively attend to one dimension while ignoring irrelevant variations, and the dimensions are considered perceptually dependent.

Address correspondence to: Rama Amishav, Department of Psychology, University of Haifa, Haifa 31905, Israel. E-mail: arama@research.haifa.ac.il
EXPERIMENT 1

Four faces (Figure 1A) were created by orthogonally combining two sets of components (eyes, nose, mouth) with two sets of configural properties (inter-eyes distance, nose–mouth distance). A single face was presented on each trial, and the participant’s task was to classify it on either component or configural properties as quickly and as accurately as possible. The experiment employed three orthogonally combined factors: Task (components judgement, configural judgement) and condition (control, filtering) as within-subjects factors, and orientation (upright, inverted) as a between-subjects factor. For components judgement, one control condition involved faces a1 and a3, and the other faces a2 and a4. The filtering condition required discriminating faces a1 and a2 from faces a3 and a4. For the configural judgement, one control condition involved faces a1 and a2, and the other faces a3 and a4. The filtering condition required discriminating faces a1 and a3 from faces a2 and a4.

Our results show no response time (RT) difference between the components and the configural judgements in their control conditions, indicating that component and configural properties were equally discriminable. Importantly, the results show a symmetric Garner-interference effect for upright faces. RTs in the filtering condition were significantly longer than RTs in the control condition by 44 ms for components judgements and by 29 ms for configural judgements. The difference in the magnitude of interference for the two tasks was not significant. Performance with inverted faces showed a significant asymmetric Garner interference (averaged 57 ms) only from component to configural properties.

These results indicate that component and configural properties interact in the processing of upright faces, and that processing of inverted faces is dominated by components. These results are apparently congruent with the holistic view. To further examine it, the next experiments tested the perceptual interaction between components (Experiment 2) and between configural properties (Experiment 3). If faces are processed as unified Gestalts, as argued by the holistic view, then selective attention to different components or to different configural properties should not be possible.

EXPERIMENTS 2 AND 3

The faces for Experiment 2 were created by orthogonally combining two eyes with two mouths (Figure 1B), and participants classified the faces on either the shape of the eyes or the shape of the mouth. The faces for Experiment 3 were created by orthogonally combining two inter-eyes distances with two nose–mouth distances (Figure 1C), and participants
Figure 1. The sets of faces for (A) Experiment 1—faces vary in components (eyes, nose, mouth) and configural properties (inter-eyes and nose-mouth distance), (B) Experiment 2—faces vary in eyes and mouth, and (C) Experiment 3—faces vary in inter-eyes and nose-mouth distance.
classified the faces on either inter-eyes distance or nose-mouth distance. In both experiments no Garner interference was observed, regardless of orientation. When faces varied only in components, participants could selectively attend to one component (e.g., eyes) while ignoring the other (e.g., mouth), and vice versa. Likewise, when faces varied only in configural properties, participants could selectively attend to one property (e.g., inter-eyes distance) while ignoring the other (e.g., nose-mouth distance), and vice versa.

**DISCUSSION**

The results of this study demonstrate perceptual dependence between component and configural properties in the processing of upright faces, with no relative dominance of one over the other. Participants are unable to selectively attend to components while ignoring irrelevant variation in configural properties, and vice versa. In contrast, the processing of inverted faces is dominated by components. The results further show that when faces vary only in components or only in configural properties, participants can selectively attend to one component while ignoring irrelevant variations in another component, or to selectively attend to one configural property while ignoring irrelevant variations in another configural property, when the task calls for it.

These results are not congruent with the holistic view, which assumes that faces are processed as undifferentiated Gestalts, nor with the configural view, in particular the dual-mode hypothesis, which assumes that component and configural properties are processed independently, with relative dominance of configural properties in processing of upright faces.

Rather, our findings clearly show that component and configural properties can be explicitly represented, and when faces vary only in components or in configural properties, different components or different configural properties can be processed independently. However, when both component and configural properties vary, as is the case with real faces, these two types of properties interact in the processing of upright faces. These findings strongly suggest that it is the mutual perceptual interaction between component and configural properties that is the hallmark of upright face perception.

**REFERENCES**

Delayed transfer to visual short-term memory follows express attentional shifts after an invalid predictive spatial cue: Evidence from the N2pc and SPCN

Benoit Brisson and Pierre Jolicœur

Centre de recherche en Neuropsychologie et Cognition, Université de Montréal, Montréal, Quebec, Canada

In spatial cueing experiments using predictive cues, response times (RT) are shorter when the target appears at the cued location (valid trials) than when it appears at an uncued location (invalid trials). This RT effect is often accompanied by an enhancement—but no effects on the latency—of early occipital P1 and/or N1 event-related potential (ERP) components (Mangun, 1995). It is still not clear where in the processing stream the first latency effects occur.

To investigate this question, we focused on two lateralized ERP components: The N2pc (N2 posterior contralateral) and the SPCN (sustained posterior contralateral negativity). The N2pc is thought to reflect visual selective attention mechanisms that separate relevant and irrelevant perceptual information in bilateral, multielement search arrays (Eimer, 1996; Luck & Hillyard, 1994). It is a negative deflection that is maximal at posterior electrode sites contralateral to an attended item (e.g., larger negativity over the right hemisphere if the target is presented in the left hemifield). It typically

Address correspondence to: Benoit Brisson, Département de Psychologie, Université de Montréal, CP 6128, succursale Centre-ville, Montréal, Québec H3C 3J7, Canada. E-mail: benoit.brisson@umontreal.ca
starts about 180 ms post-target onset, lasts about 100 ms, and can be isolated by subtracting the ipsilateral waveform (average of left-sided electrode with left visual field target and right-sided electrode with right visual field target) from the contralateral waveform (average of left-sided electrode with right visual field target and right-sided electrode with left visual field target). The SPCN, which follows the N2pc in the contralateral minus ipsilateral difference wave, is thought to reflect visual short-term memory activity (Jolicœur, Brisson, & Robitaille, 2008; Vogel & Machizawa, 2004).

**STIMULI AND PROCEDURE**

Four equiluminant colours (red, blue, green, and yellow) were used in each trial. Two of these colours were defined as target colours and two were defined as distractor colours for the entire test session for any given participant (colour assignments were counterbalanced between participants). The cue consisted of a brief colour change of two grey placeholders (one on each side of fixation) that surrounded the upcoming target and distractor locations. One placeholder turned into a target colour and the other into a distractor colour. The cue display was followed, at a stimulus–onset asynchrony of 800 ms, by a target display that contained a target and a distractor on 80% of trials (target-present trials) and two distractors on the remaining 20% of trials (target-absent trials). The items in the target display were squares with a gap in one side (different for each square), and these were always in a different colour than the colours used in the cue display.

The target-coloured placeholder, which indicated the most probable location of the upcoming target (in 75% of target-present trials), appeared randomly to the left and right of fixation and appeared randomly in each of the two possible target colours. A speeded four-choice response as to the location of the gap in the target square was required on each target-present trial. No responses were required in target-absent trials.

Target-absent trials were included to estimate cueing effects on the P1 and N1 components in absence of any target-related N2pc/SPCN. Contralateral and ipsilateral target-absent waveforms were defined in respect to the location of the target-coloured placeholder, whereas contralateral and ipsilateral waveforms were defined in respect to the target location in target-present trials. Therefore, to eliminate overlapping P1/N1 cueing effects in valid trials, the contralateral target-absent waveform was subtracted from the contralateral target-present waveform, and the ipsilateral target-absent waveform was subtracted from the ipsilateral target-present waveform, because the target-coloured placeholder and the target were in the same hemifield. In invalid trials, however, the overlapping P1/N1 cueing effects was eliminated by subtracting the ipsilateral target-absent waveform.
from the contralateral target-present waveform, and the contralateral target-absent waveform from the ipsilateral target-present waveform, because the target-coloured placeholder and the target were in opposite hemifields. N2pc and SPCN measures were obtained from these target-present minus target-absent subtraction waveforms.

RESULTS

Shorter RTs were produced in valid (685 ms) than in invalid trials (714 ms), \( F(1, 15) = 16.04, p < .002 \). As is often the case with highly visible, unmasked targets, no validity effects was observed on accuracy (95.5% in valid trials and 94.9% in invalid trials), \( F(1, 15) = 1.29, p > .27 \). Mean percentage of false alarms in target-absent trials was 1.13%.

The P1 mean amplitude (90–130 ms post-target display) was larger contralateral to the attended (cued) side, \( F(1, 15) = 17.77, p < .001 \), and this P1 cueing effect was similar across valid, invalid, and target-absent trial types \( (F < 1) \), indicating that the target-absent condition provides a good estimate of cue-related attentional activity in absence of any target related N2pc/SPCN. No correlation was observed between the P1 effect, collapsed across trial type, and the RT effect, \( r = .081, p > .76 \).

The contralateral minus ipsilateral difference waves for which the target-absent activity was subtracted is presented in Figure 1. The N2pc mean amplitude (210–290 ms) was similar in valid and invalid trials \( (F < 1) \). A jackknife analysis (see Kiesel, Miller, Jolicoeur, & Brisson, 2008) revealed no hint of any effect of cue validity on N2pc latency \( (F_{\text{adjusted}} < 1) \).

As for the SPCN, mean amplitude in the 350–400 ms time range was larger in valid trials than in invalid trials, \( F(1, 15) = 5.24, p < .04 \), but no difference was observed in the later 410–460 ms time range \( (F < 1) \). Importantly, a jackknife analysis revealed that the SPCN had an earlier onset in the valid condition than in the invalid condition, \( F_{\text{adjusted}}(1, 15) = 32.89, p < .0001 \), and the SPCN latency effect was correlated with the RT effect, \( r = .503, p < .047 \). The scalp distributions of the SPCN were typical and similar between the valid and invalid conditions.

---

1 Given that the SPCN follows the N2pc, an earlier onset of the former means a greater overlap with the latter. In this situation, the SPCN is thus superimposed on the greater negativity of the N2pc, modulating its amplitude. In order to minimize the variability caused by differences in N2pc amplitude at the onset of the SPCN, we did not use a fixed amplitude criterion to measure the SPCN latency, as usually done, but instead we measured the time at which the SPCN, in each N–1 waveform, reached half of its maximum (peak amplitude) minus minimum (the junction between N2pc and SPCN) amplitude. A 10-Hz low-pass filter was applied on the waveforms prior to the jackknife analyses, and the SPCN onsets of each participant in each condition was recovered from the jackknife values.
DISCUSSION

The locus of the spatial cueing RT effect was tracked by measuring the onset latency of two lateralized ERP components: The N2pc, an index of visual selective attention, and the SPCN, an index of maintenance activity in visual short-term memory. The first important result was that both the onset and the amplitude of the N2pc were unaffected by cue validity, despite the fact that a larger P1 amplitude was observed contralateral to the attended (cued) location. The N2pc results demonstrate that although there was a clear benefit for targets presented at the attended location rather than at the unattended location, in terms of mean RT, the amplitude change in the P1 component apparently does not necessarily translate to an immediate acceleration of visual target selection, as indexed by the N2pc. In contrast, the SPCN occurred earlier in valid trials, and a positive correlation was observed between cueing effects on SPCN latency and on RTs, suggesting that SPCN latency differences contribute to the RT effect. In sum, the present results strongly suggest that although predictive spatial cueing affects the amplitude of early ERP components, such as the occipital P1, the first stage of processing in which such early amplitude effects are translated into latency differences that could explain the observed RT effects occur after visual selection processes giving rise to the N2pc, but at (or before) transfer in visual short-term memory, as reflected by the SPCN.

Figure 1. Contralateral minus ipsilateral corrected difference waves time-locked to target onset at PO7/PO8 sites. Shown are the 10-Hz low-pass filtered waveforms.
REFERENCES


The influence of strategy on working memory guidance of attention

Nancy Carlisle, Leanne Boucher, and Geoffrey F. Woodman

*Vanderbilt University, Nashville, TN, USA*

Spatial working memory and spatial attention have been shown to rely on the same resources, but the attempts at determining the relationship between object-based working memory and attention have led to mixed results (Awh, Vogel, & Oh, 2006). The Biased Competition Model of attention (Desimone & Duncan, 1995) proposes that objects in working memory—target templates—will guide attention to similar items in a search array. Subsequent studies testing the prediction that items held in working memory will guide attention to similar items in a search array have led to disparate results. Some studies have found positive guidance where attention is guided to items matching the contents of working memory (Downing, 1999; Huang & Pashler, 2007; Olivers, Meijer, & Theeuwes, 2006) with some suggesting that such guidance is involuntary (Soto, Heinke, Humphreys, & Blanco, 2005). Others have found no relationship or evidence of negative guidance where the memory-matching item in the search array may be avoided in order to perform search more efficiently (Downing & Dodds, 2004;
Houtkamp & Roelfsema, 2006; Woodman & Luck, 2007). Woodman and Luck (2007) proposed that differences in participants' strategies might explain both positive and negative guidance, with attention being directed to the memory-matching items in order to perceptually resample the memory information in order to increase memory performance.

In the present work, we examined whether strategy can alter the relationship between objects held in working memory and the direction of attention. Participants held an object in working memory while performing a search task for a shape-defined target in an array of six objects. On half the trials, a memory-matching distractor was present in the search array. In order to assess whether attention was directed to the memory-matching distractor, we took both behavioural measures of attentional direction and directly measured overt attention by tracking eye movements. For the behavioural measures of attentional direction, we compared the reaction times to the search array when no memory-matching distractor was present in the array to the reaction time when a memory-matching distractor was in the array. If items in memory guide attention, we predicted longer search responses when the memory-matching distractor is present in the array. Similarly, we predicted more fixations to the memory-matching distractor in the search array than to the other distractors if items held in working memory guide attention.

In Experiment 1, we found large individual differences in the object first fixated by participants. Half of the participants attended the memory-matching distractor at least three times more often than the average of the non-memory-matching distractors; the other group did not show a preference for the memory-matching item over the other distractors and actually attended the memory-matching distractor less often than the average of the other distractors (Figure 1A). We grouped the participants that frequently fixated the memory-matching item into the “match-attend” group and the subjects that avoided the memory match the “match-reject” group. The groups continued their heterogeneous patterns of attention to the memory-matching item until the search response was made, at which point both groups attended the memory-matching item frequently. The manual search responses were consistent with the measures of overt attention (Figure 1B). The match-reject group had a negative guidance effect of memory match on RT, showing a decreased reaction time to the search array when a memory-matching item was present. In contrast, the match attend group showed a nonsignificant positive guidance effect of memory match on RT, with increased reaction times when the memory-matching item was present. These results are in line with the proposal of Woodman and Luck (2007) that strategy influences whether items held in working memory can guide attention.
Figure 1. Error bars represent 95% between-subjects confidence intervals for A and B, and 95% within-subjects confidence intervals for C and D. (A) The proportion of first saccades directed to the memory-matching distractor for the match attend and match reject group. (B) The reaction time guidance effect for the match attend and match reject groups. The reaction time guidance effect is computed as the increase in reaction time when the memory match is present in the array (search RT memory match present/search RT memory match absent). Positive values indicate attention being directed towards the memory-matching item; negative values indicate attention being directed away from the memory-matching item. (C) The proportion of first saccades directed to the memory-matching distractor for the baseline (white bars) and instructional (black bars) session for the memory stress and search stress groups. (D) The reaction time guidance effect for the baseline and instructional session for the memory stress and search stress groups.
In order to further explore this prediction, in Experiment 2 we directly manipulated participants’ strategies. In a baseline session, participants were told to perform the search task as quickly as possible and the memory task as accurately as possible, as in Experiment 1. In the subsequent instructional session, participants were told to stress either search speed or memory accuracy. If attention to memory-matching items could be strategically altered, we expected the instructions to alter the pattern of attention exhibited by the participants. First, we found large individual differences in whether the memory-matching item was attended on the first fixation in the baseline session, replicating Experiment 1. We then focused on the effect of the instructions given at the beginning of the instruction session. We found that the search-stress instructions reduced the proportion of first saccades directed to the memory-matching distractor compared to the baseline session (Figure 1C). Additionally, we found an interaction of session and instruction condition on the memory guidance effect in RT (Figure 1D). The search-stress group showed a reduced positive guidance effect of memory match on RT in the instruction session, whereas the memory stress group switched from a negative to a positive guidance effect of memory match on RT. Thus, both the eye-tracking and behavioural results show that instructions can influence whether attention is directed to items matching those held in working memory. These findings clearly imply that strategy has an effect on how working memory is used to guide attention.

The results of these experiments help to resolve the disparate results in the literature relating to the interaction of object-based attention and working memory. It seems that attention can be guided towards items matching those in working memory when it is beneficial for the task at hand, but that working memory does not obligatorily guide attention. The results also inform theories of attention and working memory, indicating that strategy should be included in the discussion of the link between object-based attention and working memory. Further studies will need to examine what factors influence strategies.

REFERENCES


Object-based attention is “turned off” by top-down control

Louis K. H. Chan and William G. Hayward

University of Hong Kong, Pokfulam, Hong Kong

Object-based attention is manifested by better detection performance at a spatial location within an attended object, as compared to an unattended object (e.g., Egly, Driver, & Rafal, 1994). Shomstein and Yantis (2002) discussed two basic approaches that may account for this object-based effect. First, the early sensory enhancement account suggests that attention automatically spreads along the boundary of an object. Second, the attentional prioritization account proposes that higher attentional priority is given to locations that are within an attended object. Different mechanisms for object-based attention may predict different degrees of top-down control on object-based effects. For instance, attentional prioritization is thought to be largely configurable (e.g., Wolfe, 1994), whereas sensory enhancement effects appear to be relatively automatic. Previous studies have generally observed that object-based effects are reduced or even eliminated when focal attention is guided by other space-based, contextual information (Goldsmith & Yeari, 2003; Shomstein & Yantis, 2004; Yeari & Goldsmith, 2006). However, although object-based effects can be offset by deliberately engaging other spatial cues, this does not mean that the

Address correspondence to: Louis Chan, Department of Psychology, University of Hong Kong, Pokfulam, Hong Kong. E-mail: clouis@graduate.hku.hk
operation of object-based attention itself was moderated. Therefore, it remains an open question as to whether object-based attention can be modified by top-down control.

The present study is designed to address this issue. A direct way to answer this question is to create a situation in which object-based attention will impair a task. If an object-based effect is still observed in this situation, this would mean that object-based attention is obligatory; if the object-based effect is eliminated, this would mean object-based attention is modifiable by top-down processes.

EXPERIMENTS 1 AND 2

In this study, we adopted a paradigm modified from Egly et al.’s (1994) “two-rectangle” method. In their study, one of the four ends of two vertical/horizontal rectangles was cued. The cue was valid in most trials, but the target also occasionally appeared at either of the two locations adjacent to the cue, one at the cued object and one at the uncued object. Performance was measured, and a quicker response was found for uncued targets at a cued object than at an uncued object, reflecting an object-based effect. Our method is illustrated in Figure 1. In each trial, a preview display (fixation and rectangles) was first showed for 1000 ms. After that, a red cue appeared for 50 ms. This cue matched the search target colour and so should attract attention obligatorily (Folk, Remington, & Johnston, 1992). This notion was confirmed in the first experiment. After 100 ms since the onset of the cue, four visual search items in different colours were displayed at each end of the rectangles. Participants searched for the red target and responded to its orientation.

In the first experiment, we demonstrated an object-based effect and that the cue captured attention. The target appeared at the cued, same-object, different-object, and diagonal-from-cue positions for 10%, 40%, 40%, and 10% of trials, respectively. We found the quickest response in validly cued trials (489 ms vs. 563 ms, 572 ms, and 587 ms, respectively), all \( p < .001 \), even though the cue was rarely valid. We found quicker response in same-object than different-object trials (563 ms vs. 572 ms, 9.0 ms), \( p < .03 \).

In the second experiment, we wanted to create a situation in which the use of object-based attention would only impair performance. One way to achieve this was to include a condition in which the target appeared at the unattended object in most trials. In two conditions, the target either appeared more (80% vs. 20% of trials) at the same-object position or more at the different-object position, and it did not appear at the other two positions. Since we were not studying whether space-based attention can overrule object-based effects, in the experiment the object orientations were
changing across trials so that the cue did not predict the absolute target position. As a result, we observed a same-object advantage in the more-same-object condition (576 ms vs. 586 ms, 10.7 ms), \( p < .01 \). However, in the more-different-object condition, a same-object advantage was not observed (584 ms vs. 583 ms, \( -1.5 \) ms), \( p > .3 \), and was significantly smaller than that of the more-same-object condition, \( p < .01 \). Hence, the present results suggested that the operation of object-based attention was “turned off” when it was disadvantageous to exhibit object-based attention.

Figure 1. (A) Procedures of Experiments 1 and 2. In each trial, participants searched for a red target among a green, a blue, and a yellow distractor, and indicated whether the target was a C or a mirrored-C shape. The rectangles may be vertical or horizontal, varied randomly across trials. (B) Results of Experiment 2. An object-based effect was only observed when the target was likely to occur at the cued object.
DISCUSSION

In two experiments, we demonstrated that object-based attention is configurable by top-down control. We replicated the object-based effect when it was neutral or advantageous for attention to prefer locations in an attended object. However, we failed to demonstrate any object-based effect when object-based attention was unfavourable to performance. Object-based attention can be deliberately suppressed even in the absence of other competing attentional cues. We did not observe a reversed object-based effect. This means that object-based attention was simply “turned off”, rather than an object-based process that inhibit same-object locations was at work.

These results have important implications for the understanding of the mechanisms underlying object-based attention. For instance, object-based attention may serve as a factor that influences attentional allocation, which is enabled by default. However, this factor can be flexibly weighted by top-down control. In this respect, our suggestion is more consistent with an attentional prioritization account, because a similar mechanism is commonly assumed for other prioritization principles in visual search (dimension-weighting account; Müller, Heller, & Ziegler, 1995). This analysis is also consistent with Shomstein and Yantis’ (2008) findings. They found that the magnitude of the object-based effect varied with the strength of object representation. It is conceptually similar to findings that a more salient perceptual dimension is a more dominant factor in attentional prioritization (Theeuwes, 1991).

REFERENCES


Auditory object-based attention

Lauren N. Hecht  
*University of Iowa, Iowa City, IA, USA*

Brandon Abbs  
*Harvard Medical School, Cambridge, MA, USA*

Shaun P. Vecera  
*University of Iowa, Iowa City, IA, USA*

Within the visual domain, there is ample evidence that attention selects objects. For example, when two objects are spatially superimposed, attention is most efficiently directed towards one object at a time. Dividing attention across objects, as measured by feature reports between objects rather than within, reduces accuracy in reporting information about the objects (e.g., Duncan, 1984). Although objects exist in other sensory modalities, object-based effects have not been reported beyond vision. For example, auditory attention has been studied extensively, but the focus has been selection using a single feature (e.g., a low-probability frequency among high-probability frequencies), or selection based on space.

Modern theories of auditory perception have described the representation of auditory information in terms of either streams (e.g., Bregman, 1990) or objects (e.g., Griffiths & Warren, 2004; Kubovy & van Valkenburg, 2001) within auditory scenes. Under these schemes, one goal of the auditory system is to segregate different acoustic sources and determine the features of the environment that should be attributed to each source. Behavioural
evidence supports the existence of auditory perceptual organization processes.

For example, Watkins (1991) investigated speech recognition accompanied by irrelevant spectral variation (e.g., the reverberation of sounds off of surfaces), in which the auditory system must attend to features that are created by the source and ignore those that are created by echoes in the environment. Watkins proposed that this separation is accomplished by selecting an auditory stream based on the rate of change of its features, because spectral features of the source vary at a faster rate than those associated with echoes in the environment. This difference is due to the fact that the environment itself is changing at a slower rate than the source that is creating the target sound. The auditory system can then use these temporal differences to perform an auditory equivalent of figure–ground segregation. This correspondence between the problem faced by the auditory and visual systems gives one motivation for examining object-based effects in auditory attention.

Although Watkins’ (1991) study used speech stimuli, he noted that successful segregation may be stimulus general and depend upon spectral cues. Not only is the rate of change of spectral information necessary to select the stream, the actual spectral components themselves must be selected in order to recognize the speech. Thus, multiple features (i.e., spectral components) both segregate one object (i.e., stream) from another and allow object identification, a situation reminiscent of that which occurs in vision: Perceptual organization processes segregate objects from one another, and the results of these grouping processes then influence the allocation of attention, allowing attention to be directed towards objects (Vecera, 1994, 1997; Vecera & Farah, 1994).

Given the similarities between visual object-based attention and accounts of auditory perceptual organization, we asked if object-based effects exist in audition. To avoid space-based attentional selection, our current experiment created an auditory version of Duncan’s (1984) feature report task. Participants heard two simultaneously presented sounds: A violin and a French horn. The spectral cues to the horn timbre consisted of low-frequency harmonics of its waveform and an abrupt increase in the spectral envelope (or “attack”). In contrast, the violin timbre contained additional high-frequency harmonics and a slower change in envelope (see Figure 1). After a brief orientation to the auditory stimuli and their features (pitch: High or low; intensity: Loud or soft), participants practiced identifying features of a single sound until their accuracy was greater than 75% within a block of 24 trials (see Figure 2a).

Upon meeting this criterion, participants began trials during which the instruments were played over two speakers simultaneously, giving the impression of two instruments playing from the same spatial location.
Each instrument varied along two dimensions, pitch and intensity. Participants heard both instruments and reported two dimensions that were cued; the two dimensions came from either the same object (e.g., violin pitch and violin intensity) or from different objects (e.g., French horn pitch and violin intensity). See Figure 2b for trial events and their durations. If the streams or objects established by the auditory system drive later attentional selection, then participants should exhibit an auditory object-based attentional effect. Specifically, participants should be more accurate reporting two features from the same auditory object than reporting from different auditory objects.

Data analyses revealed a significant main effect of object type, with greater accuracy when features came from the same object (78.5%) than when they came from different objects (73.5%). The first and second reports were equally accurate when averaged across object type; however, these reports did significantly interact with the object type (see Figure 3). Planned comparisons indicated that responses to the first same-object report (78.1%) were not statistically different from the second same-object report (79.0%), suggesting that participants were able to select multiple features of the same object without any cost in their performance. In contrast, responses were significantly more accurate to the first different-object report (75.8%) than the second different-object report (71.2%), demonstrating an accuracy decrement when attention was directed from one object to another. Taken together, these findings demonstrate auditory object-based attentional selection and suggest that audition may be similar to vision in that these early auditory objects direct attention in an object-based manner.

Importantly, our results demonstrate auditory attentional selection without spatial cues to differentiate objects. The two auditory objects were differentiated by their timbre, and participants were able to report the features (pitch and intensity) associated with a particular timbre. Various cues, such as temporal proximity (Näätänen & Winkler, 1999) or similarity in spectral properties (Bregman, 1990), aid in perceptual grouping of auditory features associated with auditory objects. These ideas are similar to corresponding ideas in vision, in which different Gestalt cues organize visual input. Of course, the cues for auditory organization will differ from those used in vision, although spatial cues may be used by both modalities. In both audition and vision, various stimulus cues exist to segregate objects from one another. Our findings show that in audition, these auditory objects, or perceptual groups, influence attentional selection.
Figure 1. Waveforms (top) and spectrograms (bottom) for (a) a soft, high-pitched horn, (b) a soft, low-pitched violin, and (c) both playing together. The x-axis for each of the panels represents time (0-500 ms). These two components sum to produce the overlapping sound depicted in (c).
Figure 2. Events and their durations for a trial in the (a) single-sound practice block and (b) dual-report blocks.
REFERENCES


Figure 3. Mean accuracy for dual-report trials in the same-object and different-object conditions. Error bars indicate 95% within-subjects confidence intervals.
Individual differences in perceptual sensitivity to emotional human movement

Martha D. Kaiser and Maggie Shiffrar

Department of Psychology, Rutgers University, Newark, NJ, USA

Previous research has shown that typical adults demonstrate a heightened visual sensitivity to potentially threatening biological motion (Chouchour-elou, Matsuka, Harber, & Shiffrar, 2006). Namely, typical adults are more sensitive to the presence of an angry point-light walker as compared to the presence of happy, fearful, sad, or neutral walkers. This is known as the anger-superiority effect. In a series of psychophysical studies, we asked whether individual differences in social skills influence the visual analysis of emotional human motion. Typical adults display a wide range of social abilities. Autism is a disorder that compromises social ability. Importantly, autism is a continuous disorder and autistic traits have been shown to exist in the general population. Given the associations between autism and dysfunction of the pSTS (a neural area involved in the visual detection of human movement; Boddaert et al., 2004), the amygdala (a neural area involved in emotion recognition and threat detection; Howard et al., 2000), and their interconnection, we examined the relationships between the magnitude of autistic traits in the general population and visual sensitivity to emotional human action. We predicted that individuals with excellent social skills (i.e., few autistic traits) would show levels of visual sensitivity to human motion that were emotion dependent, whereas individuals with poor social skills (i.e., more autistic traits) would show emotion-independent visual sensitivity to human motion.

METHODS

Twenty-three Rutgers undergraduates took part in the study. Participants completed the Autism-Spectrum Quotient (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), a self-administered questionnaire that
measures the amount of autistic traits in adults with normal intelligence. Participants also completed a psychophysical task in which they viewed masked point-light displays of walkers displaying anger, fear, happiness, neutrality, or sadness. The order of AQ questionnaire administration and psychophysical task were counterbalanced across participants. Participants were not informed of the emotional content in the displays. Rather, their task was simply to detect the presence of a person within the cloud of dots (point-light mask). On each trial, participants viewed a 3 s movie. Each movie contained coherent or scrambled walker within a mask of scrambled human movement. The masking elements were individually constructed for each walker so that the points in each mask had motion trajectories that were identical to the points of the scrambled or coherent walker.

RESULTS

The participants’ AQ scores fell in the expected normal range (mean = 17.35, SD = 4.69). For analysis, we created two groups of participants using a median split of AQ score. The dependent measure on the psychophysical task was perceptual sensitivity to the point-light walkers. To measure this, we calculated \( d' \) scores for individuals’ sensitivity to each emotional category of walker; these were computed by subtracting the standardized rate of false alarms from the standardized rate of hits. Overall, as expected, subjects were most sensitive to the presence of angry walkers. However, the low and high AQ groups displayed different patterns of results (see Figure 1). An ANOVA with AQ group as the between subjects variable showed a significant effect of autistic traits on sensitivity to angry walkers, \( F(21, 1) = 4.719, p < .05. \) Notably, AQ score did not influence sensitivity to walkers exhibiting any other emotions. The low AQ group demonstrated heightened sensitivity to angry walkers versus the other emotional movements, but the high AQ group failed to exhibit this perceptual advantage. Individuals with few autistic traits exhibited the anger-superiority effect, where angry walkers were detected with greater accuracy than walkers expressing other emotions.

CONTROL EXPERIMENT

Do socially adept observers demonstrate heightened visual sensitivity to angry point-light walkers because perceptual processes are enhanced for potentially threatening stimuli or because angry movement consists of velocity profiles with abrupt transitions? In other words, do AQ scores correlate with detecting the presence of angry emotional content per se, or with detecting the presence of abruptly changing velocities in general? To address this question, the previous experiment was replicated with inverted
Figure 1. Sensitivity to emotions by AQ group.
displays. Specifically, a different group of participants again reported whether or not a coherent, but now inverted, point-light walker was present within a scrambled mask. The results of this control experiment indicated that the anger-superiority effect decreased significantly with inverted displays. Since the velocity profiles are identical in the upright and inverted displays, these results suggest that visual sensitivity to angry human movements, per se, is modulated by the observer’s social abilities.

**DISCUSSION AND IMPLICATIONS**

This set of psychophysical experiments systematically examined the relationship between social skills and perceptual sensitivity to biological motion. A median split of AQ scores found that individuals with high AQ scores did not demonstrate heightened sensitivity to point-light displays of angry, and thus potentially threatening, human gaits. Conversely, our results also indicate that observers with more autistic traits exhibit relatively decreased visual sensitivity to such potential threat. Such decreased sensitivity to anger is interesting because it (1) supports the idea that autistic traits extend into the general population, (2) compliments and extends findings of impaired visual analysis of biological motion in individuals with autism (Blake, Turner, Smoski, Pozdol, & Stone, 2003; Kaiser, Delmolino, & Shiffrar, 2008), and (3) highlights the critical connection between social skills and sensitivity to negative emotions as essential for social survival.

A large body of research has shown that visual analysis of human movement differs from the analysis of object movement and is supported by distinct neural mechanisms in typical adults. Many studies have assumed that this privileged analysis of human action is characteristic of the typical population. This research indicates that individual differences in social skills may influence, or be influenced by, perceptual sensitivity to emotional movements.

**REFERENCES**


Context-free inhibition: Attentional biases transfer strongly across temporal and spatial search tasks

Brian R. Levinthal and Alejandro Lleras

University of Illinois at Urbana-Champaign, Champaign, IL, USA

Visual search has traditionally been studied in the context of bottom-up (display-driven) and top-down (knowledge-driven) factors (Treisman & Gelade, 1980; Wolfe, 1994). Recently, there has also been a growing interest in the role of recent experience on visual search performance (e.g., intertrial effects; Maljkovic & Nakayama, 1994). Our research focuses on one specific intertrial effect in visual search: The distractor previewing effect (DPE; Goolsby, Grabowecky, & Suzuki, 2005; Lleras, Kawahara, Wan, & Ariga, 2008). The DPE occurs in the context of an oddball search task, when target-present trials are preceded by target-absent trials (i.e., a trial with no oddball). The DPE refers to the finding that an observer’s ability to select an oddball target is impaired when on the immediately preceding target-absent trial, all the distractors belonged to the same visual category as the current oddball target (e.g., all prior distractors were green, and the current target is green).

The DPE and most intertrial effects have been primarily studied within the context of spatial search tasks in which all stimuli are presented simultaneously on a single display. Recently, however, an analogous pattern of results was observed in the context of temporal search tasks, using rapid serial visual presentation stimuli (RSVP; see Lleras, Kawahara, Levinthal, & Polychronopoulos, 2008): When observers are asked to identify the odd-coloured letter in an RSVP stream of homogeneously coloured letters, selection of that target is impaired if the target has the same colour as the distractors on the immediately preceding RSVP stream.

Address correspondence to: Brian R. Levinthal, Department of Psychology, University of Illinois at Urbana-Champaign, 603 East Daniel Street, Champaign, IL 61820, USA. E-mail: levintha@illinois.edu
Our goal here was to study whether the inhibitory processes at play in these two tasks (spatial and temporal searches) rely upon common representations or whether the inhibitory effects are contextually specific to each search task. Will a bias against selecting red targets created by viewing a spatial search task affect selection of red targets in an RSVP stream? We employed a mixed paradigm in which participants randomly performed either a spatial or temporal search on every trial. This paradigm allowed us to compare inhibitory effects when the search task is repeated (within-task effect) to instances in which the search task switched between trials (between-task effect). To preview, our results show strong transfer effects: Inhibitory biases created on one trial affect selection processes on the subsequent trial, regardless of whether search type is repeated or switched across trials.

METHODS

Fourteen students from the University of Illinois, aged 21–29, participated in the experiment. Stimuli were presented on 17-inch monitors, controlled by Pentium-IV PCs. Stimuli consisted of alphanumeric characters presented in Arial font in red, green, or white.

A description of the paradigm can be seen in Figure 1. On any given trial, participants randomly performed one of two tasks. In the spatial search task, three large (76 point) characters were presented along an iso-acuity ellipse surrounding the centre of the screen. In the temporal search task, six small (26 point) characters appeared at the centre of the screen in rapid succession, each appearing for 32 ms and followed by a 90 ms blank interval. Participants searched these arrays for a colour oddball, and reported whether that target was a letter or a number. A variable 800–1300 ms delay followed all target-absent trials. Note: Spatial target-absent displays were presented for 600 ms, whereas spatial target-present displays were displayed until response. During temporal searches, a 2500 ms delay followed target present trials, which was deemed sufficient time to gather responses about the RSVP stream. A failure to respond within this interval was considered an incorrect response. Similarly, RTs in the spatial search task longer than 2000 ms were also discarded.

All target-present trials were categorized according to task (spatial or temporal), whether the current task was the same as on the preceding trial (repeat or switch), and with respect to relationship between the current target colour and the colour of distractors on the preceding target-absent trial (distractor-colour previewed, target-colour previewed).
Results

Overall accuracy for the temporal search task was 76% (RT = 883 ms). Overall RT for the spatial search task was 825 ms (accuracy = 94%).
Spatial search

We performed a repeated measures ANOVA including task repetition and feature preview. The analysis revealed a main effect of preview condition, $F(1, 13) = 80.12, p < .001$, but no main effect of task repetition and no interaction. Paired comparisons indicated that a significant intertrial effect occurred following both spatial search (60 ms), $t(13) = 4.67, p < .001$, and temporal search tasks (44 ms), $t(13) = 3.04, p < .009$.

Temporal search

Analyses were constrained to targets appearing in the third position in the RSVP stream, when inhibitory effects are largest (Lleras et al., 2008). We again performed a repeated measures ANOVA including task repetition and feature preview. As before, the analysis revealed a main effect of preview condition, $F(1, 13) = 20.84, p < .001$, but no main effect of task repetition nor an interaction. Paired comparisons indicated that a significant intertrial effect occurred following both a temporal search (25%), $t(13) = 6.47, p < .001$, and a spatial search (15%), $t(13) = 5.69, p < .001$.

DISCUSSION

When observers perform an oddball visual search, their performance is influenced by the preceding trial, such that they find it more difficult to select oddballs when they are of the same colour as the distractors on the previous trial. Our results show that this inhibitory bias of selective attention is not constrained to the search context in which it was created. In fact, this bias was equally strong regardless of whether it originated from viewing a spatial or temporal search display. The extent of this transfer provides strong evidence that the bias likely arises from nontask specific mechanisms or representations, and may be related to central decision-making processes regarding how information is to be interpreted by our attentional system (as distracting or not). In other words, when our attentional system applies a “distractor” tag to a category of stimulus, the tag exerts its influence on the future workings of selective attention, irrespective of the task context.

REFERENCES


Objects in scenes: Is one system enough?

Michael L. Mack and Thomas J. Palmeri

Vanderbilt University, Nashville, TN, USA

A recent study by Joubert, Rousselet, Fize, and Fabre-Thorpe (2007) reported an interesting interaction of scene and object perception. Natural versus man-made scenes were presented for 28 ms and participants responded with a go/no-go response for the category label given at the beginning of a block. Like previous studies, they found very accurate and fast scene categorization (~95% accuracy, 405 ms average response time [RT]). The novel twist was that the scene images used in this experiment either contained objects consistent with the scenes’ category (e.g., an urban street scene with a parked car) or contained objects inconsistent with the scenes’ category (e.g., an urban street scene with a large tree) (see Figure 1A). A post hoc analysis comparing these two types of scenes showed an inconsistent-object disadvantage, such that participants made more errors and were slower to respond when categorizing a scene containing an inconsistent object.

Joubert et al. (2007) proposed a two-system model to explain this interaction. Specifically, an object recognition system operates in parallel along with a scene categorization system. Each system extracts information that provides evidence for the natural versus man-made decision. When a scene contains an inconsistent object, the object system provides evidence that conflicts with the scene system, resulting in more errors and slower
response times. Although this two-system account is compelling, we explored a simpler alternative.

Previous work has shown that rapid scene perception is largely driven by coarse, global scene statistics (Oliva & Schyns, 1997; Schyns & Oliva, 1994). Global image properties such as the pattern of energy in the spatial frequency domain capture global structural elements of scenes that are diagnostic of a scene’s category. However, since these global statistics are calculated across the entire scene, they are not immune to the local elements of a scene. For example, the global scene statistics of a forest scene that includes a small shed will be influenced by the image properties of the shed and be slightly different than a forest scene with a large bush in place of the shed. This influence of specific objects on the global representation of scenes is a potential source of the conflict found by Joubert et al. (2007). Our aim in this study was to examine whether the interaction produced by scenes with consistent versus inconsistent objects could be explained by a single system model of scene perception.
MODEL

We tested a model of scene categorization that extended the work of Oliva and Torralba (2001). They proposed an explicit computational model of scene category recognition based on global image statistics. In its simplest form, the global spatial frequency content of each scene is sampled at a range of frequencies and orientations. Classification of scene category is performed by linear discriminant analysis of these spatial frequency representations based on a labelled training set of scenes. To provide an account of reaction time data, we extended this model to include a decisional mechanism based on a drift diffusion model (Ratcliff, 1978). Specifically, the quality with which a particular scene was categorized using the linear discriminator established the drift rate of the diffusion process over time towards either a natural or a man-made decision criterion. Accuracy was determined by whether the correct criterion was reached. Reaction time was determined by a constant time for sensory processing and the motor response (a free parameter) plus the time it took for the diffusion process to reach its decision criterion. The decision criterion was also a free parameter.

METHODS

Stimuli

Stimuli came from a freely available scene database (Oliva & Torralba, 2001) and consisted of natural (coast, mountain, forest, meadow) and man-made (urban street, building, city) scenes (similar to those used by Joubert et al., 2007). Scenes were further labelled as having objects consistent or inconsistent with the scene’s category.

Procedure

Testing the proposed model consisted of the following steps. First, scenes were split into three different sets: A training set of 100 natural and 100 man-made scenes, a testing set of consistent-object scenes, and a testing set of inconsistent-object scenes. The two parameters of the model were optimized to account for the mean accuracy (96%) and RT (409 ms) associated with consistent object scenes found by Joubert et al. (2007). Fixing the two parameters, the trained model was then used to categorize the remaining set of inconsistent-object scenes. This entire process was repeated with 20 randomly selected training and testing sets.
RESULTS

As shown in Figure 1B, the model successfully accounted for the accuracy and reaction times for both consistent- and inconsistent-object scenes found by Joubert et al. (2007). Planned comparisons confirmed a significant inconsistent-object disadvantage for both accuracy, $t(19) = 3.57, p = .002$, and reaction time, $t(19) = 2.54, p = .019$, predicted by the model.

DISCUSSION

The category of objects found within a scene influences both the speed and accuracy of categorizing that scene at a superordinate level (Joubert et al., 2007). One explanation for this result is a two-system account of distinct scene and object processing mechanisms that are in conflict. In this study, we instead found that a single system scene perception model could successfully account for these results. This model married a previously successful model of rapid scene categorization with a well-known model of perceptual decision making, resulting in a model that could account for both accuracy and reaction times. The presence of inconsistent objects in a scene influences the coarse global representation of the scene enough to account for lower accuracy and slower reaction times. Instead of distinct scene and object perception systems operating in parallel and arriving at conflicting categorizations, the inconsistent-object disadvantage can be explained by a single scene perception system that processes the global statistics found in natural scenes.

REFERENCES

Face recognition with contrast chimeras

Ming Meng

Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

Sharon Gilad

Department of Neurobiology, Weizmann Institute of Science, Rehovot, Israel

Pawan Sinha

Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

In principle, a photographic negative is exactly as informative as its positive counterpart. However, as anyone who has had to search through a roll of negatives for a snapshot of a particular person knows, contrast negation has dramatically adverse consequences on our ability to identify faces. Exploring the causes of this phenomenon is important for understanding the broader issue of the nature of information the visual system uses for face identification.

In our past work, we have found that polarity of contrast around the eyes is a remarkably stable feature, with the eyes usually being darker than the forehead and the cheeks (Sinha, 2002). The absolute magnitude of contrast across different regions of a face changes greatly under different imaging/lighting conditions, but the local polarity relationships between the eye regions and their neighbourhood are maintained in all but the most unnatural lighting setups (such as lighting a face from below). Dark or oily facial complexion also does not typically disrupt these relationships. The use of ordinal brightness relationships confers significant tolerance for photometric variations in images and may help explain perceptual and neural invariance to illumination changes.

Address correspondence to: Ming Meng, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, USA. E-mail: mmeng@mit.edu

MM and SG contributed equally.
The otherwise highly consistent polarity relations are reversed in negative face images. We hypothesized that this may be a factor leading to the poor recognizability of negated faces. To test this hypothesis, we created a set of “contrast chimeras”. These are faces that are photonegatives everywhere except in the eye region (thus preserving local contrast polarity in that neighbourhood). Such faces still have unnatural pigmentation and shading cues over much of their extents and present largely the same problems to shape-from-shading processes as the full negatives do. Explanations based on disruptions due to unnatural pigmentation (Bruce & Langton, 1994; Vuong, Peissig, Harrison, & Tarr, 2005) or incorrect shape-from-shading cues (Hill & Bruce, 1996; Johnston, Hill, & Carman, 1992; Kemp, Pike, White, & Musselman, 1996) would not, therefore, predict significant improvements in performance with such chimeric faces beyond the improvements derived from the intrinsic recognizability of the eyes. However, if performance in negatives is compromised, at least in part, because of the destruction of local polarity relations between the eyes and their neighbourhood, performance with contrast chimeras is expected to be significantly better than that with contrast negatives.

We tested subjects with 24 monochrome celebrity face images spatially normalized to have the same interpupillary distance. For each original image, we created three additional variants: A full-negative (“1”), the positive eyes on a head-silhouette (“2”), and a contrast chimera (“3”). Positive eyes noticeably boost perceived animacy of the faces. Condition 2 allowed us to determine whether any observed performance gains in the chimeric condition are due merely to the intrinsic recognizability of the eyes on their own.

Fifteen subjects were assigned randomly to three equal-sized groups. Subjects in a given group were shown all faces in one of the three experimental conditions. They were asked to name the individuals shown. They were subsequently shown all faces in the full-positive condition to determine which celebrities they were familiar with. Performance for each subject was recorded as the proportion of familiar faces recognized in the experimental condition they had been assigned to. The results are shown in Figure 1(a). We found recognition performance in the full-negative condition to be 54.35%. This is consistent with results from previous studies that have compared recognizability of negatives relative to the original images. Performance with the positive eyes on a head silhouette was poor, averaging 13.37%. However, performance with contrast chimeras was close to ceiling (mean: 92.32%). All pair-wise comparisons between these groups are statistically significant at $p < .01$.

To further examine the significance of ordinal relationships, we used fMRI to determine whether the inclusion or violation of these relations modulates neural responses to face images. We recorded brain responses to nonfamous faces in negative, positive, chimeric, and eyes on head-silhouette
Figure 1. (a) Recognition performance as a function of stimulus type. Each of the three experimental conditions (negatives, eyes, chimeras) was shown to separate groups of subjects, followed by the fully positive stimuli. (b) Averaged activations in the right fusiform face area corresponding to the four different facial image types. Percentage MR signal change was calculated by using activation at $t=0$ s of each trial as the baseline. Activations corresponding to full positives and chimeras are statistically indistinguishable from each other, but are significantly higher than those corresponding to the full negatives and the eyes on head silhouettes. Error bars indicate standard error.
conditions. Eight subjects were shown 25 instances each of these four classes of face images in M-sequences (Buracas & Boynton, 2002), and their brain activations were recorded using a rapid event-related design. For any one subject, a particular face appeared in only one of the four conditions. The condition in which a particular face was shown was counterbalanced across different subjects. Subjects were asked to continuously monitor the fixation cross and were not required to make any judgements regarding the faces shown. Figure 1(b) shows the averaged fMRI responses in the right fusiform face area (FFA; Kanwisher, McDermott, & Chun, 1997; identified via separate localizer runs for each subject) as sorted by the four conditions.

Consistent with past reports (George et al., 1999), fully positive faces led to significantly increased brain activity in the FFA relative to fully negative faces, $F(1, 40) = 4.8$, $p < .05$. More interestingly, responses evoked by contrast chimeras are also significantly greater than those evoked by fully negative faces, $F(1, 40) = 4.7$, $p < .05$. In fact, they are as high as, and statistically indistinguishable from, those evoked by fully positive faces, $F(1, 40) = 0.11$, $p = .74$. By contrast, eyes embedded in the head-silhouette lead to minimal increased activity in the FFA, much less than positives, $F(1, 40) = 13.76$, $p = .001$, and chimeras, $F(1, 40) = 7.18$, $p = .01$, and similar to full negatives, $F(1, 40) = 0.89$, $p = .35$. These results suggest that the regular polarity of contrast around the eyes is important for eliciting significant neural activation to faces. Although the FFA’s functional role in face identification remains unknown, it is clear that full neural processing of faces is contingent on the ordinal contrast relationships.

Taken together, our results cast new light on the long-standing observation regarding the difficulty of recognizing photographic negatives. They suggest that that the difficulty in analysing negative facial images may be driven in large part by the destruction of 2-D contrast polarity relations between some of the key regions of the face. More generally, these relations might be embodied in the visual system’s facial representations and serve as strong determinants of recognition performance.

REFERENCES


Defining perceptual load: The role of local competitive interactions in visual cortex

Ana Torralbo

Beckman Institute, University of Illinois, Urbana-Champaign, IL, USA

Diane Beck

Department of Psychology, University of Illinois, and Beckman Institute, University of Illinois, Urbana-Champaign, IL, USA

Perceptual load has been proposed as a determining factor in the degree to which unattended information is processed (Lavie, 2005). When the perceptual task is easy (e.g., finding an X target embedded in a field of Os), it does not exhaust perceptual capacity and the remaining resources spill over to the unattended information, which then exerts a strong effect on the performance. However, when the task demands are increased and thus more resources are needed (e.g., finding an X embedded in a field of heterogeneous letters), there exist fewer resources to process the unattended information, and thus the influence of these stimuli on performance is reduced. Although manipulations that increase visual processing demands have modulated the processing of unattended items, as predicted by perceptual load theory, a clear definition of perceptual load is still needed. What are the factors that increase perceptual load and ultimately modulate the distractor processing?

Address correspondence to: Ana Torralbo, Beckman Institute, University of Illinois, 405 N Mathews Ave., Urbana, IL 61801, USA. E-mail: atorralb@uiuc.edu
COMPETITIVE INTERACTIONS AS A DETERMINING FACTOR OF PERCEPTUAL LOAD

We propose that local competitive interactions that are known to modulate the representation of stimuli in visual cortex, might have a determining role in perceptual load (Torralbo & Beck, 2008). Specially, we appeal to evidence that stimuli presented simultaneously in the visual field are not processed independently, but instead interact in a manner that suggests a competition for neural representation (Desimone & Duncan, 1995). Furthermore, by directing attention to a particular item, these competitive interactions can be biased, and thus resolved, in favour of the target (Kastner, de Weerd, Desimone, & Ungerleider, 1998). We propose that these competitive interactions and the resulting top-down biasing mechanisms needed to resolve the competition, may determine the perceptual load of the task, and thus the degree to which the unattended information is filtered. In the experiments presented here (see Torralbo & Beck, 2008), we manipulate two variables that should modulate the degree of competitive interactions in visual cortex, to assess whether they also modulate distractor processing in the manner predicted by perceptual load theory.

Distance between items

One factor that has been shown to modulate the interactions between stimuli is the distance between the items: The closer the stimuli are in the space the more likely for them to compete for representation among the same group of cells in visual cortex. Thus, we varied the distance between items (i.e., the density) in a search array, while keeping set size, heterogeneity of target and nontargets, and target-distractor separation constant (see Figure 1a). We placed a distractor at fixation that could be response compatible or incompatible with respect to the target (e.g., X or Z, respectively, when the target was X). We predict that as the density of the display, and thus the competition among items, increases, a stronger top-down bias would be needed to overcome that competition and detect the target. This strong top-down bias will result in greater filtering of the remaining items in the display and thus the influence of the distractor would be reduced.

As expected, distractor processing was modulated by the density of the search array. When the density was high (i.e., high competition) processing of the distractor was reduced compared to when the density was low, $F(1, 15) = 14.91, p < .002$; the difference in reaction times between compatible and incompatible trials was 40 ms in low density blocks but only 24 ms in
high density blocks. Thus, increasing competitive interactions among the items in the search array produced a pattern of results similar to increasing perceptual load.

**Number of items projected to the same local regions of cortex**

One of the common characteristics of the competitive interaction is that they predominately occur among cells in local regions of cortex; that is, they occur most strongly among items processed in nearby regions of cortex. Thus, another factor that should modulate the degree to which the letters of the search array compete is whether or not the letters fall within the same or different hemifields. In the early to intermediate area of visual cortex, where many of these local competitive interactions have been seen, the right and left visual hemifields are processed separately by the two cerebral hemispheres. Therefore, the more items that appear within the same hemifield, the higher the degree of competitive interactions among those items. We achieved this hemifield manipulation by presenting a search array arranged in a diamond shape (see Figure 1b), such that the...
target could appear alone in a hemifield or share the hemifield with two nontarget letters. If perceptual load is determined by local competitive interactions, as we predict, then placing multiple items within the same hemifield should result in greater perceptual load and reduced distractor processing than when the target item appears alone within a hemifield.

As expected, the influence of the distractor placed at fixation was modulated by the number of items presented within the same visual field. When the target shared a hemifield with the nontargets (i.e., high competition) the processing of the distractor was reduced compared to when the target was placed alone within a hemifield, $F(1, 19) = 6.97, p < .016$; the difference in reaction times between compatible and incompatible trials was 27 ms when the target was alone in the hemifield but only 8 ms when the target was in the same hemifield as the nontargets. These results are consistent with our hypothesis that increasing local competitive interactions in visual cortex increases perceptual load.

**PERCEPTUAL RESOURCES AND NEURAL COMPETITION**

Our results suggest that competitive interactions among cells representing near by information may determine the perceptual load of the task, and thus the degree to which distracting information is ignored. The greater the competitive interactions among items in the search array, the stronger the attentional bias needed to overcome that competition and the greater the perceptual load of the task (Torralbo & Beck, 2008).

**REFERENCES**


Perceptual organization across hemispheres

M. van Assche
Centre Hospitalier Régional de Strasbourg, Department of Psychiatry, Strasbourg, France

P. Gos
Ecole Normale Supérieure, Lyon, France

A. Giersch
Centre Hospitalier Régional de Strasbourg, Department of Psychiatry, Strasbourg, France

After extracting local information, the visual system must reconstruct forms into global configurations, by means of grouping processes. Information processed in each cerebral hemisphere also requires to be integrated in one unique visual scene. This grouping across hemispheres can be explored by comparing what happens when information is displayed in the same or in different hemifields. The interhemispheric cost that has been described until now is rather low (e.g., Murray, Foxe, Higgins, Javitt, & Schroeder, 2001). Subjects are usually slowed down by only 10–30 ms when comparing objects presented across hemifields rather than within the same hemifield. However, this has been explored with short-duration stimulus presentation, to avoid ocular saccades. Here, stimuli stayed on the screen and eye movements were recorded in order to check that subjects maintained unchanged hemifields’ location. This allowed us to observe a much larger cost of across-hemifield presentation than in previous studies. We then replicated these results in several experiments and questioned to which extent such a cost is exogenously modulated by grouping factors, or endogenously by top-down
factors. Indeed, studies conducted in brain-lesioned patients have reported a partial visual extinction recovery when connectors are used as grouping clues to group objects across hemifields (Brooks, Wong, & Robertson, 2005). We thus expected similar results in our healthy volunteers, i.e., a decrease in the response time cost observed for presentation across hemifields. Second, we tested the impact of the global configuration of the whole figure used during a trial, each figure including six objects. Global processing is indeed preferentially achieved by one cerebral hemisphere (Fink et al., 1997). We checked whether facilitating the global processing of the figure by manipulating its global configuration reduces or not the interhemispheric transfer cost. Finally, we verified the impact of top-down control.

METHOD

We derived a new paradigm from the Repetition Discrimination Task (Beck & Palmer, 2002) to explore the visual information processing within and across hemispheres, and manipulated grouping processes by connectors (Palmer & Beck, 2007; Palmer & Rock, 1994). The Repetition Discrimination Task (Beck & Palmer, 2002) allows assessing perceptual organization abilities by disentangling low-level grouping and top-down modulations. Circles are alternately presented with squares within an array. However, the alternation is never perfect and two adjacent figures are identical. Subjects have to discriminate whether the pair of identical figures (target pair) are two circles or squares. Adding a grouping factor by connectors produces two types of targets. Within-group pairs are targets linked by a connector, whereas between-group pairs are localized between two connectors. Subjects are classically faster for within-group than between-group pairs, which is accounted by the grouping by connector effect. Top-down expectations are also manipulated by varying the proportion of between-group pairs (25%, 50%, and 75% of between-group trials) relative to within-group ones. This leads to a variation of the grouping by connectors’ advantage size.

In order to assess the extent of the interhemispheric cost, we modified this original procedure. This time, figures are placed on a circle around a central fixation point. The target pair position is systematically manipulated: It can be either within the same hemifield (hemifield position) or across both hemifields (centred position). The stimuli stay on the screen until the subject gives his response. Fixation of the fixation point is verified by means of continuous eye-recording. The RT difference between both positions represents the interhemispheric cost.
EXPERIMENT 1

Subjects first performed the RDT when no grouping factor was provided (Figure 1A). Participants showed a significant interhemispheric cost of 90 ms.

EXPERIMENT 2

To evaluate if the interhemispheric cost observed in Experiment 1 could be erased by grouping factors like in brain-lesioned patients, we introduced connectors that linked figures by pairs. Hence, target pairs were within-group (linked objects) or between-group pairs (objects of two different pairs). We additionally manipulated the orientation of connectors. This allowed objects of different pairs (between-group) to be grouped by means of virtual links (Figure 1B).

We reproduced the significant interhemispheric cost’ result (87 ms) observed earlier. This cost was significantly reduced by 99 ms for within-group relative to between-group pairs. The classical advantage for within-group pairs relative to between-group pairs was reduced when connectors allowed objects of between-group pairs to be virtually related. The interhemispheric cost for between-group pairs remained constant, however (Figure 2). Regarding top-down modulations, subjects were faster to identify between-group pairs when their occurrence increased, meaning that

Figure 1. (A) Experiment 1: The target-pair is localized within the same (1) or across both hemifields (2). (B) Experiment 2: Virtual links between connectors are weakly (1), moderately (2), or strongly induced (3), increasingly facilitating the grouping of between-group objects. (C) Experiment 3: Global configuration is provided (1, 2) or not (3). Subjects have to gaze at the central fixation point and discriminate whether the pair of identical figures is two circles or squares (right or left response-keypress respectively). Figures stay displayed until subject’s response, but disappear as soon as subject stops gazing at the central fixation point.
between-group regions prioritization was at work. However, this did not allow subjects to reduce the interhemispheric cost when looking for a between-group target.

**EXPERIMENT 3**

As figures displayed in Experiment 2 always involved a global configuration, Experiment 3 involved three types of stimuli: (1) Concave connectors inducing a global configuration, (2) convex connectors inducing a global configuration, and (3) concave and convex connectors altogether to abolish the global configuration (Figure 1C). We addressed the question whether the interhemispheric cost would be even larger in this last condition.

Subjects were faster to identify a within-group relative to a between-group target pair when the global configuration was abolished, performed the other way round for the concave stimuli, and equally fast in both cases for convex stimuli. This shows that the manipulation of the global configuration was efficient.

Here again, a significant interhemispheric cost of 83 ms was observed. This cost was significantly reduced by 80 ms for within-group relative to between-group pairs. However, no other manipulation had any effect on this cost, neither the global configuration (Figure 3), nor the attention manipulation.

**DISCUSSION**

An interhemispheric cost was consistently observed in all three experiments. We suggest that the interhemispheric transfer cost might be larger than...
previously described. Interestingly, the effect of connectors reproduces results observed in brain-lesioned patients. Like in Brooks et al. (2005), only connectors, i.e., the presence of information on the midline, helped to reduce the interhemispheric cost. The lack of impact of other manipulations suggests that grouping effects mediated by attention or virtual links involves more complex mechanisms than the enhancement of direct links between neurons coding separate objects.

**REFERENCES**


Multisensory synchrony guides attention in dynamic cluttered environments

Erik van der Burg and Christian N. L. Olivers

*Vrije Universiteit, Amsterdam, The Netherlands*

Adelbert W. Bronkhorst

*Vrije Universiteit, Amsterdam, and TNO Human Factors, Soesterberg, The Netherlands*

Durk Talsma and Jan Theeuwes

*Vrije Universiteit, Amsterdam, The Netherlands*

Every day, we receive a bulk of information from different sensory inputs. This information often interacts when presented in close temporal or spatial proximity. For instance, nonspatial auditory signals enhance the perception of synchronized visual events (Vroomen & de Gelder, 2000). Furthermore, Spence and Driver (1997) have shown that responses to visual targets are faster when such a visual event is preceded by an auditory signal from the same location, compared to signals coming from a different location.

Most studies investigating interactions between different senses have typically explored single events, presenting one object at a time. Recently, van der Burg, Olivers, Bronkhorst, and Theeuwes (2008b) investigated how auditory signals affect competition between visual objects in more dynamic, cluttered environments. Figure 1A provides an example of the search display used in these studies (see also www.psy.vu.nl/pippop for a demonstration). Participants searched for a horizontal or vertical line among distractor lines (24, 36, or 48) of various orientations. At random intervals, a random number of items changed colour. On average once every 900 ms, the target changed colour, and it always did so alone. The important manipulation involved the presentation of a tone. In the tone absent condition, participants were instructed to search for the target, and to respond as fast as possible to its orientation. In the tone present condition, the
task remained the same, but the target colour change was accompanied by a tone.

The results are presented in Figure 1B. Although the target was the only item changing colour at its moment of change, finding it required strong attentional effort when the tone was absent. Search dramatically improved when the target colour change was accompanied by a tone. We dubbed this effect the “pip and pop” phenomenon, as the presentation of a tone (“a pip”) gives the impression of a visual pop-out. This effect occurred even though the pip was uninformative about the location or the identity of the target. The experiments also showed that the pip and pop effect is not due to temporal top-down knowledge, as visual temporal cues did not result in search benefits whatsoever. What we propose is that the auditory event is efficiently integrated with the visual event, boosting the saliency of the latter, which therefore attracts attention.

**IS IT EARLY AND AUTOMATIC?**

In a follow-up study, we investigated whether auditory-visual events guide attention in an automatic fashion (van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008a). Participants were asked to make a temporal order
judgement (TOJ) about which of two dots (left or right) appeared first. Such a task has been shown to be affected by attention (e.g., Shore, Spence, & Klein, 2001). Lateral to each of the dots, irrelevant distractors continuously changed colour. Prior to the presentation of the first dot, a tone carrying no relevant spatial information was synchronized with the colour change of one of these distractors, either on the same or on the opposite side of the first dot. The TOJs were affected by the location of the synchronized distractor, indicating that irrelevant auditory–visual synchrony guides attention in an automatic manner. Although we believe that the pip and pop effect has a strong automatic component, we have evidence that the effect is also susceptible to some top-down influences. For instance, it is less strong when participants adopt a small, more focused attentional window as compared to a more diffuse attentional window (van der Burg, Olivers, & Theeuwes, 2008e).

What are the underlying neural mechanisms behind the pip and pop effect? How early does the audio-visual integration occur? In an event-related potential (ERP) study (van der Burg, Olivers, Talsma, & Theeuwes, 2008d), diagonally oriented lines were presented in the left and right visual space, all continuously changing orientation. Participants searched for a line that changed to a horizontal or vertical orientation and made an unspeeded response to this target orientation. Accuracy was better when the target orientation change was accompanied by an auditory signal compared to when no sound was present, or the sound was synchronized with a distractor instead. We found evidence for relatively early (before 100 ms from tone-target onset) modulation of the ERP wave when a tone was synchronized with the target (AV condition), as compared to the summed A and V conditions (when only the auditory or only the visual signal was present). Later in the ERP wave, a lateralized ERP component (N2pc; Luck & Hillyard, 1994) that reflects the locus of attention in the visual field was observed when the target event was synchronized with the auditory signal. Interestingly, a similar N2pc component was observed when an irrelevant distractor was accompanied by an auditory signal, confirming our idea that, through rapid integration of the auditory and visual signals, audiovisual synchrony guides attention in cluttered continuously changing environments.

OTHER MODALITIES?

We have shown that auditory–visual synchrony guides attention in a stimulus-driven manner. So far, it remains unclear whether other modalities can affect spatial visual selection as well. In another study, we used a similar experimental set-up as in van der Burg, Olivers, Bronkhorst, and Theeuwes
(2008b), except that we replaced the auditory signal by a tactile signal. Search again dramatically improved when the target colour change was accompanied by a tactile signal as compared to when no such signal was present (“poke and pop”; van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008c). The result indicates that other modalities than audition can also affect spatial visual selection.

All in all, by using different methodologies (RTs, accuracy, TOJ, and ERPs), we clearly show that multisensory synchronization guides attention in visual multiple object displays. Whereas most studies have shown that spatial auditory signals affect spatial visual processing (e.g., Spence & Driver, 1997), the present experiments provide evidence that spatially nonspecific auditory or tactile events affect spatial visual processing as well, as long as they are synchronized. We suggest that the observed effects are due to multisensory integration, generating a salient visual event that captures attention.

REFERENCES


Selective attention of working memory contents by 11-month-old infants

Mariko Yamaguchi and Lisa Feigenson

*Johns Hopkins University, Baltimore, MD, USA*

Working memory (WM) allows us to mentally maintain a small number of items for storage and further processing. However, processing often requires selection of a subset of the total WM items for manipulation. Several studies (Garavan, 1998; Kessler & Meiran, 2006; Oberauer, 2003) suggest the existence of an internal focus of attention that permits just this type of selection. When asked to maintain multiple items in WM, adults are faster to update the same WM item twice in succession than to shift attention and update a different item. This suggests that adults must first select the appropriate item in WM before updating it.

These findings indicate an intriguing intersection between WM and attention. It appears that items can be selected from within WM, just as they are selected from an external array. However, it remains unknown whether: (1) The re-updating cost is specific to abstract numerical information and mental arithmetic, and (2) whether selection within WM is an explicit process that is exclusive to adults, or whether it is also observable in spontaneous early behaviour.

Here, we examined 11-month-old infants’ ability to switch their attention between items stored in WM. We relied on infants’ demonstrated ability to represent up to three hidden objects in WM (Feigenson & Carey, 2003; Wynn, 1992), then asked whether infants can also *re-update* their WM representations—that is, whether they can update an existing WM array that is no longer in the current focus of attention. Re-updating requires that WM representations remain flexible enough to be altered over time even when attention has shifted away to track objects in a different part of the scene.

Forty-eight 11-month-old infants (range: 11;0–11;30, average: 11;15; 26 males; n = 16 in each experiment) participated in a looking time experiment in which they saw displays of three objects serially hidden behind two
opaque screens. The screens were then lifted to reveal either the correct (three objects) or the incorrect (two objects) outcome (incorrect outcomes were created by surreptitiously removing one object from the side with two items), and infants’ total looking time to the display was measured.

In Experiment 1, infants saw a sequence of events that did not require re-updating the WM array. One object was hidden behind screen A on the left, then the remaining two objects were hidden behind screen B on the right (order and side counterbalanced). This sequence required no switches of attention to existing WM representations and thus no re-updating. Replicating previous work, infants succeeded in remembering the three objects, looking longer to the incorrect outcome of two objects (11.79 s), than at the correct outcome of 3 (7.64 s); paired two-tailed t-test, t(15) = 2.338, p < .05.

In Experiment 2, infants were required to re-update the WM array. Again, three objects were hidden behind two screens, but this time in alternating order. One object was first hidden behind screen A, then a second object was hidden behind screen B, and finally the third object was hidden behind screen A. Thus, infants had to shift their attention away from screen A in order to create a representation of the object hidden at screen B, and then had to return their attention to re-update their representation of what was behind screen A. Here, infants failed to look longer at the incorrect outcome of two objects (8.91 s) than at the correct outcome of three (10.85 s), t(15) = −1.110, ns, suggesting that they had failed to accurately re-update their WM representation.

Figure 1. Infants’ looking times to expected and unexpected outcomes. In Experiments 1 and 3 where infants were not required to re-update, they successfully represented three objects in working memory. In Experiment 2, where re-updating was required, infants failed to represent all three objects.
Was infants’ failure in Experiment 2 due to the simple act of alternating their eye gaze, or was it instead due to the attempt to re-update a WM representation? Experiment 3 was a control condition in which the two screens were connected by a horizontal bar to create a single, dumbbell-shaped screen. In previous work, two connected objects have often been represented as one (for review, see Scholl, 2001). Our aim was to replicate the required eye movements of Experiment 2, but to eliminate the requirement to re-update by encouraging infants to represent all of the objects as belonging to the same array. The spatiotemporal features of the presentation were exactly as in Experiment 2, in which infants failed to remember the three objects. Infants saw one object hidden behind the left side of the screen, then one object hidden behind the right side, and finally one object hidden behind the left side. Despite having identical hiding motions to those in Experiment 2, infants here successfully looked longer at the incorrect outcome of two objects (11.74 s) than at the correct outcome of three objects (7.73 s), $t(15) = 2.377$, $p < .05$. In summary, similar to adults, infants show a cost to shifting attention within WM.

Was infants’ difficulty in re-updating limited to mentally adding an object to an array, or does it also extend to other types of transformations? Experiment 4 explored a subtraction re-update (two objects behind screen A, one object behind screen B, then one object taken away from behind screen A). Paralleling the results of Experiment 2, 11-month-olds looked equally at the correct (9.44 s) and incorrect (8.58 s) outcomes, $t(15) = -1.145$, $ns$. This failure did not arise from a difficulty with subtraction events in general; infants succeeded in Experiment 5, a control subtraction event that required no re-updating. Infants looked longer to the incorrect outcome (8.79 s) than to the correct outcome (6.35 s), $t(15) = 2.757$, $p < .05$. The data demonstrate that re-updating remains difficult for infants, regardless of the type of update required.

These results reveal a striking similarity between infant and adult WM. An internal focus of attention exists early in development and allows the selection of relevant items from the contents of WM. Despite this developmental continuity, some developmental change is suggested. Adults can successfully re-update WM representations, but at a cost to reaction time (Garavan, 1998; Kessler & Meiran, 2006; Oberauer, 2003); in contrast, here infants failed entirely to re-update. Future studies should explore ways in which selective attention within WM may change in more subtle ways over development.
REFERENCES


