Chapter2

Face-gender Discrimination in the Near-Absence of Attention

2.1 Introduction

Visual properties that can be extracted with little attention have long been under investigation. As discussed in the introduction to this thesis, performance on these pre-attentive tasks does not suffer, even when they are carried out in parallel with known attentionally demanding tasks (e.g., in the dual-task paradigm). Typically, these tasks involve simple feature discriminations, such as discriminating between two colors or two orientations. However, once tasks involve discriminating conjunctions of simple features, attention can become a necessary requirement. Thus, when subjects are asked to discriminate between a bisected color pattern and its mirror image, performance is strongly contingent on attention.

Even these "more complex" discrimination tasks, however are leagues away from the sophistication and complexity[3] involved in everyday vision where richly detailed scenes must be constantly inspected. The question of the attentional demands of the processing of these scenes is an interesting one, because on the one hand, classical theories of attention [4] would predict that the visual system must rely on attention to bind together the different features of the objects in the scenes. On the other hand, however, as we have discussed, evidence exists that natural scenes can be classified with

[3] We use the notion of "complexity" here and in the following chapter as a conceptual argument, in a computational sense. For instance, complexity can be defined loosely as the number of elementary computer operations it takes to solve a problem.
[4] e.g., the FIT discussed in Chapter 1.

remarkable speed (Thorpe, Fize et al., 1996). Furthermore, our everyday experience is not that of tunnel vision—rather, even when deeply engaged in some task, we can be aware of objects outside the focus of attention.

This issue was recently addressed by Li and colleagues who found that observers were able to discern the gist of natural scenes in the near-absence of attention (Li, VanRullen et al., 2002). Using a dual-task paradigm, they presented subjects with a variety of images of natural scenes and asked them to report the presence of an animal or vehicle (Figure 2.1a). Although each image was only presented for less than 30 ms, subjects were able to achieve surprisingly high levels of performance on these tasks, regardless of the availability of attention (Figure 2.1b). These data thus demonstrate that natural stimulus processing is not limited by the availability of the attentional resource.

From a computational point of view, the result that natural scene categorization can be performed in the near-absence of focal attention is surprising since such categorization tasks are substantially more "complex" than discriminating between two-color patterns. It therefore appears that the "complexity" of a visual discrimination task does not necessarily determine its attentional demands. Instead, the nature of the stimuli involved (natural stimuli versus artificial combinations of individual features) also seems to influence a task's attentional requirements.

Where does this ability to process natural stimuli in the absence of focal attention break down? Would attention become necessary if the natural stimulus discrimination task became more "complex"—i.e. if targets and distractors are made more similar to each other and share the same features? To address this issue, we chose to determine the attentional demands of a face-gender discrimination task. Since both male and female faces share the same features (i.e., eyes, noses, mouths, etc.) and these features are similarly arranged in both sets of faces, this task could help us determine if the limits of pre-attentive processing of natural stimuli are determined merely by the complexity of the stimuli, or if even more complex natural stimuli can be processed pre-attentively.

Numerous experiments have explored the attentional demands of face processing. Although faces are believed to be of particular importance to the visual system (Farah, 1995; Kanwisher, McDermott, & Chun, 1997; Farah, Wilson, Drain, & Tanaka, 1998; Ro, Russell, & Lavie, 2001), most studies have failed to demonstrate a pop-out effect for faces or other natural objects in visual search (Nothdurft, 1993; Kuehn & Jolicoeur, 1994; Purcell, Stewart, & Skov, 1996; Brown, Huey, & Findlay, 1997; VanRullen, Reddy, & Koch, 2004). This suggests that face processing requires some form of attention. However, this result is still controversial (Hansen & Hansen, 1988; Suzuki & Cavanagh, 1995; Hochstein & Ahissar, 2002), and it is hoped that the present experiments will contribute to resolving this debate.

2.2 Methods

2.2.1 Participants

Six participants, including the author, were tested in Experiments 1, 2, and 3. Six additional participants were tested on Experiment 4 while another six were tested on Experiment 5. All participants (age range from 22 to 31 years) were undergraduate and graduate students or staff at the California Institute of Technology, who signed a consent form and were paid $13.50 per hour. By self-report they had normal or corrected-to-normal acuity.

2.2.2 Face database

The face stimuli used were obtained from the Max Planck Institute, Tübingen, Germany and contained 7 views of 100 male and 100 female faces (Vetter, 1998) (http://faces.kyb.tuebingen.mpg.de). This database of color photographs is well-matched for low-level features such as color, size, and illumination. All obvious gender cues, such as facial hair, have been removed from these images. Pilot experiments showed that the gender of some faces in the database was ambiguous with overall discrimination performance around 65%. Therefore, eight additional participants were asked to judge the gender of each face and rate their confidence on a 3-point Likert scale (this experiment was conducted by Patrick Wilken). The mean of these responses was converted into Z-scores. Each face was randomly presented 10 times for 1000 ms. The present gender-discrimination experiment used the 50 male and 50 female individuals that produced the highest mean male and female ratings. Examples of the faces used are shown in Figure 2.2b.

2.2.3 Apparatus

Participants were seated approximately 120 cm from a computer monitor connected to a Silicon Graphics (O2) computer for the dual-task experiments. The refresh rate of the monitor was 75 Hz. The face rating (described above) and face recognition experiments (Experiment 3) were performed on a Macintosh G4 computer; the refresh rate of the monitor was 75 Hz. The display was synchronized with the vertical retrace of the monitor.

2.2.4 Experiment 1: face-gender discrimination

The experiment consisted of two distinct tasks: an attentionally demanding, central letter discrimination task, and a peripheral, face-gender discrimination task. These tasks were performed in three conditions: blocks of the central or peripheral task alone, or a dual-task condition in which both central and peripheral tasks were performed concurrently. Subjects were instructed to be as accurate as possible, and no constraint was imposed on their reaction times. An auditory tone was provided as feedback on incorrect trials. The experimental timeline for one trial is illustrated in Figure 2.2a. In all three conditions the trials were arranged as shown in the figure and only the specific instructions to participants differed.

Central letter discrimination task

The attentionally demanding central task consisted of letter discrimination. Each trial started with a fixation cross presented 300 ± 100 ms before the onset of the first stimulus. At 0 ms, five randomly rotated letters (Ts and Ls, either all the same or one different from the other four) were presented at the center of the display at nine possible locations within 1.2° of fixation. Participants were required to report whether all five letters were identical or not by pressing one of two keys on the keyboard. The letters were individually masked by an "F," rotated by an angle corresponding to the "T" or "L" it replaced. The SOA (“Stimulus Onset Asynchrony”) was determined individually for each participant (see “Training” below). This task has been shown to be effective in engaging spatial attention at the center of the display (Braun & Julesz, 1998; Lee, Koch et al., 1999). In particular, Braun has shown that reducing the SOAs for each subject by as little as 30 ms can result in substantial drops in performance on this task (Braun & Julesz, 1998); see also (Li, VanRullen et al., 2002). Thus, the SOAs obtained for each subject by using the training procedure described below ensure that subjects do not have time to shift attention to and from the letters during the task, without impairment in performance.

Peripheral face-gender discrimination task

A face subtending approximately 2.5° of visual angle was presented peripherally 26 ms following the onset of the central stimulus. The face appeared at a random location centered on an edge of an imaginary rectangle subtending 8° x 10° of visual angle. Each face was masked by a pattern mask composed of scrambled faces; the peripheral stimulus was always masked before the central stimulus. The peripheral SOA was determined individually for each participant (see “Training” below). Participants were required to report the gender of the face using two keys on the keyboard.

Dual-task

In the dual-task condition, participants were asked to respond to first the central task (with the left hand) and then the peripheral one (with the right hand), and fixate at the center.

Training

At the beginning of training, the letters were displayed for 500 ms and the faces for 160 ms before the mask appeared (Figure 2.2, (see also Movie 1))[5]. Through the course of training, “letter” and “face” SOAs were decreased (when mean performance in a 48-trial block exceeded 90%). In order to limit the possibility of eye movements, the letter SOA was decreased to below 250 ms for all subjects. Thus, training was complete when participants’ letter SOA had stabilized below 250 ms for a one-hour session. After training, over the group of participants, the “face” SOA varied between 133 ms-160 ms and the “letter” SOA between 173 ms–240 ms. This procedure, coupled with the high motor demands of the dual-task paradigm, meant that participants required extensive training (between 6 and 12 hours per participant). For three of the six participants, one

[5] See http://journalofvision.org/4/2/4/article.aspx for the movie

set of 350 randomly selected faces (7 views of 50 individuals) was used as stimuli, while the other three participants were trained on a different set of 350 faces. Participants received the same amount of training in all tasks.

Data collection

Once training was complete, the letter and face SOAs were fixed for each subject and data was collected over five one-hour sessions. Each session consisted of four blocks of 48 trials in each single-task condition and six blocks of 48 trials of the dual-task condition. A session was considered valid if dual-task letter performance was not significantly lower (t test, $p > .05$) than single-task letter performance. This served to ensure that participants were effectively focusing attention on the central letter task. Over the six participants, only two sessions were rejected as a result of this criterion.

2.2.5 Experiment 2

In a separate dual-task session, all six participants from Experiment 1 were asked to perform gender discrimination on a set of novel stimuli (7 views of the 50 individuals they had not seen in Experiment 1) using the same method as previously used, but with no further training. Participants performed only one session of this type.

2.2.6 Experiment 3

Experiment 3 was performed on the same day as Experiment 2, with participants (except the author) from Experiments 1 and 2. In two separate sessions of this experiment, participants were presented with the faces they had viewed during Experiment 1 (the “familiar” images) and Experiment 2 (“control” faces), respectively, along with an equal number of faces they had never seen before. Each face was shown centrally for 1000 ms. Participants reported whether they recognized the face or not using two keys on the keyboard. The first session was run before, and the second after, Experiment 2.

2.2.7 Experiment 4

In a separate set of experiments, six participants who had been trained on a different dual-task experiment (Li, VanRullen et al.) were tested in our paradigm for one hour per day for two consecutive days. These participants had been trained on the same central letter discrimination task, but a different peripheral task (animal vs. non-animal and vehicle vs. non-vehicle discrimination). In the paradigm they had been trained on, these participants responded to the peripheral stimulus by releasing the mouse button. Thus, instead of reporting whether the face presented was male or female with different keys on the keyboard, three of these participants were asked to release the mouse button if the face was male while the other three released the button if the face was female. During the experiment, the participants viewed a different image set each day.

2.2.8 Experiment 5

In Experiment 5, six new participants were tested. They were trained on three different peripheral tasks: upright face-gender discrimination (i.e., the same task as in Experiments 1 and 2), inverted face-gender discrimination (i.e., where each face was rotated by 180°), and a discrimination between two color patterns—a vertically bisected disk with red and green halves or such a disk rotated by 180° (i.e., they had to discriminate a red-green disk from its mirror image, a green-red disk). In individual dual-task blocks, participants performed the central letter discrimination task together with one of the three peripheral tasks. Each session consisted of four blocks of the single central-letter task, two blocks of each single peripheral-task, and three blocks of each dual-task. The faces were masked by a pattern mask composed of scrambled faces (as before) while the disks were masked by a disk divided into four red and green alternating quadrants. The tasks were matched for difficulty such that single-task performance for all three peripheral tasks was on average 75%. Participants received an equal amount of training on the three peripheral tasks. The same face set and training and data collection methods were used as in Experiment 1. The SOA for the disks was 98.1 ± 4.1 ms.

2.2.9 Data analysis

A 1-way ANOVA and paired t-tests were computed for each experiment to compare single and dual-task performance. An alpha value of .05 was used for all statistical tests. Normalized performances in the dual-task experiment were calculated by a simple linear scaling of the mean value of each participant's performance. The scaling mapped the mean single-task performance to 100% leaving chance at 50%: Normalized performance = 1/2 + 1/2. $[(P_2 – 1/2) / (P_1 – 1/2)]$, where P_2 and P_1 refer to performance in the dual-task and single-task conditions, respectively.

2.2.10 Eye movement control

In all our dual-task sessions, the presentation times of the central and peripheral stimuli were less that 250 ms, and the peripheral stimuli were always presented at random locations. These constraints were imposed in order to limit the possibility of eye movements, and it is indeed unlikely that eye movements contributed significantly to the observed behavioral performance of our subjects. However, to directly confirm this, in a separate control experiment 3 subjects performed the face-gender dual task experiment while their eye movements were monitored. They were seated 75 cm from a computer monitor on which the stimuli were presented. The monitor measured 26º x 19º of visual angle. Subjects rested their chins on a chinrest in order to minimize head movements. An infrared (IR) eye-tracking system obtained from ISCAN, Inc. was used to measure and record eye positions. The camera and IR beam generator were placed approximately 60 cm away from the subject, and the right eye was illuminated with invisible IR light (~850 nm). The camera sampled the image of the eye at 120 Hz, and the image was processed in real time to obtain the position of two features of the eye: the IR-dark spot at the center of the pupil and the IR-bright spot on the cornea (the corneal reflection). The difference of these two values, in principle, gives a measure of the position of the eye independent of head position.

During calibration trials that were run prior to the start of the experiment, and after every two blocks, subjects fixated 9 calibration points for 2 s each. The calibration points were presented on a 3 x 3 grid on the monitor. These calibration trials gave us the value of the eye position in camera coordinates. Using these values and the position of the fixation points in the stimulus display, a correspondence between camera coordinates and the stimulus display coordinates could be established. This correspondence was then used to ascertain eye positions during the dual-task experiment in stimulus display coordinates. The mean error of the scanner over all subjects during the calibration runs was estimated to be about 0.40 º.

In the dual-task experiment, the eye positions were measured during each trial (from the appearance of the first fixation cross up until when the letter masks appeared) at 120 Hz. The median and mean of these values over all trials in each of the three experimental conditions were computed. This data is discussed in the results section below and is presented in Figure 2.7.

2.3 Results

The effects of attentional manipulation on face-gender discrimination were studied with a dual-task paradigm in which participants performed a central, attentionally demanding, letter discrimination task as well as a second, peripheral face-gender discrimination task either concurrently or separately. The role of attention on gender discrimination was measured by comparing performance on the peripheral task, when this task was performed alone (single-task condition), with performance under dual-task conditions. If gender discrimination requires little or no attentional resources, peripheral performance will suffer minimally in the dual-task condition compared to the single-task condition. If, however, the peripheral task does require attention, performance should be severely impaired under the dual-task condition (Sperling & Melchner, 1978; Braun & Sagi, 1990; Braun & Julesz, 1998).

In Experiment 1, six participants were tested on this paradigm (Figure 2.3). Their performance on the central letter discrimination task when performed alone was on average 83.1% ± 4.1% (mean ± s.d.). This value can be compared with performance on this task in the dual-task condition (83.4%± 5.6%): if a participant's attention is engaged by the central letter task, performance in the dual-task condition should be equivalent to performance in the single-task condition; otherwise there should be a significant decrease in performance levels. For our participants, there was no significant difference in performance on this task between the single and dual-task conditions (*t*-test, $p > .05$). When participants performed the face-gender discrimination task alone, performance was on average 77.6% ± 3.8%. This comparatively lower value reflects the short stimulus exposure and the fact that obvious gender cues, such as the presence of facial hair, were removed from the images. Performance on this task in the dual-task condition (74.9% ± 4.0%) was also not significantly different (*F* (1, 10) =1.52, *p* = .2) from performance in the single-task condition over the group of six participants (Figure 2.3a). For five of the six participants, individual *t*-tests revealed no significant difference in performance (*p*>.05) between these two conditions. Figure 2.3b summarizes these results: in the face-gender discrimination task, performance for all six participants in the dual-task condition was above 90% of their performance in the single-task condition (normalized plot, see Methods). These results indicate that, although there may be a small decrement in the dual-task condition, face-gender discrimination can still be performed efficiently with little or no spatial attention resources available and constitutes the main finding of this study.

In order to limit the possibility of eye movements, the central SOA was maintained below 250 ms for all participants, and the peripheral stimulus could appear anywhere at one of eight peripheral locations. This constraint, together with the high motor demands of the dual-task procedure, meant that participants required extensive training (between 6 and 12 hours per participant) with the same set of male and female images (referenced hereafter as the “familiar” face set). Consequently, it could be argued that instead of performing gender discrimination as required, participants were actually using a strategy akin to face recognition. To control for this potentially confounding effect, the same participants were tested on a set of novel faces (“control” faces) in Experiment 2 (Figure 2.4a). Despite the novelty of the control face set, over the group of participants, the difference in performance on gender discrimination between single and dual-task conditions was not significant (*F*(1,10)=1.43, *p*=.3). Individually, for five of the six participants, performance was not significantly different between these two conditions (79.1% ± 4.8% and 75.6% ± 5.1%, respectively; paired *t*-test, *p* > .05).

Although there was a modest decrement in the dual-task condition, face-gender discrimination performance for all six participants was above 85% of their original single-task performance (normalized plot, Figure 2.4a). Note that the central task performance in the dual-task condition was not significantly lower than performance in the single-task condition for each participant (*t*-test, $p>.05$), indicating again that attention was effectively engaged at the center in the dual-task condition. From this control experiment, it appears that familiarity with the face set is not critical to the observed performance.

In fact, results from an additional control experiment (Experiment 3), indicate that participants had not gained any appreciable familiarity with either of the face sets they had viewed during Experiments 1 or 2. In separate sessions, participants were presented with the faces viewed extensively during the training and data collection phases in Experiment 1 ("familiar" faces) or faces viewed in Experiment 2 ("control" faces), as well as an equal number of completely novel faces. (The "familiar" faces had been viewed between 18 and 30 times, while the "control" faces had each been viewed twice during the course of experiments 1 and 2, respectively. Each presentation of the face had lasted between 143 ms and 160 ms, depending on the participants' SOA; see Methods). Participants were asked to report for each face whether they had seen it at least once during Experiment 1 or 2. Surprisingly, for both the "familiar" and the "control" sets of images, participants' performance on this recognition task (52.1 ± 3.4% for the familiar face set and 51.1 ± 2.2 % for the control face set, Figure 2.5) was not significantly different from chance levels ($p = .2$, $p = .4$, respectively, paired *t*-test). Thus, it appears that despite having viewed some of the faces several dozen times, participants were unable to differentiate the stimuli in either face set. These results confirm that the pattern of performance observed in Experiments 1 and 2 cannot be accounted for by familiarity with the stimuli used.

Since participants had been extensively trained on the face-gender discrimination task, it could still be argued that they had learned low-level features in the image set, which would contribute significantly to the observed performance. To control for this, six new participants were tested on our gender-discrimination task (Experiment 4). They had been trained on a completely different dual-task experiment (natural scene categorization: animal vs. non-animal or vehicle vs. non-vehicle) (Li, VanRullen et al., 2002). Data was collected over two days with a new set of stimuli on each day. Despite the novelty of the peripheral task, participants performed comparably well in the dual-task and single-task conditions (Figure 2.4b). While performance on the gender-discrimination task was significantly lower (F (1, 10) =5.4, p=.04) in the dual-task (69.7 ± 5.6%) vs. single-task (75.91 ± 6.2%) condition over the group of participants, there was, individually, no significant difference in performance for four of the six participants (p>.05, paired t-tests). The normalized results shown in Figure 2.4b indicate that despite the novelty of the task, performance in the dual-task condition was above 80% of performance on the single-task condition for all six participants. We conclude therefore that there was no strong or consistent confounding effect of training in our gender discrimination task.

Thus, whether involving highly familiar or completely novel faces, or even a completely novel discrimination task, there is only a modest decrement in performance on face-gender discrimination in the near-absence of focal attention.

Finally, in order to rule out the possibility that low-level cues in the face dataset could account for the observed results, we tested six additional participants in Experiment 5. In this experiment, participants were required to perform face-gender discrimination on both upright and inverted faces, using the same method as Experiment 1. Inverted faces provide a suitable control for basic low-level characteristics (e.g., contrast, luminance, spatial frequency, etc.), which might aid gender discrimination. If the observed results were due to low-level statistical properties of male and female faces, equally high levels of performance would be observed in both the upright and inverted face-gender discrimination tasks.

Participants received the same amount of training in both the upright and inverted face-gender discrimination tasks, and the level of difficulty was matched so that the mean single-task performance was about 75% for both tasks. Consistent with the results of Experiments 1 and 2, participants again achieved a high level of performance on upright face-gender discrimination in the dual-task condition compared to the single-task condition (Figure 2.6a; 71.3% ± 3.4%, 75.5% ± 4.0%, respectively). Over the group of six participants, a 1-way ANOVA revealed no significant difference in performance in the dual and single-task conditions (F (1, 10) =3.62, p = .09). Individually, there was no significant difference between these two conditions for four of the six participants (t-test, p>.05), and all six participants performed above 85% of their original single-task performance. In contrast, based on a 1-way ANOVA, the six participants showed a significant decrease in performance (F(1,10) = 25.7, p < .001) in the inverted face-gender discrimination task when attention was unavailable compared to the single-task condition (59.7% ± 4.7%, 71.7% ± 3.3%, respectively), and individual tests for each participant revealed a significant decrease in performance in this dual-task condition for all six participants (t-test, p<.05; Figure 2.6b). Further, for each of the six participants, performance in the inverted dual-task condition was significantly lower than performance in the upright dual-task condition (p <.05, t-test). We conclude that the observed performance in upright face-gender discrimination cannot be accounted for by the low-level statistical properties of the stimulus set.

The interpretation of the results reported here relies on the assumption that the central letter task efficiently engages attention in the dual-task condition and that performance on attentionally-demanding tasks should suffer dramatically in the dual-task condition. As a further control, we verified that performance on a known attentionally-demanding task would indeed be severely impaired in the dual-task condition (Braun & Julesz, 1998; Li, VanRullen et al., 2002). We had the same six participants discriminate between a masked color disk and its mirror image in the dual-task condition. In our experiment, participants received the same amount of training in all three discrimination tasks (upright face-gender, inverted face-gender, and colored-disk discrimination), and task difficulty was matched so that single-task performance was about 75% for all three tasks. In contrast to the results observed for upright face-gender discrimination and consistent with previous studies (Braun & Julesz, 1998; Lee, Koch et al., 1999; Li, VanRullen et al., 2002), we observed (Figure 2.6c) a dramatic decrease in performance over the group of six participants when the colored-disk discrimination task was performed in the dual-task versus single-task condition (51.8% ± 3.4%, 76.1% ± 8.5% respectively, $F(1,10) = 42.0$, $p < 10^{-4}$). As shown in Figure 2.6c, normalized performance values were between 45% and 60% of single-task levels, and for five of the six participants these values were not significantly different from chance levels of performance (paired *t*-test, $p > 0.1$). These results confirm that under our experimental conditions, the attentional requirements of the central task can result in a clear decrease in dual-task performance.

To determine whether eye movements contributed significantly to the observed performance, three subjects performed the dual-task experiment while their eye positions were monitored. For each of the three experimental conditions (central, peripheral, and dual), the median and maximum eye positions were calculated on each trial for each subject. The data for one subject is presented in Figure 2.7a. In this figure (drawn to scale), the outermost gray rectangle depicts the extent of the computer monitor on which the stimuli were presented. The yellow hashed area represents the region occupied by the face stimuli. It is important to note that no part of the face stimuli extended beyond this yellow area. The innermost rectangle, delineated by a solid black line, represents the 1.2º window in which the letter stimuli were presented. The red crosses mark the median and maximum eye positions on individual trials in each of the three conditions. As the figure illustrates, eye positions in each of the three conditions were limited to the area occupied by the letters and never entered the area occupied by the face stimuli. Thus, this subject did not make large eye movements in the direction of the peripheral stimuli in any of the three experimental conditions. Comparable results were obtained for other subjects in this experiment.

Of course, this eye movement control is only meaningful if subjects are performing the task with comparable performance levels as those obtained in our main experiment. The behavioral data for the three subjects tested with the eye-tracker are shown in Figure 2.7b. The first panel presents the normalized performance on face-gender discrimination over all trials in the experiment. In the second and third panels, we have imposed an upper bound on the permissible eye movements. Thus in the second panel, all trials on which the maximum eye position during the trial exceeded 2º were rejected. Behavioral performance over the remaining trials was calculated. In the third panel, all trials on which the median eye position exceeded 1º were rejected. After imposing these limits on eye positions, the behavioral performance observed on the face-gender discrimination task (the second and third panels) is not significantly different from performance shown in the first panel. Thus, this data confirms that eye movements do not make a significant contribution to the behavioral performance of subjects on this task.

2.4 Discussion

Our findings demonstrate that telling a very briefly flashed male from female face, a fine discrimination task, can be performed remarkably well when spatial attention is engaged elsewhere. We have shown that participants can achieve a high level of performance in the presence of little or no focal attention when they are tested on a set of completely unfamiliar faces, even when they are unfamiliar with the task itself. In control experiments, we have shown that these results cannot be explained by eye movements. Further, when participants perform the same face-gender discrimination task in the near-absence of spatial attention on a set of inverted faces, performance is significantly impaired compared to performance on this task with upright faces. These results demonstrate that the observed findings cannot be attributed to low-level characteristics of the image set. Previous psychophysical studies have shown that face recognition is impaired when the faces are inverted rather than upright (Yin, 1969; Valentine, 1988; Valentine & Bruce, 1988; Brown, Huey et al., 1997). Additionally, even though functional imaging studies have suggested that inverted face processing recruits additional brain areas compared to upright face processing (Haxby, Ungerleider et al., 1999), electrophysiology in monkeys has revealed that face-specific cells maintain responses to inverted faces, but that these responses are weaker and longer in latency compared to those evoked by upright faces (Perrett, Mistlin et al., 1988). Our results suggest a differential requirement of spatial attention by these two tasks: the absence of spatial attention has a pronounced effect on the processing of inverted faces, but not upright faces.

It should be noted that while 19 of the 24 datasets we obtained did not demonstrate any significant decrease in performance in the dual-task conditions, the remaining five did show some decrement. However, some decrement in performance is expected to occur when participants perform two demanding tasks concurrently, compared to when the tasks are performed alone. These performance decrements do not necessarily imply competition for an attentional resource, but could be attributed to other factors such as having to maintain two sets of task goals, or having to encode and produce two sets of motor responses (Allport, 1980; Duncan, 1980a; Pashler, 1984, 1994). In addition to comparing single and dual-task performance, it is revealing to compare the dual-task performance of our participants in the face-gender discrimination task with dual-task performance on tasks that are known to require attention (Braun & Julesz, 1998; Li, VanRullen et al., 2002). As we have shown, performance on a known attentionally demanding task (discriminating a red-green from a green-red disk) drops to chance levels when the available spatial attention is severely reduced. In contrast, performance on our face-gender discrimination task remains consistently above 80% of single-task performance when attention is engaged elsewhere. Indeed, a statistical comparison of all 24 datasets we collected indicates that all our participants perform face-gender discrimination in the dual-task condition significantly above chance (*t*-test, $p < 10^{-16}$).

From a computational perspective, we designed our peripheral task to be challenging: this task did not merely involve the discrimination of targets and distractors at a basic level of categorization, but required a fine discrimination within a category level, between male and female faces that share the same overall structure and lack hair and other obvious gender cues. In essence, this meant a fine discrimination of the spatial arrangement of highly similar features present in both targets and distractors. Our results indicate that such discrimination can be carried out in the presence of a primary task highly effective in requiring attentional resources (Braun & Julesz, 1998; Lee, Koch et al., 1999; Li, VanRullen et al., 2002). This supports the notion that the "complexity" of a task as measured by its computational demands does not necessarily determine its attentional requirements. Classical views of selective, visual attention have suggested that while simple salient stimuli can be detected outside the focus of attention, attention plays a key role in the recognition of more complex stimuli. In other words, it has been proposed that attention is necessary to combine the different low-level features of a stimulus into a coherent representation of the object (Treisman & Gelade, 1980). Access to this representation is supposed to be necessary for object recognition and behavior. Our findings argue that face-gender discrimination is possible in the near-absence of attention. Although this conclusion cannot be generally extended to other subordinate—level categorization tasks involving natural stimuli, our approach shows that attention is not always necessary for such tasks. The possibility that faces hold a special status for the visual system is still under debate (Farah, 1995; Gauthier & Tarr, 1997; Kanwisher, McDermott et al., 1997; Farah, Wilson et al., 1998; Tovee, 1998; Gauthier, Skudlarski, Gore, & Anderson, 2000; Ro, Russell et al., 2001; Bogen & Berker, 2002). It would thus be interesting to test the role of attention in other “complex” discrimination tasks and to determine whether expertise in other areas yields similar results.

If a failure to pop-out during a search task is taken to indicate the necessity of focal attention for recognition, then our results appear to contradict a number of studies that have shown that facial information does not "pop-out" in a visual search situation (Nothdurft, 1993; Kuehn & Jolicoeur, 1994; Purcell, Stewart et al., 1996; Brown, Huey et al., 1997). However, it is worth noting that earlier studies had suggested that faces can be processed in parallel (Hansen & Hansen, 1988), and this issue is still controversial and open to debate (Hochstein & Ahissar, 2002). Furthermore, the correspondence between dual-task and visual search results has recently been called into question (VanRullen, Reddy et al., 2004). More supportive evidence for the pre-attentive processing of faces comes from clinical reports of patients with visual neglect (Vuilleumier, 2000; Vuilleumier, Sagiv et al., 2001). For these patients, extinction was less likely to occur for faces presented in the neglected hemifield than other objects (e.g., meaningless shapes). In other words, faces can attract attention more efficiently and thus probably have a competitive advantage at the pre-attentive level. Such observations are compatible with ERP and magneto-encephalography (MEG) investigations of the latency of face or face-gender selective responses, which was found to be on the order of 100 ms–150 ms (Schendan, Ganis, & Kutas, 1998; Yamamoto & Kashikura, 1999; Mouchetant-Rostaing, Giard, Bentin, Aguera, & Pernier, 2000; Liu, Harris, & Kanwisher, 2002). Given this remarkable speed, one wonders whether such processing can depend critically on visual attention.

In neural terms, several electrophysiological investigations have found single neurons responsive to faces in the infero-temporal cortex of monkeys, the "end-point" of the ventral visual hierarchy (Gross, Rocha-Miranda, & Bender, 1972; Bruce, 1982; Perrett, Rolls, & Caan, 1982; Desimone, Albright, Gross, & Bruce, 1984; Perrett, Smith et al., 1984; Rolls, 1984; Perrett, 1987). Similar observations have been made in humans in medial temporal lobe structures ((Kreiman, Koch, & Fried, 2000a; Quian Quiroga, Reddy, Kreiman, Koch, & Fried, 2005); see also Chapter 5). Several neuro-imaging studies have shown the existence of higher-level brain regions (such as the Fusiform Face Area, FFA) that selectively process facial information (Sergent, Ohta, & MacDonald, 1992; Haxby, Horwitz et al., 1994; Kanwisher, McDermott et al., 1997; Kanwisher, Stanley, & Harris, 1999) although some models of face recognition have conjectured that gender discrimination could occur in more posterior temporal areas (Bruce & Young, 1986). Consequently, it is not unreasonable to suppose that our stimuli differentially activate neurons in such high-level areas and that gender discrimination can rely on the selectivity of these neurons. Some evidence shows that these areas can be modulated by attention (Wojciulik, Kanwisher, & Driver, 1998; O'Craven, Downing et al., 1999; Pessoa, McKenna, Gutierrez, & Ungerleider, 2002), but the present results indicate that the residual activity in the near absence of attention is sufficient for the efficient processing of faces. Our findings, together with those of Li *et al*. (Li, VanRullen et al., 2002) and Rousselet *et al*. (Rousselet, Fabre-Thorpe et al., 2002), suggest that the activation of such high-level neuronal populations can take place in the near-absence of attention.

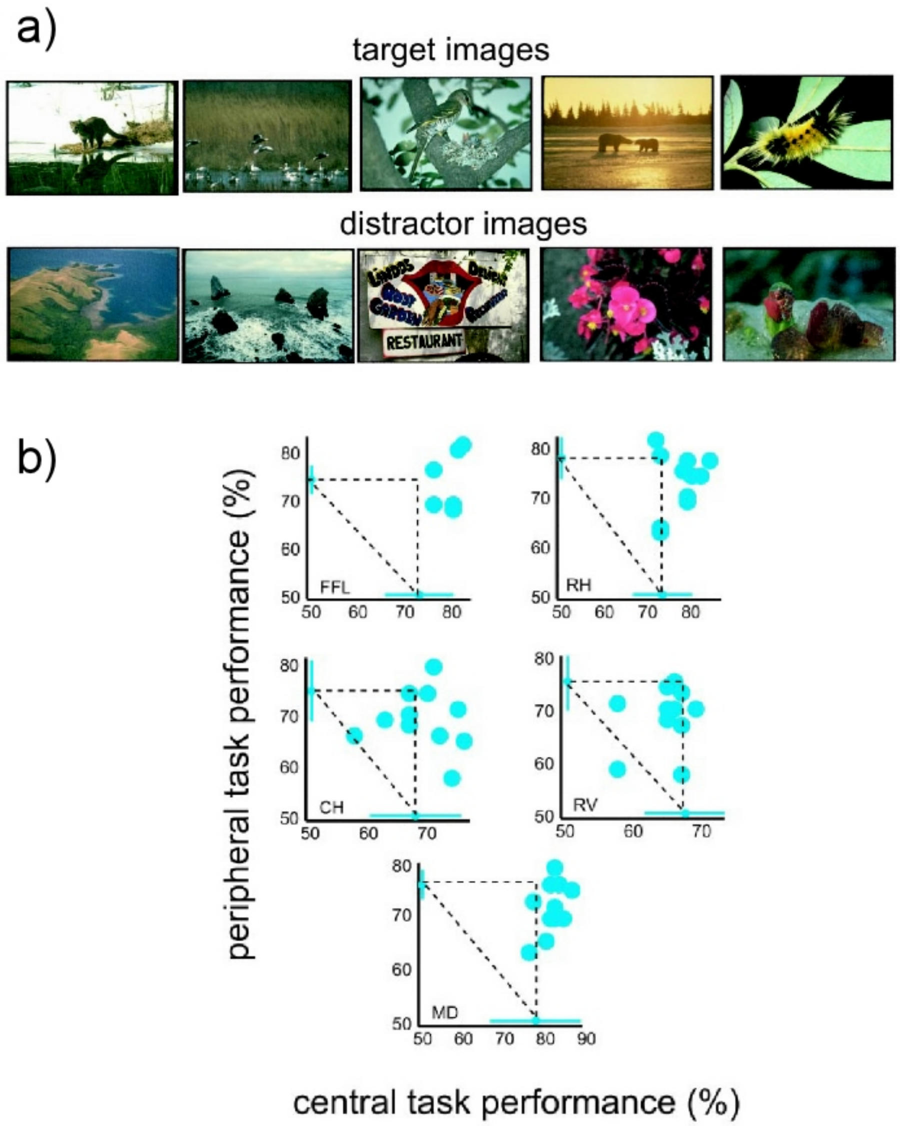

Figure 2.1: Natural scene categorization in the near-absence of attention (Adapted from Li et al., 2002). a) Examples of target (scenes containing animals) and distracter scenes (scenes with no animals) used in the experiment. b) Performance of 5 subjects on the natural scene categorization task. Performance on the peripheral task (natural scene categorization) is plotted against performance on the central letter discrimination task for each subject. The filled circles represent performance on individual blocks in the dual-task condition. Performance on the two tasks in the single task condition is plotted as a line along the respective axes, where the length of the line represents the standard deviation in performance. For each subject the difference in performance on both tasks was not significantly different ($p > .05$) between single and dual-task conditions indicating that the animal detection task is performed well in the near-absence of attention. Copyright 2002, National Academy of Sciences, USA.

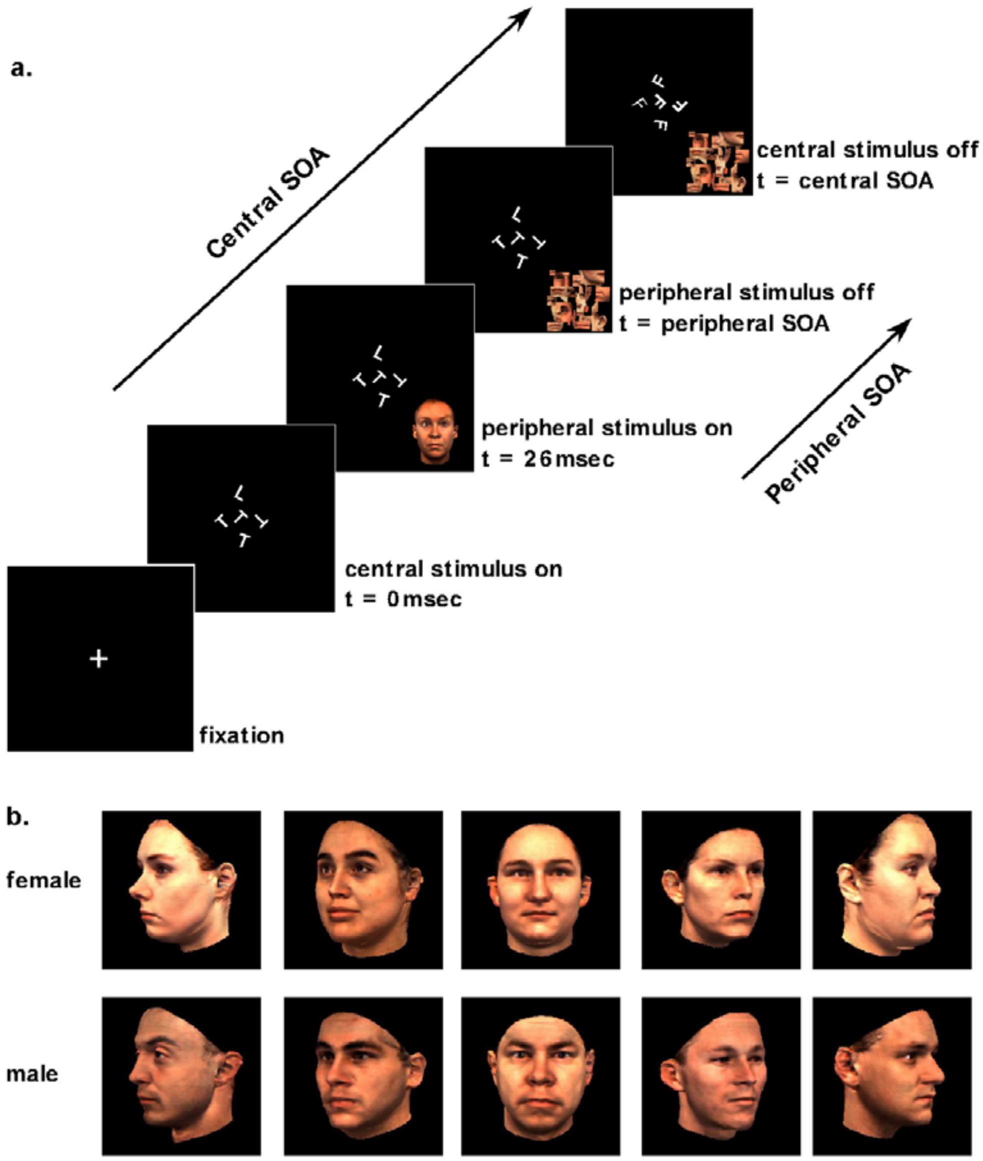

Figure 2.2: Face-gender discrimination dual-task experiment. (a) Schematic timeline for one trial in the dual-task experiment. At the end of a trial, participants are required to report the gender of the face presented and/or whether the 5 central letters were the same (either 5 Ts or 5 Ls) or different (4 Ts and 1 L or 4 Ls and 1 T). All trials are arranged similarly, independent of the specific instructions. Both letters and faces were masked individually. Central SOA (~200 ms) and peripheral SOA (~145 ms) indicate the presentation time for letters and faces, respectively. (b) Exemplars of male and female faces used in the experiment. (From (Reddy, Wilken, & Koch, 2004); copyright, Journal of Vision, 2004)).

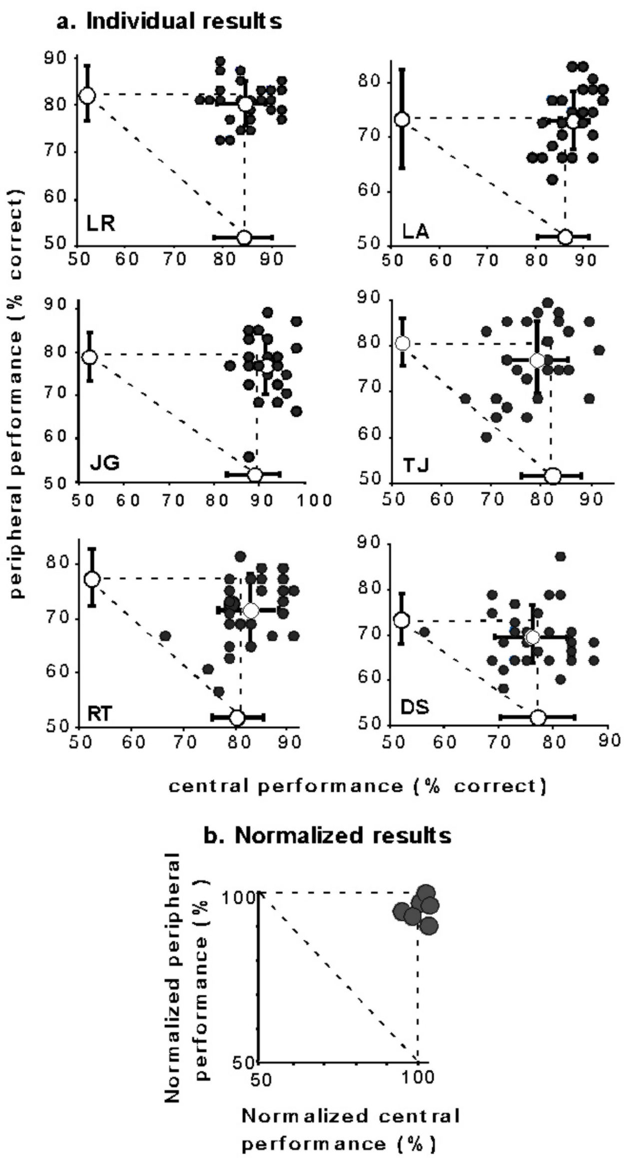

Figure 2.3. Results from six participants in the dual-task paradigm. *(a)* The horizontal axis represents performance on the attentionally demanding central letter task. The vertical axis represents performance on the peripheral gender discrimination task. Each filled circle is the participant's mean performance in the dual-task in one block of 48 trials, while an open circle represents mean performance in the three experimental conditions: single-central task, single-peripheral task, and the dual-task. By default, performance of the "to-be-ignored" task is assumed to be at chance level (50%) in the single-task condition. Error bars represent standard deviation. For all participants except one, (RT), face-gender discrimination performance in the dual-task condition is not significantly worse (*t*-test, $p>0.05$) than performance in the single-task condition indicating that face-gender discrimination suffers only minimally when performed concurrently with an attentionally-demanding task. *(b)* Normalized average performance for each participant in the dual-task paradigm. Each point represents a participant's performance in the dual-task normalized to their single-task performance. Normalized values are obtained by a linear scaling that maps the average single-task performance to 100% leaving chance at 50%(Methods). Normalized gender-discrimination performance values lie above 90% of single-task performance, suggesting that participants can perform face-gender discrimination remarkably well in the near-absence of attention. (From (Reddy, Wilken et al., 2004); copyright, Journal of Vision, 2004)).

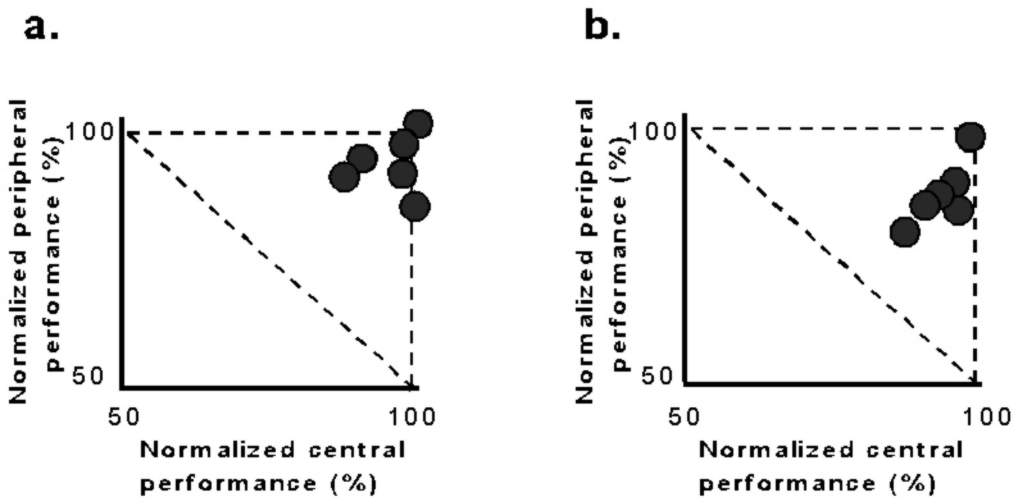

Figure 2.4. Control experiments for familiarity with images and training. *(a)* Normalized average performance for six participants in the dual-task paradigm using a completely novel set of faces. (Notation as in Figure 3b). Normalized dual-task performance lies above 85% of single-task performance for all participants, indicating that even with a novel set of faces, gender discrimination is performed well under the dual-task condition. *(b)* Normalized average performance for six participants who had been trained on a completely different dual-task paradigm. Normalized dual-task performance lies above 80% for all participants. This suggests that in spite of unfamiliarity with the gender discrimination task, performance was only marginally impaired in the dual-task condition. (From (Reddy, Wilken et al., 2004); copyright, Journal of Vision, 2004)).

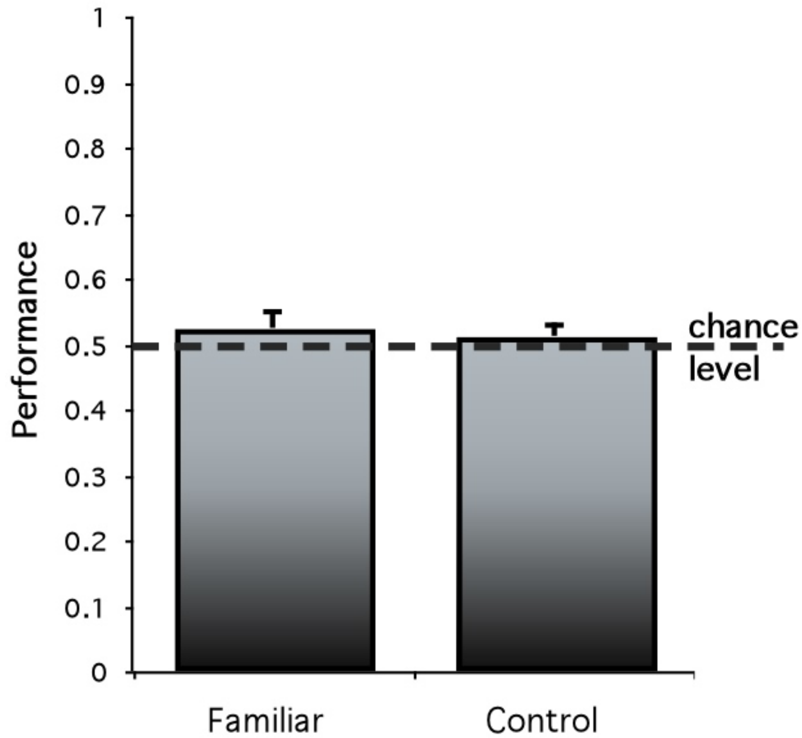

Figure 2.5. Face recognition experiment. Participants (the same as those shown in Figures 3 and 4a) were presented with faces they had viewed during the study and an equal number of novel faces and asked to report whether they recognized the face or not. The "familiar" image set is the one participants were trained on whereas the "control" faces had been viewed only twice each. In both cases, participants are at chance level at discriminating previously seen faces from novel faces, indicating that they had formed no explicit representation of the face sets. Error bars represent standard deviation. (From (Reddy, Wilken et al., 2004); copyright, Journal of Vision, 2004)).

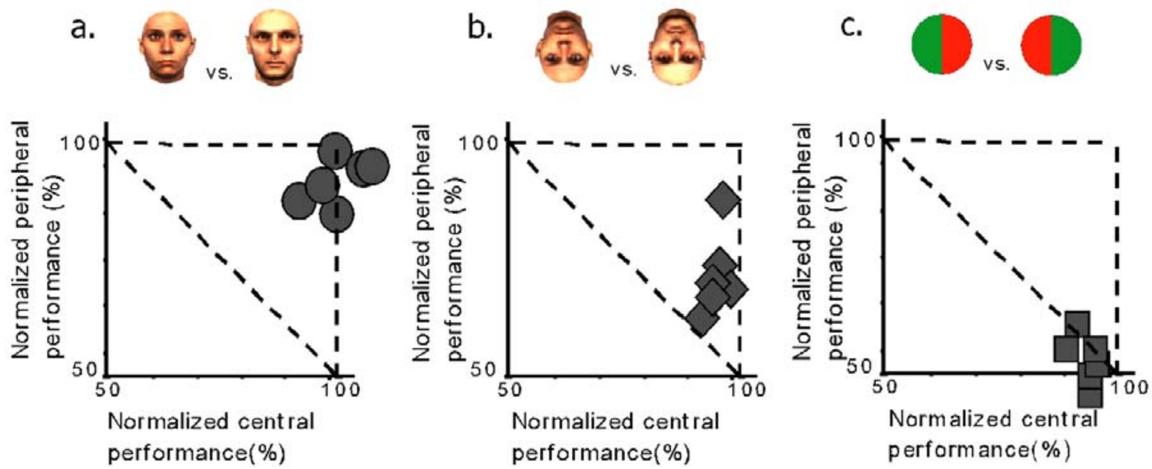

Figure 2.6. Control experiments for low-level visual cues and to test the efficacy of the central task in withdrawing attention. Normalized dual-task results of six participants in three tasks. *(a)* Upright face-gender discrimination task. Normalized dual-task performance values are on average 92% of single-task performance levels for upright face-gender discrimination, as expected from results shown in Figure 3b. *(b)* On the other hand, in the inverted face-gender discrimination task, normalized dual-task performance values are on average 72% of single-task performance levels, demonstrating that in the near-absence of attention, performance is impaired. In addition, for each participant, there is a significant decrease in performance when the task involves inverted faces compared to upright faces. Thus low-level visual cues cannot account for the pattern of results obtained in the upright face-gender discrimination task. *(c)* Color pattern discrimination task. Normalized dual-task values are on average 53%, demonstrating that attention is effectively withdrawn by the central letter task in dual-task conditions. (From (Reddy, Wilken et al., 2004); copyright, Journal of Vision, 2004)).

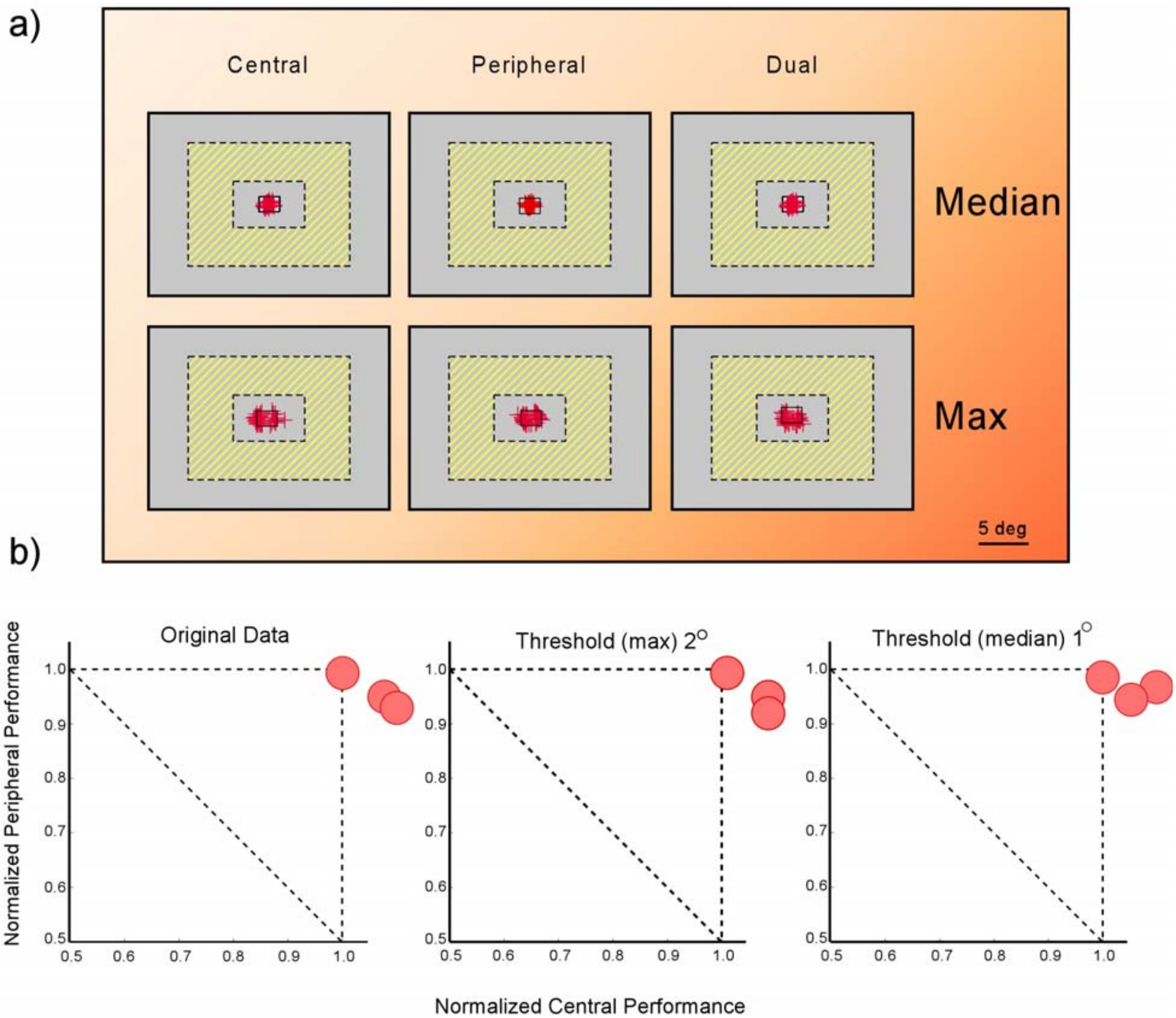

Figure 2.7 Eye movement control for 3 subjects. a) The median and maximum eye position during individual trials in all three conditions is presented for one subject. The figure is to scale. The grey area represents the extent of the computer monitor on which stimuli were displayed. The yellow-hashed area marks the inner and outer limits of the region occupied by the faces. Thus no part of the face stimuli was presented outside this area. The innermost rectangle, delineated by a solid black line, marks the extent of the letter stimuli (1.2º). Each red cross is the median and maximum eye position on individual trials. The data presented here shows that the subject's eye movements never entered the region occupied by the face stimuli but were instead limited to the area occupied by the letters. b) Behavioral performance for the three subjects. In the first panel, performance over all trials is shown. In the second panel, trials on which the maximum eye position was larger than 2º were rejected, and performance over the remaining trials was calculated. In the third panel, trials on which the median eye position was larger than 1º were rejected. The behavioral performance obtained after imposing these restrictions is not significantly different from performance shown in panel 1. Thus, the data indicate that eye movements do not contribute significantly to the high levels of behavioral performance we have observed.