Change detection in an attended face depends on the expectation of the observer

Erin L. Austen   Psychology, University of British Columbia, Vancouver, Canada

James T. Enns   Psychology, University of British Columbia, Vancouver, Canada

Sensitivity to a scene change during a brief interruption depends critically on a match between what the observer expects to see and the kind of change that occurs (Austen & Enns, 2000). The present study tested the generality of this conclusion using human faces, which are both socially more relevant and perceptually more configural than the compound letters tested previously. An experiment using the flicker technique examined sensitivity to two types of change: facial identity and emotional expression. Change detection was assessed when attention was either focused or distributed, the change was either expected or unexpected, and the faces were either upright or inverted. The main finding was that detection was expectation-dependent, even when only a single upright face was presented. Secondary findings with regard to attentional distribution and face inversion confirmed that observers were indeed engaged in face processing. We conclude that observer expectations critically influence the perception of single and fully attended human faces.

Keywords: change detection, observer-expectation, perception, visual representation, focused attention, divided attention, visual search, human face processing

Introduction

Perception is not uniformly detailed over the visual field for several reasons: cones are distributed to maximize spatial resolution near the retinal fovea, a disproportionate number of cortical neurons are devoted to the center of gaze, and multiple objects in simultaneous view cannot all be attended uniformly. These anatomical and cognitive considerations combine to place severe limitations on what can be seen in a glance.

The limits on perception are well illustrated by 'change-blindness' (Rensink, 2002). Briefly interrupting the view of a scene by an eye blink, a brief flicker in the image, or a change in viewing position, renders the viewer profoundly insensitive to changes in the location and identity of objects that are not at the current focus of attention. This has prompted researchers to try to quantify the number of objects that can be seen in a glance, by measuring the speed and/or accuracy in the report of object attributes. Display presentation is usually limited to a brief period that confines eye fixations to a single location. Care is also taken to ensure that the visual items cannot be verbally rehearsed or perceptually grouped during the period that intervenes between the original scene and the test. Such studies point to a four-item upper limit on the short-term visual memory for scene contents at a glance (Rensink, 2002; Sperling, 1960; Vogel, Woodman, & Luck, 2001).

But this does not answer the question of whether the representation of all four of these objects is equally detailed. To address this question, researchers have focused on whether attention can be devoted equally well to one versus two objects. This research clearly indicates that even two objects are not represented as richly in a glance as a single object. For example, in one study observers made a speeded decision about the spatial relations among the tips of two arrowheads (< vs. >) (Baylis and Driver, 1993). When the two arrowheads were perceived as belonging to the same object (a central hexagon) this decision was made more efficiently than when they were seen as belonging to different objects (two K-shapes flanking a central hexagonal background region). Similar 'two-object costs' in perception have been documented using a wide variety of perceptual tasks and displays (Baylis, 1994; Davis et al., 2000; Duncan, 1984; Driver & Baylis, 1995).

Based on this research, one might be tempted to conclude that the visual representation of a single object is rich and detailed (Duncan, 1984). Yet there are numerous hints that even the perception of an object in isolation — one that is fully attended — does not involve a completely detailed representation. Admittedly, this may be hard to accept because it runs counter to our subjective experience of what it means to see. Yet, in a generally stable world, there may be no need for a detailed representation to be constructed in the mind (Rensink, 2000). Instead, sensory information can be consulted on a 'need-to-know' basis. The less-than-

pictorial schemas of our mind may work simply because they permit us to link appropriately to a visual field location or object when the need arises, giving rise to an illusion of detail.

This point was made rather elegantly over 30 years ago in a demonstration involving a variant of the Necker cube shown in Figure 1 (Hochberg, 1968). This drawing shows a wire cube with one solid side on the right, as implied by the fact that the wire portion of the cube is occluded on that side. Most observers report that when they fix their gaze on the corner labeled '1' the cube is seen as though it is being viewed from below. What is surprising is that when the same viewers fixate the corner labeled '2,' the cube appears as though it is being viewed from above. These results are the same as those obtained with the original version of the Necker cube, despite the fact that in this altered version the solid side is inconsistent with a viewpoint from above. This indicates that even for a prolonged view of a single, unambiguous object at the center of gaze, perception is not uniformly detailed.
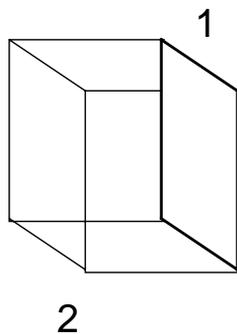
Figure 1. This modified Necker cube changes its perceived orientation depending on whether the eyes are fixated on corner '1' or '2'. The fact that perception varies for this unambiguous object suggests that even the perception of a single attended object is not as rich in visual detail as we like to think.

This point has also been made more recently in studies of change detection involving scenes of single actors in real world interactions (Simons & Levin, 1998) and in movies (Levin & Simons, 1997). Observers who are fixating these single actors, and attending to them, still often fail to notice an identity switch to the actor during a brief viewing interruption. Both the older Necker cube demonstration, and these more recent change blindness results indicate that neither single spatial locations nor single objects appear to be the basic unit of visual representations. If they were, we would expect observers of the Necker cube in Figure 1 to detect the inconsistencies in their perceptions, and we would expect changes to a single object, one that is both foveated and attended, to be detected reliably.

If the basic unit of perception is not the single object, what is it? In an effort to address this question, we recently reported results of a change detection task involving compound letters (Austen & Enns, 2000). These are stimuli that consist of two independent levels of structure. At the 'local' level of detail are small letters that together form the shape of a larger letter at a 'global' level of detail. An example of a compound letter is shown in Figure 2. We chose these stimuli in order to disentangle the spatial distribution of attention — is attention focused or distributed over the field of view? — from both the detail level of the stimuli — is the target at the local or global level? — and the expectations of the observer — is a target expected at the local or global level?.
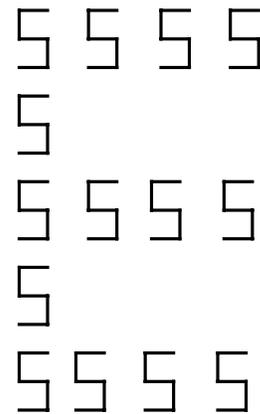
Figure 2. The compound letters used in the change detection study of Austen & Enns (2000).

We employed the flicker method of studying change detection (Rensink, 2000). The observer's task was to indicate whether any one of the items was changing from frame to frame. The to-be-detected change could involve either the global letter or the local letters. In one half of the display sequences, a single compound letter changed at one of the levels between frames, and in the other half of the displays there was no change. Fixing the overall proportion of change and no-change trials at 50% meant that, regardless of the expectation of the observer, or the likelihood of a change occurring at either level, the response biases of the observers were controlled.

In the first experiment of Austen and Enns (2000), observers were informed that when a change was present, it was equally likely to be at either level. The results showed that when attention could be focused on a single compound letter, changes at the local and global levels of structure were equally detectable. However, when attention had to be distributed among three or five compound letters to determine whether one was changing, then changes at the global level were detected more readily than changes at the local level. This pair of findings was thus consistent with the idea that single attended objects are richly represented following a brief

glance, and that there is a bias favoring the global or 'gist' level of structure when attention is distributed among multiple objects (Navon, 1977).

This interpretation had to be modified, however, by the results of a second experiment in which expectations were varied systematically. In a global bias condition, 75% of the change trials involved a change to one of the global letters while the remaining 25% of those trials involved letter changes at the local level. A local bias condition involved the complementary arrangement of probabilities (25% changes were to global letters and 75% were to local letters). Under these conditions, observers were both faster and more accurate to detect changes at the level that was most likely to change. This was true even when there was only a single item to monitor for change and despite the fact that the expectation bias was not linked to a response bias (there was still an equal likelihood of change versus no change, as in Experiment 1). This strongly suggests that the limiting factor on attention is not the number of items to be examined, but rather the detail level within an object that is consistent with the current expectation of the observer.

But how general is this conclusion? Is it limited by the particular stimuli that have been tested so far? For instance, some have criticized Hochberg's (1968) Necker cube because it is an impoverished and artificial stimulus. The lines on the page must be interpreted as a three-dimensional object, with some of these lines representing edges of an unseen surface and others representing wires. Others have criticized the real world interaction and movie experiments involving single actors because the observer's perception of these critical figures is so uncontrolled. There is no independent way to verify where the attention or the eyes of the observers were actually focused in those studies. Finally, even the compound letters used by Austen and Enns (2000) can be called into question. For one, compound letters have little ecological or social significance as 'objects.' If anything, they are among the most arbitrary and overlearned symbols that can be tested. Also, unlike most natural objects, the levels of a compound letter are independent of one another, meaning that it is possible to change the global level with little effect on the local level, and vice versa.

Rationale for Testing Change Detection in Faces

Our aim was to test the hypothesis of expectation-dependent perception using human faces. Faces are ideal for several reasons. For one, they are a class of objects with unique social and biological significance for humans. They are among the earliest objects humans learn to read (for emotional expression) and to recognize (for identity). Humans are 'experts' at face processing, both in the sense that they are able to rapidly assign meaning to hundreds

of closely similar stimuli (faces) and in the sense that this is done with little if any conscious awareness of the underlying factors involved. Neuropsychological conditions (Farah, Levinson, Klein, 1995), behavioral data (Ro, Russell & Lavie, 2001; Tanaka & Farah, 1993), and brain imaging evidence (Haxby, Hoffman, & Gobbini, 2000; Kanwisher, McDermott & Chun, 1997) all support the idea that faces form a coherent and distinct class of objects with special relevance for humans.

Second, human faces are hierarchical in their structure, in that they are comprised of 'parts' such as eyes and mouth and 'spatial relations' among parts such as eye-mouth and eye-eye distance. But, unlike compound letters, the levels of structure in a face are completely interdependent, in that changing the specific parts of a face will also change its identity (Tanaka & Sengco, 1997). The face stimulus forms a tightly knit package of features in which almost every nuance has an influence on the perception of other aspects of the face (Perrett, Benson, Hietanen, Oram & Dittrich, 1995).

Third, human faces are processed for multiple sources of information. At a first approximation, these sources can be characterized as the three-dimensional physical structure of the face, the analysis of familiar identity, and the emotional expression of the face. Although theoretical treatments differ in assigning these three functions to either a hierarchical scheme (Zeki, 1999) or to independent parallel processes (Bruce & Young, 1986; Young, 1998), all agree that very different kinds of information are analyzed when a face is evaluated for identity versus expression. Neuropsychological studies show that patients can be left with severe deficits in one function and yet show relatively intact performance in the other (Young, Newcombe, de Haan, Small & Hay, 1993).

The Design of the Present Study

The present experiment followed the same design as Austen and Enns (2000), with the exception that the compound letters were replaced by the faces of two individuals (person 1 and 2) posing with one of two emotional expressions (happy and sad). We expected face changes to be detected generally more efficiently than letter change detection because of the special status afforded to faces in human perception. Among the two kinds of face changes we implemented, we expected expression changes to be detected more readily than identity changes simply because the happy versus sad discrimination could be based on a salient visual feature of the face (mouth shape) whereas detecting an identity change required a more subtle analysis of relations among features (Suzuki & Cavanagh, 1995). Having a different baseline of change detection for the two features also allowed us to test whether expectations would play as strong a role when the two features were not balanced.

Perhaps change detection is expectation-dependent only when the two kinds of change are roughly equal to begin with.

There were three main questions of interest. First, we tested the detection of face changes under both focused and divided attention, allowing us to determine whether the detectability of the two change types varied with attentional focus. We reasoned that if face identity requires configural processing which is more complex than the featural processing associated with face expression, visual search for identity change should become increasingly more difficult with increases in display size than would search for expression change (Wolfe, 1996).

Second, we tested change detection in faces that were either upright or inverted. This is an additional way to confirm that the pictures were being processed as faces and not merely as arbitrary stimuli in which some pixels could differ from frame to frame. When faces are upside down it is especially difficult to determine the identity of an individual, largely because the configuration of the face is orientation dependent (Carey & Diamond, 1977; Murray, Yong, & Rhodes, 2000; Yin, 1969). Thus, detecting a change to facial identity in an inverted face is likely to be difficult. In contrast, one might expect an expression change in an upside down face to be detected more easily, since attention would need only to be focused on a single feature such as mouth curvature or eye shape. We therefore predicted that identity change detection should be disproportionately more difficult in upside down faces, while there should be little to no effect of inversion on the detection of expression changes. Incidentally, inverting the faces also controls for low-level stimulus differences between the stimuli. If observers are relying on average luminance or contrast differences between the faces to detect a change, then we should expect the same pattern of data when the faces are inverted, since inversion has no effect on average image luminance or contrast.

Third, and most importantly, we tested change detection under three biasing conditions: neutral (identity and expression changes were equally likely), identity (an identity change occurred on 75% of the change trials), and expression (an expression change occurred on 75% of the change trials). We kept the overall proportion of change to no-change trials constant at 50% in each condition to prevent any response biases from influencing the results. We reasoned that if observer expectations were an important factor in face processing, then detection would be best when changes occurred at the expected level. If, on the other hand, all detail levels were attended and represented simultaneously because faces are richly represented in a glance, then we should find no effect of biasing to one aspect of the face over another.

Methods

Participants

One hundred and fifty undergraduates from the University of British Columbia participated in a 1-hr session in return for partial course credit. Participants were randomly assigned to one of six conditions formed from 3 Biases (Expression, Neutral, Identity) x 2 Orientations (Upright, Inverted). All participants reported normal or corrected-to-normal vision.

Stimuli and Apparatus

Displays were controlled by a Macintosh computer and presented on a monitor set to 256 levels of gray. A chin rest was used to maintain a viewing distance of 57 cm. Photographs of the faces of two females (footnote 1), each posing for two separate expression shots (one happy, one sad), were digitally altered so that each face was the same oval shape (2.2 x 3.0 degrees of visual angle). The photos were cropped to remove any information conveyed by hair and accessories, and were presented on a medium gray background. The set of four female faces therefore allowed for all combinations of identity (person 1 versus 2) and emotional expression (happy versus sad). An example each possible change type used in the experiment is shown in Figure 3.

Displays consisted of alternating frames of 1, 3 or 5 randomly chosen faces for 200 ms, followed by a blank frame of 200 ms, and then followed again by faces in the same locations for 200 ms. This sequence continued until observers pressed one of two keys, indicating that a change had been detected in one of the faces. On one half of the trials, a change was present, such that a different face in Frame B replaced one of the faces in Frame A. These two frames continued to alternate until a response was made. Note that since the same two frames were alternated on any given trial, the evidence for 'change' was present until the observer responded.

The change, when it occurred, could involve either the identity or the expression of any face in the display. Observers in the Neutral Bias condition were informed that these changes were equally likely, while observers in the Identity Bias condition were informed that 75% of the changes would involve identity, and the remaining 25% of changes would be to expression. Observers in the Expression Bias condition were informed of the reciprocal probabilities.

Feedback was presented in the form of a plus (correct) or minus (incorrect) sign at the center of the screen following each response. This also served as the fixation and warning symbol for the start of the next trial. Observers were given a time window of 13 s in which to make a response. If none was made, a timeout symbol
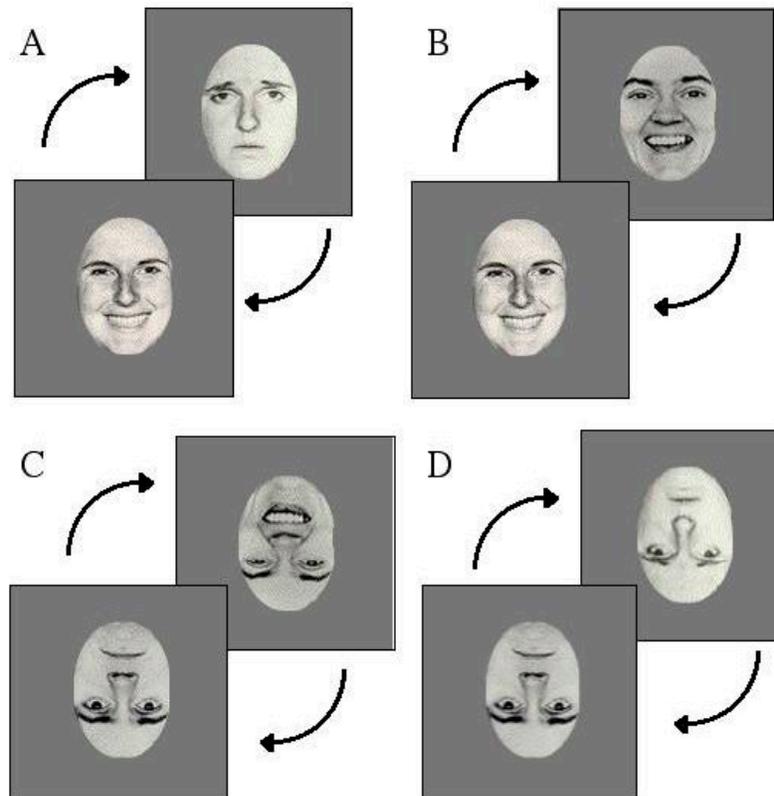
Figure 3. Examples of the four possible change types: (A) emotional change in upright face, (B) identity change in upright face, (C) emotional change in inverted face, (D) identity change in inverted face.

appeared at the center of the screen, and was followed by a new trial.

Faces appeared randomly in one of nine locations of an imaginary 3 x 3 matrix (18 x 19.2 degrees overall, each cell measured 6 x 6.4 degrees). Face locations were jittered randomly, with the constraint that a minimum distance of 1 degree separated the faces.

Procedure

Participants indicated whether a change was present by pressing a designated key with an index finger as rapidly and accurately as possible. If no change was detected they pressed a different key with the other index finger. Participants were told that a change was present in one of the faces on 1/2 of the trials and that the three display sizes were randomly intermixed in a block of trials. Participants were given printed and verbal instructions, before beginning a practice block of 10 trials. A testing session consisted of eight blocks of 60 trials. At the end of each block, a dialogue box on the screen indicated the error rate, and a warning message was presented if errors exceeded 10%. Participants were instructed to slow down on the next block if this warning message was presented. Response time (RT) was measured in milliseconds (ms).

Results

It was necessary to conduct several preliminary analyses, to confirm our assumptions about the way these stimuli were processed, before turning to the analyses of primary interest involving the role of expectations on change detection. A first analysis compared letter change detection from the Austen and Enns (2000) study with the present face change detection results. Overall, responses were found to be faster (by 700 ms) and more accurate (by 4%) for face change than for letter change detection (RT: $F(1, 174) = 106.66$, $p < .01$; accuracy: $F(1, 174) = 15.56$, $p < .01$).

The remaining analyses were conducted on the face change data in the present study. These data were subjected to analyses of variance involving the within-subjects factors of Change Type (None, Identity, Expression) and Display Size (1, 3, 5), and the between-groups factors of Bias (Neutral, Identity, Expression), and Orientation (Upright, Inverted). The dependent variables were correct RT and accuracy. Because these two measures revealed the same patterns (footnote 2), they were combined for presentation purposes in the form of inverse efficiency scores (Townsend & Ashby, 1983). This involves forming a ratio of RT over proportion correct, for each observer and condition. It is a compact
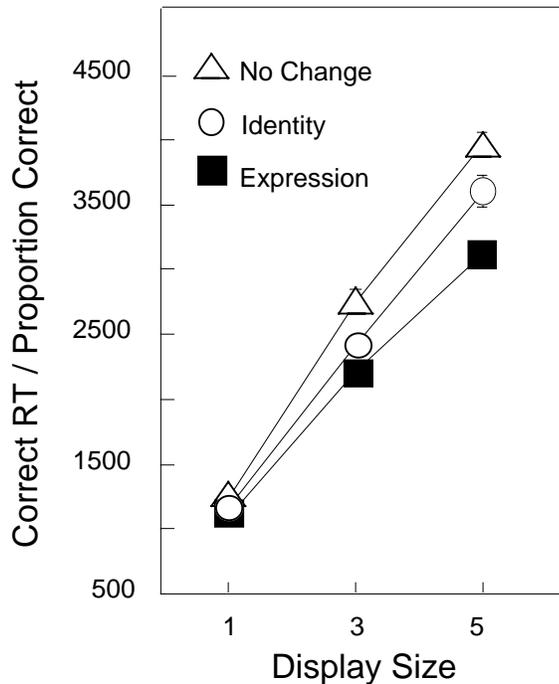
Figure 4. Search efficiency for each of the three change types (identity expression and no-change) across display size. Most of the standard error bars are smaller than the data symbols.

and easily interpretable index of performance, whose only assumption is that there is a linear relationship between correct response time and errors (footnote 3). Efficiency scores are especially useful when error rates are variable across conditions. They are interpreted in the same way as correct RT, being in fact identical when accuracy is perfect, and growing proportionately with increases in errors.

Change Detection and the Distribution of Attention

One preliminary analysis examined the influence of display size on face change detection, testing the assumptions that change detection was less efficient as the number of faces was increased and that expression change was more readily detectable in these stimuli than identity change. The efficiency of detecting identity change, expression change, and no change in upright faces is shown in Figure 4 as a function of display size, averaged over all three conditions of bias. Overall detection efficiency was better for expression changes (1087) than for identity changes (1128) when the display size was one, $F(1, 76) = 11.69$, $p < .01$. As display size increased, identity change RT increased linearly, as would be expected when searching for targets that do not have 'pop out' features (average $R^2 = .998$). As expected, search rates were also less efficient for identity change (612 ms for each additional item) than expression change (506

ms/item), and this was reflected in a significant effect of Change Type on the slope of the efficiency scores, $F(1, 76) = 18.18$, $p < .001$.

Change Detection in Upside Down Faces

A second preliminary analysis examined the effects of face inversion, to test the assumption that identity processing was more dependent on configural processing than expression processing. A significant Change Type x Orientation interaction, $F(1, 144) = 5.59$, $p < .02$, shown in Figure 5, revealed that turning the faces upside down had the predicted effect of increasing search difficulty for identity change and leaving search for expression change comparably easy in both orientations. The Orientation x Bias interaction was not significant, $F < 1$.
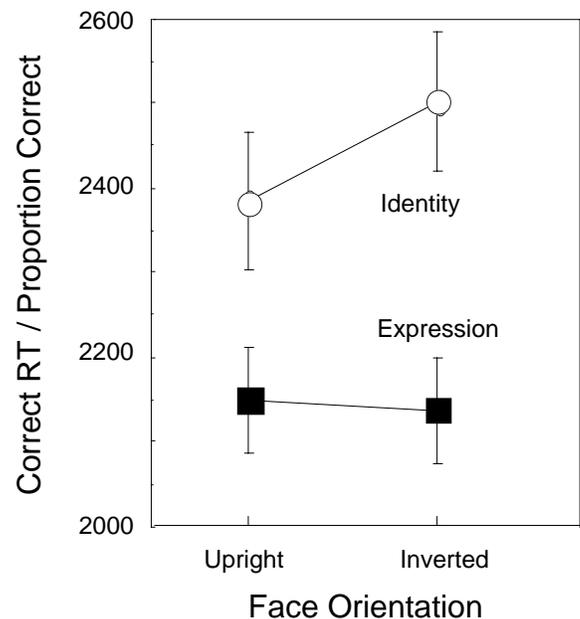

Figure 5. Search efficiency for the identity and expression changes across orientation. Error bars depict 1 SE.

Change Detection and Bias: Focused Attention

Our primary interest was in testing change detection for an expected (75% likely) versus unexpected (25% likely) type of change while the overall likelihood of any change remained constant at 50%. We made these comparisons separately for focused attention (display size = 1) and distributed attention (display sizes 3 and 5) because of our primary interest in the perceptual representations of fully attended objects.

Mean efficiency scores for detecting change in a single face is shown in Figure 6 for the three bias conditions, separately for each orientation. Detection of expression

change was generally more efficient than detection of identity change, $F(1, 147) = 48.34$, $p < .001$, but this main effect was tempered by a significant interaction between change type and expectation, $F(2, 147) = 11.84$, $p < .001$. We examined this interaction more closely with simple effects, comparing the detection of identity change across the three biasing conditions, and then the detection of expression change across the same conditions. The detection of each type of change was most efficient when observers were biased to detect it (expression change in the expression vs. neutral bias, $F(1, 147) = 9.85$, $p < .01$, and identity change in the identity vs. neutral bias, $F(1, 147) = 10.22$, $p < .01$). Detection of the unexpected change type within each of the two biasing conditions did not differ from its detection in the neutral condition (both $Fs < 1$). Thus, even when attention was focused on a single face, the detection of change was dependent on the expectations of the observers.

Detection of expression changes were generally more efficient than the detection of identity changes, $F(1, 147) = 101.88$, $p < .001$, and this main effect was again tempered by a significant interaction between change type and expectation, $F(2, 147) = 33.06$, $p < .001$. Simple effects tests confirmed that biasing observers to attend to a particular change type improved the efficiency of its detection relative to the neutral condition (expression change, $F(1, 147) = 10.56$, $p < .01$, and identity change, $F(1, 147) = 49.07$, $p < .001$). Interestingly, in the identity biasing condition, expecting to see changes in identity was not only a benefit to the detection of identity change, but it also benefited expression change as well, $F(1, 147) = 14.60$, $p < .01$. As described in the previous section, this did not occur when attention was focused. Another difference from those results was that expecting to see a change in expression not only benefited expression changes, but it impaired the detection of the unexpected identity changes, $F(1, 147) = 22.17$, $p < .01$.
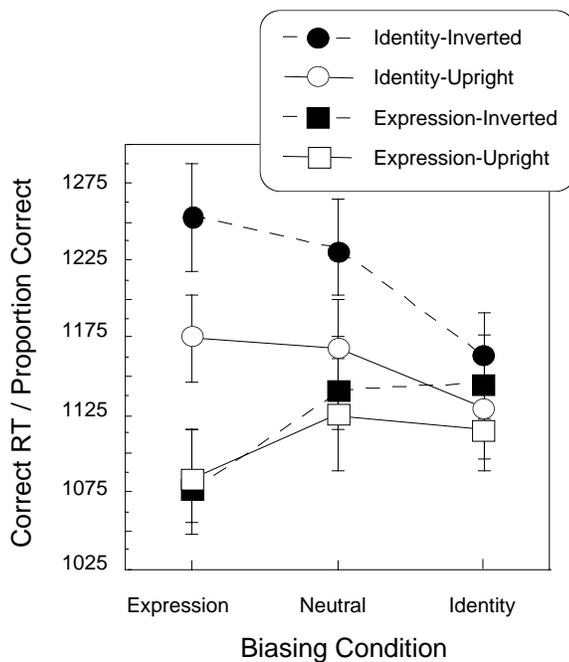

Figure 6. Search efficiency for identity and expression changes in the focused spatial attention condition. Error bars depict 1 SE.

Change detection and Bias: Distributed attention

Mean search efficiency for the distributed attention conditions (display size = 3 and 5) is shown in Figure 7. These results provided an important context for the focused attention results. We were interested to know, for example, whether facial identity, like the global level of compound letters, enjoys a global-precedence effect when attention is distributed or whether the feature of emotional expression guides attention more effectively.
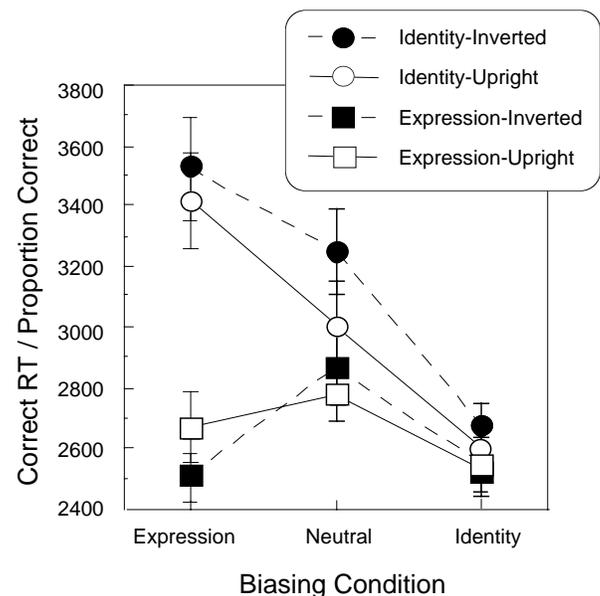

Figure 7. Search efficiency for identity and expression changes in the distributed spatial attention condition. Error bars depict 1 SE.

Discussion

We began this study by asking whether an attended face seen in a glance is richly represented. Are all of the attributes of a face available for report, once attention has been focused on it, or is the representation of a face dependent on the expectations of the observer? This question was prompted by a recent study in which the detection of change in a single compound letter was found to be highly dependent on the expectations of the observer about what kind of changes were likely (Austen & Enns, 2000).

We tested the generality of this interpretation with the detection of change in human faces for several reasons. First, humans are experts at making the subtle visual discriminations required to identify a face. Second, faces are treated as special objects in the sense that there are regions of the brain devoted to their processing, as indicated by neuropsychology and brain imaging. Third, unlike most objects, faces are defined by specific configurations and relational properties. Thus, if face perception showed the same expectation-dependence as the perception of compound letters, we would be able to conclude that even their perception was not fully detailed.

It was necessary to first conduct several preliminary analyses in order to establish a context in which the main results could be properly evaluated. These included an analysis comparing our previous letter detection results (Austen & Enns, 2000) with the results of the present face detection task. It revealed that face change detection was indeed more efficient than letter change detection. They also included an analysis of the visual search slopes for the face detection task. It revealed that although face detection may have been easier than letter detection, it nonetheless still became increasingly inefficient as the number of faces in the display increased. This indicated that our attempt to vary the distribution of attention across faces was successful.

Another important aspect of the search slope analysis was that increases in display size had a larger influence on the detection of changes in identity than on changes in expression. This is consistent with expression being coded as a simpler and more distinctive feature (e.g., mouth curvature) than identity, which is likely coded as a more complex configuration of features for which spatial relations are important. Finally, the finding that inverting the faces impaired identity change detection while leaving expression change detection unaffected supported this interpretation independently. These findings converge on the conclusion that these stimuli were being processed as human faces (Murray, Yong, & Rhodes, 2000).

The most important and novel result of this study was the influence of observer expectations on face change detection. Observers detected an expected change in a face more rapidly and accurately than an unexpected change in the same face. This was observed not only when attention was distributed across a number of faces, as would be expected by almost all theories of perception, but it was seen when observers were monitoring for a change in a single, fully attended and foveated face. Not only that, but observer expectations influenced change detection to a similar degree for features that differed in their baseline level of change detectability. This latter finding is an important contribution, since it rules out an interpretation premised on the more discriminable features of change simply being detected more readily.

This finding therefore generalizes our previous interpretation of the compound letter results (Austen & Enns, 2000) to objects — human faces — that are of biological and social significance to observers. It also generalizes it to a class of objects that are more likely to be represented in the visual brain as integral configurations rather than as patterns with separable elements.

The Role of Expectation in Perception

The idea that expectations play an important role in perception has a long history (James, 1890). In the current literature one can point to at least four distinct paradigms that illustrate this point in different ways, including covert orienting (visual targets that appear at expected locations are processed more efficiently than those that appear at unexpected locations, even when retinal location and response priming are controlled, Jonides, 1981; Downing, 1988); contingent capture (distractor objects are processed involuntarily as a direct function of the degree to which they share visual features that are relevant to the task at hand, Folk, Remington, & Johnston, 1992); change blindness (the detection of a change to a scene occurs more rapidly and reliably if the change is one that is anticipated, Rensink, 2002); and inattentional blindness (objects that appear at the center of gaze can go completely unnoticed if attention is concurrently being directed to another stimulus that is also in view, Mack & Rock, 1998).

What is the contribution of the present data, when viewed against this long legacy of research on the role of expectation in perception? A first point is that the main finding in the present study did not involve a misdirection of visual attention in space. In each of the previous paradigms, unexpected objects are not processed as efficiently, in large part, because they are presented to locations in the visual field that are not fully attended. Covert orienting involves the spatial misdirection of attention, contingent capture involves active ignoring of stimuli at a given location or time, change blindness typically involves attention distributed widely over a scene, and inattentional blindness also involves spatial and featural misdirection of attention.

A second point is that in the present study visual attention could be devoted entirely to a single object. Covert orienting, contingent capture, change blindness and inattentional blindness have all depended on multiple objects for their effects. In many cases, attending to multiple objects means the same thing as attending to multiple locations in space. However, as the literature on object-based attention shows, dividing attention between objects results in performance deficits even when the number of spatial locations is held constant (Davis et al., 2000).

The present results show clearly that even when attention is devoted to a single object, that is, to a familiar face presented in a very compact region of the visual field against an otherwise blank screen, it still does not

guarantee a visual representation in which all aspects of the object are uniformly available to consciousness. What is noticed first about a face depends strongly on what observers expect to see, even in these most minimal of visual settings. This suggests that attention to a specific set of features changes the spatial filtering of the stimulus in much the same way that attending to a location alters the spatial filtering of stimuli at that location (Yeshurun & Carrasco, 1998; 2000).

Future Directions

Although the present findings indicate that the identity and expression of an attended face are not simultaneously available in perception, they do not provide much guidance as to what information is being used to evaluate change in each feature. One promising approach to this question is the 'bubbles' technique (Schyns, Bonnar & Gosselin, 2002). This is a procedure in which various regions of a face are presented to an observer, each containing a range of spatial frequencies, in an effort to determine the most diagnostic aspects of a stimulus for any given task.

In the Schyns et al (2002) study, for example, observers either identified faces, discriminated their gender, or evaluated their emotional expression. The main findings included that the optimal spatial frequency for face identification was in the range of 12-22 cycles per face, with the most important spatial regions including the eyes, nose and mouth. In comparison, the optimal scale for the expression task was shifted toward the lower spatial frequencies (6-12 cycles per second) and the critical region was centered more exclusively on the mouth. This trend toward lower spatial frequency information in facial expression tasks is consistent both with other studies of face perception (Morrison & Schyns, 2001) and with the present finding of generally more efficient change detection for expression than identity. The application of selective filtering or the 'bubbles' technique in future studies has the potential to reveal which information is being used selectively when change of a certain kind is detected in a face.

An unexpected finding that is deserving of further study is the asymmetry in the effects of bias to each change type when attention was distributed. Biasing for identity change improved detection of both identity and expression changes. In contrast, biasing for expression change improved its detection, but it also had a large negative effect on detection of identity change. This suggests that the attentional setting best suited for identity processing also benefits expression processing, and does so most strongly when attention is distributed. One way this might come about is that processing the configural properties of the face results in automatic benefits for the processing of any specific features that are part of it. Another possibility is that identity processing requires a wider spatial focus of attention for each face,

thereby benefiting detection of incidental expression changes (Schyns et al, 2002).

What makes this asymmetry so interesting is that it is opposite to the interactions between levels in compound letters as reported by Austen and Enns (2000). That study found a global precedence effect, where attention to the large letter configuration could be achieved with little or no interference from letters at the local level. At the same time, identification of the local letters was affected by the identity of the global letter. In the present change detection task, it is tempting to link the small letters to the emotional expression in faces (both local features) and the large letters to the identity of the face (a global configuration). If so, then similar asymmetric processing relations between levels would predict that face identity would interfere more with expression detection rather than the other way around. If anything, the pattern obtained was opposite to this prediction, in that the identity bias (global) benefited expression detection (local) more than the reverse. However, given this very limited set of data on change detection (involving only two different faces and two extreme emotional expressions), we want to urge caution in any interpretations regarding the more general issues involved in processing facial emotion and identity (Young, 1998). This study was designed to use faces as a unique tool; it was not intended as a thorough study of face perception. Yet, the pattern is intriguing and warrants further study. At a minimum, it may point to important differences between the processing of faces and other objects.

Conclusion

The present findings, along with the previous results of Austen and Enns (2000), indicate that even isolated and fully attended objects are not represented in the visual system with uniformly rich detail. Visual perception is selective, not only with respect to a limited region of the visual field, and with respect to a limited number of objects, but also with respect to a limited range of all the visual attributes that together comprise the 'object.' To reiterate Julian Hochberg (1968) in his discussion of the Necker cube shown in Figure 1, the perception of even single foveated objects is 'not everywhere dense.' The present study extends this insight by showing that the non-uniformity in the perceptual details of an object can be predicted by the expectation of the observer.

Footnotes

1. The photographs used were digitally manipulated versions of images used by Ekman and Friesen (1976). Photos used with permission.
2. The patterns of significance obtained for the efficiency scores mirrored the analyses of correct RT and

accuracy in all the important respects. Only three discrepancies were observed, all because of ceiling effects in accuracy when display size =1. The accuracy of identity versus expression change did not differ significantly for upright faces, F (1, 76) = 1.04 (Figure 4), nor was the main effect of Change Type or the Change Type x Bias interaction significant, Fs < 1 (Figure 6).

3. The correlation between correct RTs and errors was r(148) = .329. This supports the assumption of linearity underlying the use of inverse efficiency scores (Townsend & Ashby, 1983).

Acknowledgements

References

Austen, E. L., Enns, J. T. (2000). Change detection: Paying attention to detail. *Psyche: An Interdisciplinary Journal of Research on Consciousness, 6 (11).*

Baylis, G. C. (1994). Visual attention and objects: Two-object cost with equal convexity. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 208-212.

Baylis, G. C., & Driver, J. (1993). Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 451-470. [PubMed]

Bruce, V., & Young A. (1986). Understanding face recognition. *British Journal of Psychology, 77,* 305-327 [PubMed].

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science, 195,* 312-314. [PubMed]

Davis, G., Driver, J., Pavani, F, & Shepherd, A. (2000). Reappraising the apparent costs of attending to two separate visual objects. *Vision Research, 40,* 1323-1332. [PubMed]

Downing, C. J. (1988). Expectancy and visual-spatial attention: Effects on perceptual quality. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 188-202. [PubMed]

Driver, J., & Baylis, G. C. (1995). One-sided edge assignment in vision: 2. Part decomposition, shape description, and attention to objects. *Current Directions in Psychological Science, 4,* 201-206.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113* (4), 501-517. [PubMed]

Ekman, P. & Friesen. W. V. (1976). *Pictures of Facial Affect.* Palo Alto. CA: Consulting Psychologists Press.

Farah, M. J., Levinson, K. L., Klein, K. L. (1995). Face perception and within-category discrimination in prosopagnosia. *Neuropsychologia, 33,* 661-671. [PubMed]

Folk, C. L., Remington, R. W. & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 1030-1044. [PubMed]

Haxby, J. V., & Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4,* 223-233. [PubMed]

Hochberg, J. (1968). In the mind's eye. In R. N. Haber (Ed.). *Contemporary theory and research in visual perception* (pp. 309-331). New York: Holt, Rinehart & Winston.

James, W. (1890). *The principles of psychology.* New York: Holt, Rinehart & Winston.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. B. Long & A. D. Baddely (Eds.), *Attention & Performance Vol. 9* (pp. 187-203). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17,* 4302-4311. [PubMed]

Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin and Review, 4,* 501-506.

Mack, A. & Rock, I. (1998). *Inattentional blindness.* London: MIT Press.

McConkie, G. W., & Currie, C. B. (1996). Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 563-581. [PubMed]

Morrison, D. J., & Schyns, P. G. (2001). Usage of spatial scales for the categorization of faces, objects and scenes. *Psychonomic Bulletin & Review, 8,* 454-469. [PubMed]

Murray, J. E., Yong, E., & Rhodes, G. (2000). Revisiting the perception of upside-down faces. *Psychological Science, 6,* 492-496. [PubMed]

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9,* 353-383.

Perrett, D, Benson, P. J., Hietanen, J. K., Oram, M. W., & Dittrich, W. H. (1995). When is a face not a face? In Gregory, R., Harris, J., Heard, P., & Rose, D. (Eds.) (1995). *The artful eye* (Ch 5, pp. 95-124). Oxford University Press.

Rensink, R. A. (2000). Visual search for change: A probe into the nature of attentional processing. *Visual Cognition*, 7, 345-376.

Rensink, R. A. (2002). Change detection. *Annual Review of Psychology, 53*, 245-277. [PubMed]

Ro, T., Russell, C., & Lavie, N. (2001). Changing faces: A detection advantage in the flicker paradigm. *Psychological Science, 12*, 94-99. [PubMed]

Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features!: Understanding recognition from the use of visual information. *Psychological Science, 13*, 402-409. [PubMed]

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin and Review, 5*, 644-649.

Sperling, G. (1960). The information available in brief visual presentation. *Psychological Monographs, 74*, 29.

Suzuki, S., & Cavanagh, P. (1995). Facial organization blocks access to low-level features: An object inferiority effect. Journal *of Experimental Psychology: Human Perception and Performance, 21*, 901-913.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology, 46A*, 225-245. [PubMed]

Tanaka, J. & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition, 25,* 583-592. [PubMed]

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modelling of elementary psychological processes.* London: Cambridge University Press.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 92-114. [PubMed]

Wolfe, J. M. (1996). Visual search. In H. Pashler (Ed.), *Attention* (pp. 13-74). London, UK: University College London Press.

Yeshurun, Y. & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature, 396,* 72-75. [PubMed]

Yeshurun, Y. & Carrasco, M. (2000). The locus of attentional effects in texture segmentation. *Nature Neuroscience, 3*, 622-627. [PubMed]

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*, 141-145.

Young, A. W (1998). *Face and mind.* Oxford: Oxford University Press.

Young, A. W., Newcombe, F., de Haan, E. H. F., Small, M., & Hay, D.C. (1993). Face perception after brain injury: Selective impairments affecting identity and expression. *Brain, 116*, 941-959. [PubMed]

Zeki, S. (1999). *Inner Vision.* Oxford: Oxford University Press.