Flawless visual short-term memory for facial emotional expressions

Éva M. Bankó

Faculty of Information Technology,
Pázmány Péter Catholic University,
Budapest, Hungary

Viktor Gál

Faculty of Information Technology,
Pázmány Péter Catholic University,
Budapest, Hungary, &
Magnetic Resonance Research Center,
Szentágothai J. Knowledge Center,
Semmelweis University,
Budapest, Hungary

Zoltán Vidnyánszky

Faculty of Information Technology,
Pázmány Péter Catholic University,
Budapest, Hungary, &
Magnetic Resonance Research Center,
Szentágothai J. Knowledge Center,
Semmelweis University,
Budapest, Hungary

Facial emotions are important cues of human social interactions. Emotional expressions are continuously changing and thus should be monitored, memorized, and compared from time to time during social intercourse. However, it is not known how efficiently emotional expressions can be stored in short-term memory. Here we show that emotion discrimination is not impaired when the faces to be compared are separated by several seconds, requiring storage of fine-grained emotion-related information in short-term memory. Likewise, we found no significant effect of increasing the delay between the sample and the test face in the case of facial identity discrimination. Furthermore, a second experiment conducted on a large subject sample (N = 160) revealed flawless short-term memory for both facial emotions and facial identity also when observers performed the discrimination tasks only twice with novel faces. We also performed an fMRI experiment, which confirmed that discrimination of fine-grained emotional expressions in our experimental paradigm involved processing of high-level facial emotional attributes. Significantly stronger fMRI responses were found in a cortical network—including the posterior superior temporal sulcus—that is known to be involved in processing of facial emotional expression during emotion discrimination than during identity discrimination. These findings reveal flawless, high-resolution visual short-term memory for emotional expressions, which might underlie efficient monitoring of continuously changing facial emotions.

Keywords: emotion, faces, social cognition, facial identity, superior temporal sulcus

Introduction

Facial emotional expressions are crucial components of human social interactions (Ekman, 1973; Fridlund, 1994; Izard, 1977). Among many of its important functions, facial emotions are used to express the general emotional state (e.g., happy or sad); to show liking or dislike in everyday life situations or to signal a possible source of danger. Therefore, it is not surprising that humans are remarkably good at monitoring and detecting subtle changes in emotional expressions.

To be able to efficiently monitor emotional expressions they must be continuously attended to and memorized. In accordance with this, extensive research in recent years provided evidence that emotional facial expression can capture attention (Eastwood, Smilek, & Merikle, 2003; Ohman, Flykt, & Esteves, 2001) and thus will be processed efficiently even in the presence of distractors (Lucas & Vuilleumier, 2008) or in cases of poor visibility (Lee, Dolan, & Critchley, 2008). Surprisingly, however, visual short-term memory for facial emotions received far less attention. To date there has been no study that was aimed at investigating how efficiently humans can

store facial emotional information in visual short-term memory.

In contrast to the continuously changing emotional expression there are facial attributes—such as identity or gender—that on the short and intermediate timescale are invariant (Calder & Young, 2005; Haxby, Hoffman, & Gobbini, 2000). Invariant facial attributes do not require constant online monitoring during social interaction. After registering a person's identity at the beginning of a social encounter there is little need to monitor it further. Indeed, consistent with the latter point, one study showed that a remarkable 60% of participants failed to realize that a stranger they had began a conversation with was switched with another person after a brief staged separation during the social encounter (Simons & Levin, 1998). Further-more, it was shown that the processing of changeable and invariant facial attributes takes place on specialized, to some extent independent functional processing routes (Calder & Young, 2005; Haxby et al., 2000). Functional neuroimaging results suggest that facial identity might be processed primarily in the inferior occipito-temporal regions, including the fusiform face area (Haxby et al., 2001; Kanwisher, McDermott, & Chun, 1997), whereas processing of the information related to emotional expressions involves the superior temporal cortical regions (Andrews & Ewbank, 2004; Hasselmo, Rolls, & Baylis, 1989; LoPresti et al., 2008; Narumoto, Okada, Sadato, Fukui, & Yonekura, 2001; Vuilleumier, Armony, Driver, & Dolan, 2001; Winston, Henson, Fine-Goulden, & Dolan, 2004). Based on these, it is reasonable to suppose that the functional and anatomical differences in the processing of changeable and invariant facial attributes might also be reflected in the short-term memory processes for these different attributes.

The goal of the present study was to investigate how efficiently humans can store facial emotional expressions in visual short-term memory. In addition, we also aimed at testing the prediction that short-term memory for information related to changeable facial emotional expressions might be more efficient than that related to invariant facial attributes, such as identity. Using a two-interval forced choice facial attribute discrimination task we measured how increasing the delay between the subsequently presented face stimuli affects facial emotion and facial identity discrimination. The logic of our approach was as follows: if there is a high-fidelity short-term memory capacity for a facial attribute then observers' discrimination performance should be just as good when the faces are separated by several seconds as when the delay between the two faces is very short (1 s). However, if part of the information about facial attributes used for the discrimination is lost during the process of memory encoding, maintenance or recall then increasing the delay between the faces to be compared should impair discrimination performance.

Experiment 1

Previous research investigating short-term memory for basic visual dimensions (e.g., spatial frequency and orientation) using delayed discrimination tasks (Magnussen, 2000; Magnussen, Idås, & Myhre, 1998; Reinvang, Magnussen, Greenlee, & Larsson, 1998) found a significant increase in reaction times (RT) at delays longer than 3 s as compared to shorter delays. It was proposed that increased RTs at longer delays might reflect the involvement of memory encoding and retrieval processes, which are absent at delays shorter than 3 s. To test whether increasing the delay leads to longer RTs also in the case of delayed facial emotion discrimination we performed a pilot experiment. The results revealed that in delayed facial emotional discrimination tasks—similarly to discrimination of basic visual dimensions—there is a significant increase in RTs when the faces are separated by more than 3 s. Furthermore, it was also found that RTs saturate at 6-s delay, since no further increase in RTs was observed at delays longer than 6 s.

Based on these pilot results, in the main experiments aimed at testing the ability to store facial emotional expressions and facial identity in visual short-term memory, we compared participants' discrimination performance when the two face stimuli to be compared—the sample and test face image—were separated by 1 s (SHORT ISI) to that when the delay was 6 s (LONG ISI; Figure 1a).

Methods

Ten subjects (6 females, mean age: 24 years) gave their informed and written consent to participate in Experiment 1. Three of them also participated in a pilot experiment. None of them had any history of neurological or ophthalmologic diseases and all had normal or corrected-to-normal visual acuity.

Stimuli

Stimuli were front view images of female faces with gradually changing facial attributes of happiness, fear, and identity. Faces were cropped and covered with a circular mask. Images of two females (Female 1 and 2) were used for creating stimuli for the emotion discrimination, while for the identity discrimination task they were paired with two additional females (Female 3 and 4), yielding two different sets of images for all discrimination conditions. Test stimuli of varying emotional intensity were generated with a morphing algorithm (Winmorph 3.01; Kovács et al., 2006; Kovács, Zimmer, Harza, Antal, & Vidnyánszky, 2005; Kovács, Zimmer, Harza, & Vidnyánszky, 2007) by
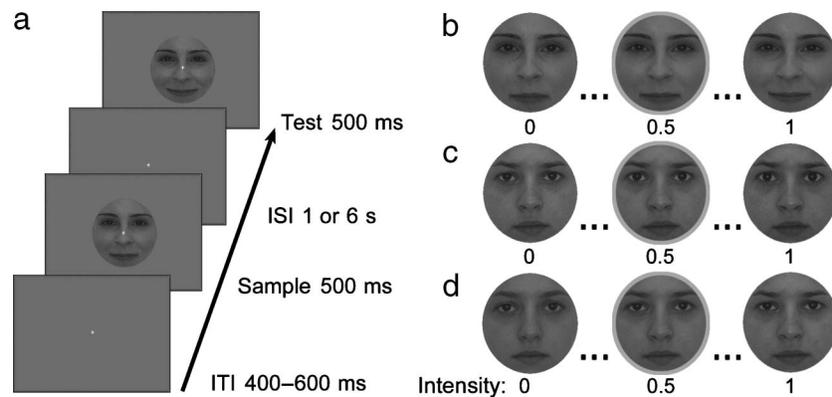
Figure 1. Experimental design and sample morphed face sets used in Experiment 1. (a) Stimulus sequence showing the happiness discrimination task. Stimulus sequence was similar for Experiments 2 and 3. Exemplar (b) happy, (c) fearful, and (d) identity morphed face sets used in Experiment 1. Each face pair consisted of the midpoint face—indicated by gray circles—and one of eight predefined stimuli. 0 and 1 show the typical two extremes while the other six face stimuli were evenly distributed in between. The morph continua used were assigned to the [0 1] interval for analysis and display purposes.

pairing a neutral and a happy/fearful picture of the same facial identity (Female 1 and 2), creating two sets of intermediate images. For the identity discrimination condition, two identity morph lines were created by morphing neutral images of two facial identities. As reference identity Female 1 and 2 were chosen, which were also used to create the morphed stimuli for the emotion discrimination task (Figures 1b–1d). Each set was composed of 101 images, 0% (neutral/Female 3 or 4) and 100% (happy/fearful/Female 1 or 2) being the original two faces. Stimuli (8 deg) were presented centrally (viewing distance of 60 cm) on a uniform gray background. Emotion and identity discrimination were measured by a two-interval forced choice procedure using the method of constant stimuli. In the emotion discrimination task, subjects were asked to report which of the two successively presented faces, termed sample and test, showed stronger facial emotional expressions (happy or fearful). In the identity discrimination task, subjects were required to report whether the test or the sample face resembled more to the reference identity. Subjects indicated their choice by pressing either button 1 or 2. Two interstimulus intervals (ISI) were used for testing: a short 1-s (SHORT ISI) and a long 6-s (LONG ISI) delay.

In each emotion discrimination trial, one of the face images was the midpoint image of the emotion morph line, corresponding to 50% happy/fearful emotional expression strength, while the other face image was chosen randomly from a continuum of eight predefined images of different emotional strength (Figures 1b and 1c). In the case of identity discrimination trials, one of the images was a face with 75% reference identity strength from the identity morph line. The other image was chosen randomly from a set of eight predefined images from the respective morph line, ranging from 50 to 100% reference identity strength. (Figure 1d). The rationale behind choosing the 75% instead of the 50% reference identity as the midpoint image for identity discrimination was to

have test faces that clearly exhibit the reference identity as pilot experiments revealed that using test stimuli with uncertain identity information leads to much poorer and noisier identity discrimination performance in our experimental paradigm.

The used continua for each attribute were determined individually in a practice session prior to the experiment. Each continuum was assigned to the [0 1] interval—0 and 1 representing the two extremes—for display and analysis purposes.

Procedures

A trial consisted of a 400–600 ms blank intertrial interval (ITI) followed by 500-ms presentation of the sample face, then either a short or a long delay with only the fixation cross present, finally 500 ms of the test face (Figure 1a). The two faces of the pair were randomly assigned to sample or test. Subjects initiated the trials by pressing one of the response buttons. In the identity condition the two reference faces of the two identity morph lines were presented for 5 s at the beginning of each block. The different facial attribute and ISI conditions were presented in separate blocks, their order being randomized across subjects. Each subject completed three of the 64-trial blocks, yielding 192 trials per condition for the experiment and underwent a separate training session prior to the experiment.

In a pilot emotional expression (happiness) discrimination experiment four different ISIs (1, 3, 6, 9 s) were used. Otherwise the experimental procedure was identical to the main experiment.

Data analysis

Analysis was performed on fitted Weibull psychometric functions (Wichmann & Hill, 2001). Performance was

assessed by computing just noticeable differences (JNDs, the smallest morph difference required to perform the discrimination task reliably), by subtracting the morph intensity needed to achieve 25% performance from that needed for 75% performance and dividing by two. JNDs have been used as a reliable measure of sensitivity (Lee & Harris, 1996). Reaction times were calculated as the average of the reaction times for stimuli yielding 25% and 75% performances. Single RTs longer than 2.5 s were excluded from further analysis. All measurements were entered into a 3×2 repeated measures ANOVA with attribute (happy vs. fear vs. identity) and ISI (SHORT vs. LONG) as within-subject factors. Tukey HSD tests were used for post-hoc comparisons.

Results

The pilot experiment revealed that increasing the delay between the face pairs leads to longer reaction times. However, the effect saturated at 6 s since no further increment in RT was found at the delay of 9 s compared to 6 s (Figure 2a). ANOVA showed a significant main effect of ISI ($F_{(3,6)} = 62.26$, $p < 0.0001$) and post-hoc tests revealed significant difference in all comparisons with the exception of the 6 s vs. 9 s ISI contrast ($p = 0.0025$, $p = 0.014$, $p = 0.78$ for 1 vs. 3, 3 vs. 6, and 6 vs. 9 s delays, respectively). Contrary to the RT results, however, participants' emotion (happiness) discrimination performance was not affected by the ISI (main effect of ISI: $F_{(3,6)} = 0.18$, $p = 0.90$).

In the main experiment, observers performed delayed discrimination of three different facial attributes: happiness, fear, and identity. In accordance with the result of the pilot experiment, in all three discrimination conditions reaction times were longer by approximately 150–200 ms in the LONG ISI (6 s) than in the SHORT ISI (1 s) conditions (Figure 2b), providing support for the involvement of short-term memory processes in delayed facial attribute discrimination in the case of LONG ISI conditions. ANOVA performed on the RT data showed a significant main effect of ISI (SHORT vs. LONG ISI, $F_{(1,9)} = 54.12$, $p < 0.0001$), while there was no main effect of attributes (happiness vs. fear vs. identity, $F_{(2,18)} = 2.15$, $p = 0.146$) and no interaction between these variables ($F_{(2,18)} = 0.022$, $p = 0.978$).

In contrast to the RT results, increasing the delay between the face images to be compared had only a small effect on observers' performance in the identity discrimination condition but not in the two facial emotion discrimination conditions (Figures 3a–3c, see also Figure 3d for the JND values used in the analysis). ANOVA showed that the main effect of ISI (SHORT vs. LONG ISI, $F_{(1,9)} = 4.24$, $p = 0.069$), the main effect of attributes (happiness vs. fear vs. identity, $F_{(2,18)} = 3.29$, $p = 0.061$), and the interaction between these variables

($F_{(2,18)} = 3.29$, $p = 0.061$) all failed to reach the significance level. In the case of facial identity discrimination, post-hoc analysis showed a non-significant trend of decreased performance in the LONG as compared to the SHORT ISI condition (post-hoc: $p = 0.07$). On the other hand, discrimination of facial emotions was not affected by the ISI (post-hoc: $p = 0.999$ and $p = 0.998$ for happiness and fear, respectively). These results suggest that fine-grained information about facial emotions can be stored with high precision, without any loss in visual short-term memory.

Experiment 2

The results obtained in Experiment 1 reflect visual short-term memory abilities in case of familiar face stimuli and in extensively practiced task conditions
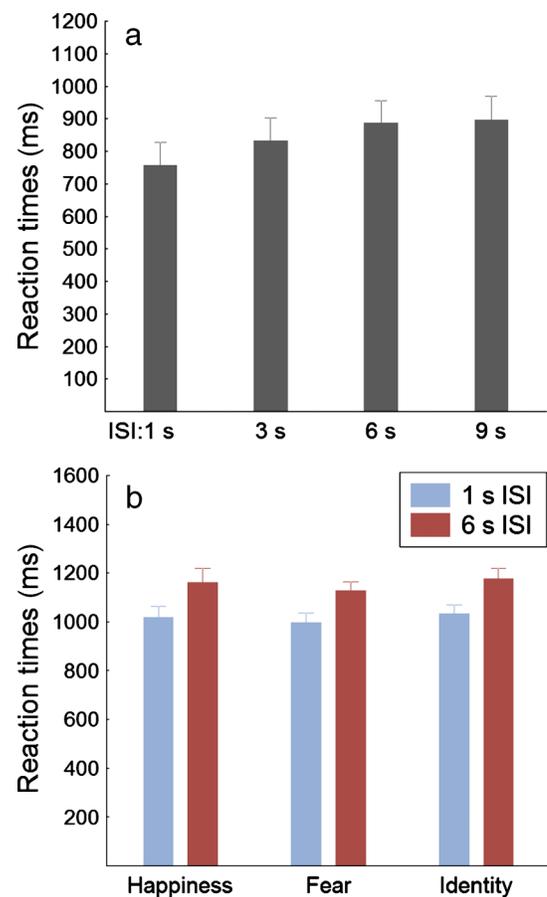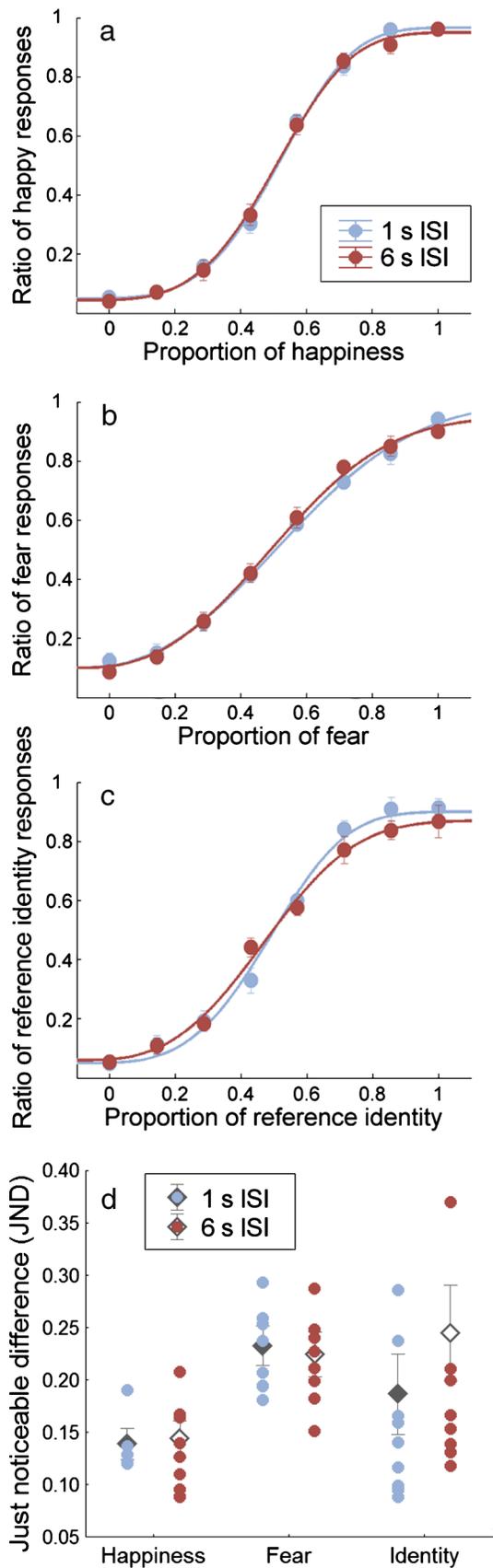

Figure 2. Reaction times for delayed emotion (happiness) discrimination measured during (a) a pilot experiment and (b) Experiment 1. Mean RTs were calculated from trials with face pairs yielding 25% and 75% performances. Error bars indicate ±*SEM* (*N* = 3 and 10 for the pilot experiment and Experiment 1, respectively).

Figure 3. Effect of ISI on the performance of facial emotion and identity discrimination. Weibull psychometric functions fit onto (a) happiness, (b) fear, and (c) identity discrimination performances. Introducing a 6-s delay (brown line) between sample and test faces had no effect on emotion discrimination and did not impair identity discrimination performance significantly, compared to the short 1-s interstimulus interval (ISI) condition (blue line). The *x*-axis denotes morph intensities of the constant stimuli. (d) Just noticeable differences (JNDs) obtained in Experiment 1. Diamonds represent mean JNDs in each condition while circles indicate individual data for short (blue) and long (brown) ISIs. Error bars indicate $\pm SEM$ ($N = 10$).

(observers performed 3 blocks of 64 trials for each attribute and ISI). Therefore, a second experiment was performed to test whether high-precision visual short-term memory for facial emotions also extends to situations where the faces and the delayed discrimination task are novel to the observers. In this experiment, each participant ($N = 160$) performed only two trials of delayed emotion (happiness) discrimination and another two trials of delayed identity discrimination. For half of the participants the sample and test faces were separated by 1 s (SHORT ISI) while for the other half of participants the ISI was 10 s (LONG ISI).

Importantly, this experiment also allowed us to test whether in our task conditions delayed facial attribute discrimination is based on the perceptual memory representation of the sample stimulus (Magnussen, 2000; Pasternak & Greenlee, 2005) or it is based on the representation of the whole range of the task-relevant feature information that builds up during the course of the experiment, as suggested by the Lages and Treisman's (1998) criterion-setting theory. This is because in Experiment 2 observers performed only two emotion and two identity discrimination trials with novel faces and thus the involvement of criterion-setting processes proposed by Lages and Treisman can be excluded.

Methods

Altogether 206 participants took part in Experiment 2. They were screened according to their performance and were excluded from further analysis if overall performance did not reach 60 percent, yielding 160 subjects altogether (78 females, mean age: 22 years), 80 for each ISI condition.

Stimuli and procedures

In Experiment 2 only two facial attributes were tested: happiness and identity. The same face sets were used as in Experiment 1 but only one of them was presented during the experiment, while the other set was used in the short practice session prior to the experiment, during which

subjects familiarized themselves with the task. In each trial similarly to Experiment 1, one image was the midpoint face (see Experiment 1 Methods section for details) while the other image was one of two predefined test stimuli. Thus only two face pairs were used in both identity and emotion discrimination conditions: in one face pair the emotion/identity intensity difference between the images was larger, resulting in good discrimination performance, whereas in the other face pair the difference was more subtle, leading to less efficient discrimination. Subjects performed a single discrimination for each of the two face pairs (Magnussen, Greenlee, Aslaksen, & Kildebo, 2003) of the two facial attribute conditions. The identity reference face was presented before the identity block. Subjects initiated the start of the block after memorizing the identity reference face by pressing a button. Stimulus sequence was identical to Experiment 1. Subjects were randomly assigned an ISI (either short or long) and a starting stimulus out of the two test faces and shown the happy and identity stimuli in a counterbalanced fashion. Presentation order of the two face pairs was also counterbalanced across subjects. Every other parameter and the task instructions were identical to Experiment 1.

Data analysis

For analyzing Experiment 2 the individual data points were insufficient for a proper fit so to test whether the distributions were different, we applied χ^2 tests to performance data obtained by pooling correct and incorrect responses for trials with face pairs having small and large intensity differences separately (Magnussen et al., 2003). Reaction times were averaged over face pairs. Similarly to Experiment 1, single trial RTs exceeding 2.5 s were excluded from further analysis leaving unequal number of RT measurements per conditions ($N = 74$ and $N = 60$ for SHORT and LONG ISI conditions). RT data were analyzed with a 2×2 repeated measures ANOVA with attribute (happiness vs. identity) as within-subject and ISI (SHORT vs. LONG) as between-subject factors.

Results

Reaction times, similarly to that in Experiment 1, were longer in the LONG ISI than in the SHORT ISI condition by 180–240 ms. Moreover, subjects were faster in responding in the happy than in the identity discrimination condition (Figure 4a). Statistical analysis revealed a significant main effect of ISI ($F_{(1,132)} = 13.09$, $p = 0.0004$) and a significant main effect of attribute ($F_{(1,132)} = 11.03$, $p = 0.001$).

The results also revealed that subjects' emotion and identity discrimination performance was not affected by the delay between the face stimuli to be compared, even though the faces were novel (Figure 4b). There was no

significant difference between the SHORT ISI and LONG ISI conditions in the case of happiness discrimination performance ($\chi^2_{(1,N=160)} = 0.493$, $p = 0.482$ and $\chi^2_{(1,N=160)} = 0.00$, $p = 1.00$ for the image pair with large and small differences, respectively) as well as in the case of identity discrimination performance ($\chi^2_{(1,N=160)} = 0.028$, $p = 0.868$ and $\chi^2_{(1,N=160)} = 0.028$, $p = 0.868$ for the image pair with
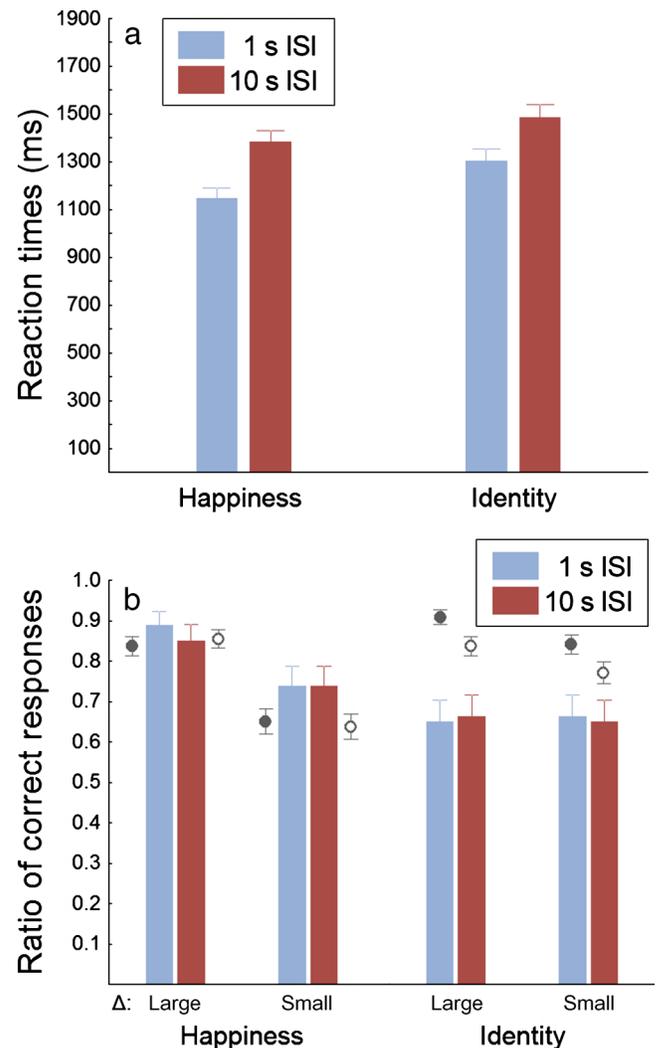

Figure 4. Reaction times and discrimination performance in Experiment 2. (a) There was a significant RT increase in the LONG compared to the SHORT ISI condition in the case of both attributes (Valid number of measurements: $N = 74$ and $N = 60$ for the SHORT and LONG ISI conditions, respectively). (b) Performance did not show any significant drop from 1 s to 10 s ISI (blue and brown bars, respectively) in either discrimination conditions, neither for face pairs with large nor with small difference. For comparison of the overall discrimination performance in Experiments 1 and 2, gray circles represent the mean performance in Experiment 1 for the corresponding face pairs in the short (filled circles) and long (circles) ISI conditions. Error bars indicate $\pm SEM$ ($N = 160$ and 10 for Experiments 1 and 2, respectively).

large and small differences, respectively). These results suggest that humans can store fine-grained information related to facial emotions and identity without loss in visual short-term memory even when the faces and the task are both novel.

Since the face images used in Experiment 2 were selected from the same image set that was used in Experiment 1, it is possible to compare the overall discrimination performance across the two experiments. As shown in Figure 4b in Experiment 2 discrimination of facial emotions in case when both the task and the faces are novel was just as good as that found after several hours of practice in Experiment 1. On the other hand, overall identity discrimination performance in Experiment 2 was worse than in Experiment 1, suggesting that practice and familiarity of faces affect performance in the facial identity discrimination task but not in the facial emotion discrimination task.

Experiment 3

To confirm that emotion discrimination in our short-term memory paradigm involved high-level processing of facial emotional attributes, we performed an fMRI experiment. Previous studies have shown that increased fMRI responses in the posterior superior temporal sulcus (pSTS) during tasks requiring perceptual responses to facial emotions compared to those to facial identity can be considered as a marker for processing of emotion-related facial information (Andrews & Ewbank, 2004; Hasselmo et al., 1989; LoPresti et al., 2008; Narumoto et al., 2001; Vuilleumier et al., 2001; Winston et al., 2004). Therefore, we conducted an fMRI experiment in which we compared fMRI responses measured during delayed emotion (happiness) discrimination to that obtained during identity discrimination. Importantly, the same sets of morphed face stimuli were used both in the emotion and in the identity discrimination tasks with slightly different exemplars in the two conditions. Thus the major difference between the two conditions was the task instruction (see Experimental procedures for details). We predicted that if the delayed emotion discrimination task used in the present study—requiring discrimination of very subtle differences in facial emotional expression—involved high-level processing of facial emotional attributes then pSTS should be more active in the emotion discrimination condition as compared to the identity discrimination condition. Furthermore, finding enhanced fMRI responses in brain areas involved in emotion processing would also exclude the possibility that discrimination of fine-grained emotional information in our emotion discrimination condition is based solely on matching low-level features (e.g., orientation, spatial frequency) of the face stimuli with different strength of emotional expressions.

Methods

Thirteen subjects participated in this experiment. fMRI and concurrent psychophysical data of three participants were excluded due to excessive head movement in the scanner, leaving a total of ten right-handed subjects (6 females, mean age: 24 years).

Stimuli

Like in Experiment 2, we tested two facial attributes: happiness and identity. We used the same face sets for both tasks to ensure that the physical properties of the stimuli were the same (i.e., there were no stimulus confounds) and the conditions only differed in which face attribute subjects had to attend to make the discrimination. To do this we created face sets where both facial attributes changed gradually by morphing a neutral face of one facial identity with the happy face of another identity and visa versa: the happy face of the first identity with the neutral face of the second to minimize correlation between the two attributes (Figure 5a). There were two composite face sets in the experiment: one female and one male. In the main experiment, six (3 + 3) face pairs yielding 75% performance were used from each composite face set, selected based on the performance in the practice session. The chosen pairs slightly differed in the two conditions—e.g., 48 vs. 60% and 42 vs. 60% for the emotion and the identity discrimination, respectively— since subjects needed bigger differences in the identity condition to achieve 75% performance. The emotion intensity difference between conditions averaged across subjects and runs turned out to be 6%, the emotion discrimination condition displaying the happier stimuli. Trials of emotion and identity discrimination tasks were presented within a block in an optimized pseudorandom order to maximize separability of the different tasks. For each subject the same trial sequence was used.

Visual stimuli were projected onto a translucent screen located at the back of the scanner bore using a Panasonic PT-D3500E DLP projector (Matsushita Electric Industrial, Osaka, Japan) at a refresh rate of 75 Hz. Stimuli were viewed through a mirror attached to the head coil with a viewing distance of 58 cm. Head motion was minimized using foam padding.

Procedure

The task remained identical to that of Experiments 1 and 2 but the experimental paradigm was slightly altered to be better suited for fMRI. A trial began with a task cue (0.5 deg) appearing just above fixation for 500 ms being either 'E' for emotion or 'I' for identity discrimination. Following a blank fixation of 1530 ms the faces appeared successively for 300 ms separated by a long ISI of varied length. The ITI was fixed in 3.5 s, which also served as the
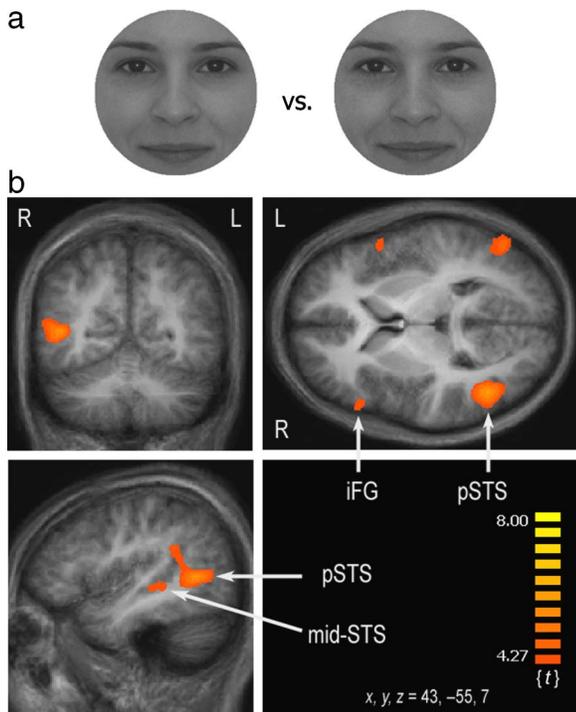
Figure 5. Stimuli and results of Experiment 3. (a) An exemplar face pair taken from the female composite face set, which differs slightly along both the facial identity and emotion axis. (b) fMRI responses for sample faces. Emotion vs. identity contrast revealed significantly stronger fMRI responses during emotion than identity discrimination within bilateral superior temporal sulcus (STS; two clusters: posterior and mid) and bilateral inferior frontal gyrus (iFG). Coordinates are given in Talairach space; regional labels were derived using the Talairach Daemon (Lancaster et al., 2000) and the AAL atlas provided with MRIcro (Rorden & Brett, 2000).

response window. The ISI varied between 5 and 8 s in steps of 1 s to provide a temporal jitter. Subjects performed 24 trials for each of the seven functional runs (12 trials of emotion and 12 trials of identity discrimination), for a total of 168 trials.

Before scanning, subjects were given a separate practice session where they familiarized themselves with the task and the image pairs with approximately 75% correct performance were determined. Eye movements of five randomly chosen subjects were recorded in this session by an iView XTM HI-Speed eye tracker (Sensomotoric Instruments, Berlin, Germany) at a sampling rate of 240 Hz. In all experiments the stimulus presentation was controlled by MATLAB 7.1. (The MathWorks, Natick, MA) using the Psychtoolbox 2.54 (Brainard, 1997; Pelli, 1997).

Behavioral data analysis

Responses and reaction times were collected for each trial during the practice and scanning sessions to ensure subjects were performing the task as instructed. Accuracy and mean RTs were analyzed with paired t-tests.

Analysis of eye tracking data

Eye-gaze direction was assessed using a summary statistic approach. Trials were binned based on facial attribute (emotion vs. identity) and task phase (sample vs. test) and mean eye position (x and y values) was calculated for periods when the face stimulus was present on each trial. From each of the four eye-gaze direction data set, spatial maps of eye-gaze density were constructed and then averaged to get a mean map for comparison. Subsequently, each of these maps was compared with the mean map and difference images were computed. The root mean squares of the density difference values for these latter maps were entered into a 2×2 ANOVA (Winston et al., 2004).

fMRI imaging and analysis

Data acquisition

Data were collected at the MR Research Center of Szentágothai Knowledge Center (Semmelweis University, Budapest, Hungary) on a 3.0-Tesla Philips (Best, The Netherlands) Achieva scanner equipped with an eight-channel SENSE head coil. High-resolution anatomical images were acquired for each subject using a T1 weighted 3D TFE sequence yielding images with a $1 \times 1 \times 1$ mm resolution. Functional images were collected using 31 transversal slices (4-mm slice thickness with 3.5 mm $\times$ 3.5 mm in-plane resolution) with a non-interleaved acquisition order covering the whole brain with a BOLD-sensitive T2*-weighted echo-planar imaging sequence (TR = 2 s, TE = 30 ms, FA = 75°, FOV = 220 mm, 64 $\times$ 64 image matrix, 7 runs, duration of each run = 516 s).

Data analysis

Preprocessing and analysis of the imaging data were performed using BrainVoyager QX (v 1.910; Brain Innovation, Maastricht, The Netherlands). Anatomicals were coregistered to BOLD images and then transformed into standard Talairach space. BOLD images were corrected for differences in slice timing, realigned to the first image within a session for motion correction and low-frequency drifts were eliminated with a temporal high-pass filter (3 cycles per run). The images were then spatially smoothed using a 6-mm full-width half-maximum Gaussian filter and normalized into standard Talairach space. Based on the results of the motion correction algorithm runs with excessive head movements were excluded from further analysis leaving 10 subjects with 4–7 runs each.

Functional data analysis was done by applying a two-level mass univariate general linear model (GLM) for an event-related design. For the first-level GLM analysis, delta functions were constructed corresponding to the onset of

each event type (emotion vs. identity discrimination × sample vs. test face). These delta functions were convolved with a canonical hemodynamic response function (HRF) to create predictors for the subsequent GLM. Temporal derivatives of the HRFs were also added to the model to accommodate different delays of the BOLD response in the individual subjects. The resulting β weights of each current predictor served as input for the second-level whole-brain random-effects analysis, treating subjects as random factors. Linear contrasts pertaining to the main effects were calculated and the significance level to identify cluster activations was set at $p < 0.01$ with false discovery rate (FDR) correction with degrees of freedom $df_{(random)} = 9$.

Results

Subjects' accuracy during scanning was slightly better in the identity than in the emotion discrimination task (mean $\pm SEM$: 79.7 ± 1.4% and 83.0 ± 2.0% for emotion and identity tasks, respectively; $t_{(9)} = -2.72$, $p = 0.024$). Reaction times did not differ significantly across task conditions (mean $\pm SEM$: 831 ± 66 ms and 869 ± 71 ms for emotion and identity, respectively; $t_{(9)} = -1.49$, $p = 0.168$).

To assess the difference between the neural processing of the face stimuli in the emotion and identity discrimination tasks, we contrasted fMRI responses in the emotion discrimination trials with those in the identity trials. We found no brain regions where activation was higher in the identity compared to the emotion discrimination condition, neither during sample nor during test face processing. However, our analysis revealed significantly higher activations for the sample stimuli in the case of emotion compared to identity discrimination in the right posterior superior temporal sulcus (Br. 37, peak at $x, y, z = 43, -55, 7$; $t = 6.18$, $p < 0.01_{FDR}$, Figure 5b). This cluster of activation extended ventrally and rostrally along the superior temporal sulcus and dorsally and rostrally into the supramarginal gyrus (Br. 22, $x, y, z = 42, -28, 0$; $t = 4.43$; Br. 40, $x, y, z = 45, -42, 25$; $t = 4.87$, $p < 0.01_{FDR}$, centers of activation for mid-STS and supramarginal gyrus, respectively). Furthermore, we found five additional clusters with significantly stronger activations in: left superior temporal gyrus (Br. 37, $x, y, z = -51, -65, 7$; $t = 4.91$, $p < 0.01_{FDR}$), left superior temporal pole (Br. 38, $x, y, z = -45, 18, -14$; $t = 4.70$, $p < 0.01_{FDR}$), bilateral inferior frontal cortex: specifically in right inferior frontal gyrus (triangularis) (Br. 45, $x, y, z = 51, 26, 7$; $t = 4.73$) and in left inferior frontal gyrus (opercularis) (Br. 44, $x, y, z = -51, 14, 8$; $t = 4.53$ $p < 0.01_{FDR}$), and finally, in left insula (Br. 13, $x, y, z = -36, 8, 13$; $t = 4.65$ $p < 0.01_{FDR}$). This network of cortical areas showing higher fMRI responses in the emotion than in the identity task is in close correspondence with the results of earlier studies investigating processing of facial emotions. Interestingly, in the case of fMRI responses to the test face stimuli, even

though many of these cortical regions, including pSTS, showed higher activations in the emotion compared to the identity task these activation differences did not reach significance, which is in agreement with recent findings of LoPresti et al. (2008). Furthermore, our results did not show significantly higher amygdala activations in the emotion discrimination condition as compared to the identity discrimination condition. One explanation for the lack of enhanced amygdala activation in the emotion condition might be that in our fMRI experiment we used face images with positive emotions and subjects were required to judge which face was happier. This is supported by a recent meta-analysis of the activation of amygdala during processing of emotional stimuli by Costafreda, Brammer, David, and Fu (2008), where they found that there was a higher probability of amygdala activation: 1) for stimuli reflecting fear and disgust relative to happiness; 2) in the case of passive emotion processing relative to the case of active task instructions.

As overall intensity of emotional expressions of the face stimuli used in the emotion discrimination task was slightly higher (6%) than that in the identity task we carried out an analysis designed to test whether the small difference in emotional intensity of the face stimuli can explain the difference in strength of pSTS activation between the emotion and identity conditions. We divided the fMRI data obtained both from the emotion and the identity discrimination conditions separately into two median split subgroups based on emotion intensity of the face stimulus in the given trial. Thus, we were able to contrast the fMRI responses arising from trials where faces showed more intense emotional expression with trials where faces showed less emotional intensity separately for emotion and identity discrimination conditions. The difference in emotion intensity of the face stimuli was 13% in the case of the two subgroups of emotion discrimination trials and 17% in the case of identity discrimination trials; that is in both cases the intensity difference between the respective subgroups was larger than the difference in emotion intensity of face stimuli between the two task conditions (6%). The contrast failed to yield difference in the STS activations between the two subgroups in either task condition even at a significance level of $p < 0.01$ uncorrected. These results clearly show that the small difference in the emotional intensity of the face stimuli between the emotion and identity discrimination conditions cannot explain the higher STS activations found during emotion discrimination as compared to the identity discrimination.

Furthermore, since subjects' performance during scanning was slightly better in the identity discrimination condition than in the emotion discrimination we performed an additional analysis to exclude the possibility that the observed differences in fMRI responses between the two conditions are due to a difference in task difficulty. For this, we selected three runs from each subject in which accuracy for the two tasks was similar and reanalyzed the

fMRI data collected from these runs. Even though there was no significant difference between subjects' accuracy in the emotion and identity tasks in these runs (mean ±*SEM*: 82.2 ± 1.7% and 81.9 ± 2.0% for emotion and identity tasks, respectively; $t_{(9)} = 0.145$, $p = 0.889$), the emotion vs. identity contrast revealed the same clusters of increased fMRI responses as when all runs were analyzed; including a significantly higher activation during the emotion discrimination task in the right posterior STS (peak at *x, y, z* = 45, −52, 4; $t = 5.08$, $p < 0.03_{FDR}$). Thus, our fMRI results provide evidence that discrimination of fine-grained emotional information required in our experimental condition led to the activation of a cortical network that is known to be involved in processing of facial emotional expression.

Although we did not track eye position during scanning, it appears highly unlikely that the difference between the fMRI responses in the emotion and identity discrimination task could be explained by a difference in fixation patterns between the two tasks. Firstly, we recorded eye movements during the practice sessions prior to scanning for 5 subjects and the data revealed no significant differences between the facial attributes (emotion vs. identity, $F_{(1,4)} = 1.15$, $p = 0.343$) or the task phases (sample vs. test, $F_{(1,4)} = 0.452$, $p = 0.538$) and there was no interaction between these variables ($F_{(1,4)} = 0.040$, $p = 0.852$). These indicate that there was no systematic bias in eye-gaze direction induced by the different task demands (attend to emotion or identity). Secondly, in the whole-brain analysis of the fMRI data we found no significant differences in activations of the cortical areas known to be involved in programming and execution of eye movements (i.e., in the frontal eye field or parietal cortex; Pierrot-Deseilligny, Milea, & Müri, 2004) in response to emotion and identity discrimination tasks.

Discussion

The results of the present study provide the first behavioral evidence that the ability to compare facial emotional expressions (happiness and fear) of familiar as well as novel faces is not impaired when the emotion-related information has to be stored for several seconds in short-term memory; i.e., when the faces to be compared are separated by up to 10 s. Furthermore, it was also found that discrimination of facial emotions is just as good when the observers perform the task only twice with novel faces as it is after extensive practice. Importantly, high-fidelity short-term memory found in Experiment 2 cannot be accounted for by the criterion-setting theory proposed by Lages and Treisman (1998) to explain results of delayed discrimination of basic visual dimensions (Magnussen, 2000). According to this theory, in delayed discrimination tasks observers' decision on a given trial is based on the

representation of the whole range of the task-relevant feature information that builds up during the course of the experiment, rather than based on the perceptual memory representation of the sample stimulus. Since in our experiment observers performed only two emotion and two identity discrimination trials with novel faces the involvement of criterion-setting processes proposed by Lages and Treisman (1998) can be excluded. Thus our results provide direct evidence that humans can store with high precision fine-grained information related to facial emotions and identity in short-term memory.

Based on the known functional and anatomical differences in the processing of changeable and invariant facial attributes (Calder & Young, 2005; Haxby et al., 2000), we assumed that short-term memory for facial emotions might be more efficient than that for facial identity. However, our results failed to provide support for such a functional dissociation in short-term memory processes for changeable and invariant facial attributes. We found that humans are also able to store fine-grained information related to facial identity without loss in visual short-term memory. There was only a small, non-significant decrease in identity discrimination performance at longer delay in Experiment 1 where observers performed several blocks of identity discrimination tasks with the same face stimuli. A possible explanation for this is that learning processes might affect identity discrimination and the ability to store information related to facial identity in short-term memory differently. Taken together, our results suggest that even though processing of emotion and identity is accomplished by specialized, to some extent anatomically segregated brain regions—as it was shown previously (Haxby et al., 2000) and also supported by our control fMRI experiment—short-term memory for both of these attributes is highly efficient.

How and where in the human brain visual information related to facial identity and facial emotions is represented during short-term memory maintenance is still an open question. Several previous studies suggested (Druzgal & D'Esposito, 2003; Postle, Druzgal, & D'Esposito, 2003; Yoon, Curtis, & D'Esposito, 2006) that the visual association cortex, including the fusiform face area (Kanwisher et al., 1997) and the lateral prefrontal cortex is involved in active maintenance of facial information during memory delays. It was also proposed that these two regions have different functions in active maintenance: the lateral prefrontal cortex codes for abstract mnemonic information, while sensory areas represent specific features of the memoranda (Yoon et al., 2006). However, a recent fMRI study (LoPresti et al., 2008) investigating short-term memory processes for facial identity and emotions failed to find sustained activity in the fusiform face area and in the superior temporal cortex; i.e., in the regions of the visual association cortex specialized for the processing of these attributes. Instead, they showed overlapping delay related activations in case of facial identity and facial emotions in the orbitofrontal cortex, amygdala,

and hippocampus and suggested that this network of brain areas might be critical for actively maintaining and binding together information related to facial identity and emotion in short-term memory. Further research is needed to explain these discrepancies and to uncover how and where in the human brain fine-grained information related to facial identity and facial emotions is represented during maintenance in visual short-term memory.

From an ecological point of view, our results showing highly efficient short-term memory for both facial identity and facial emotional expression might appear rather surprising. As we reasoned above, in contrast to the continuously changing emotional expression, facial identity is invariant on the short and intermediate timescale (Calder & Young, 2005; Haxby et al., 2000). Thus, in everyday life situations there is little need to store fine-grained identity related information in short-term memory. One possible explanation that might help to reconcile this apparent contradiction is that even though humans possess a high fidelity short-term memory system for both emotions and identity, only facial emotions are monitored, memorized, and compared continuously during social intercourse. This is supported by previous findings showing that facial emotions can automatically capture attention and receive prioritized processing (Eastwood et al., 2003; Lee et al., 2008; Ohman et al., 2001). Facial identity, on the other hand, might be attended, memorized, and monitored only in cases when identity changes are expected to take place. Assuming that identity information is not automatically monitored and memorized might help explain previous findings, showing that humans are surprisingly bad at noticing identity changes; such as when a stranger they had began a conversation with was switched with another person after a brief staged separation during a social encounter (Simons & Levin, 1998).

Conclusions

To conclude, the findings of the present study show that humans possess flawless visual short-term memory for facial emotional expressions and facial identity. Such high-fidelity short-term memory is inevitable for the ability to efficiently monitor emotional expressions and it is tempting to propose that impairment of such high-precision short-term memory storage of emotional information might be one of the possible causes of the deficits of emotional processing found in psychiatric disorders including autism and schizophrenia (Humphreys, Minshew, Leonard, & Behrmann, 2007; Kosmidis et al., 2007; Sasson et al., 2007).

Acknowledgments

Commercial relationships: none.
Corresponding authors: Éva M. Bankó and Zoltán Vidnyánszky.
Email: banko.eva@itk.ppke.hu; vidnyanszky@digitus.itk.ppke.hu.
Address: 50/a. Práter u., Budapest 1083, Hungary.

References

Andrews, T. J., & Ewbank, M. P. (2004). Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *Neuroimage, 23,* 905–913. [PubMed]

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436. [PubMed]

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews, Neuroscience, 6,* 641–651. [PubMed]

Costafreda, S. G., Brammer, M. J., David, A. S., & Fu, C. H. (2008). Predictors of amygdala activation during the processing of emotional stimuli: A meta-analysis of 385 PET and fMRI studies. *Brain Research Reviews, 58,* 57–70. [PubMed]

Druzgal, T. J., & D'Esposito, M. (2003). Dissecting contributions of prefrontal cortex and fusiform face area to face working memory. *Journal of Cognitive Neuroscience, 15,* 771–784. [PubMed]

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2003). Negative facial expression captures attention and disrupts performance. *Perception & Psychophysics, 65,* 352–358. [PubMed]

Ekman, P. (1973). *Darwin and facial expression: A century of research in review.* New York: Academic Press.

Fridlund, A. J. (1994). *Human facial expression.* New York: Academic Press.

Hasselmo, M. E., Rolls, E. T., & Baylis, G. C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioural Brain Research, 32,* 203–218. [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]

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science, 293,* 2425–2430. [PubMed]

Humphreys, K., Minshew, N., Leonard, G. L., & Behrmann, M. (2007). A fine-grained analysis of facial expression processing in high-functioning adults with autism. *Neuropsychologia, 45,* 685–695. [PubMed]

Izard, C. E. (1977). *Human emotions* (1st ed.). New York: Plenum.

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

Kosmidis, M. H., Bozikas, V. P., Giannakou, M., Anezoulaki, D., Fantie, B. D., & Karavatos, A. (2007). Impaired emotion perception in schizophre-nia: A differential deficit. *Psychiatry Research, 149,* 279–284. [PubMed]

Kovács, G., Zimmer, M., Bankó, É., Harza, I., Antal, A., & Vidnyánszky, Z. (2006). Electrophysiological correlates of visual adaptation to faces and body parts in humans. *Cerebral Cortex, 16,* 742–753. [PubMed] [Article]

Kovács, G., Zimmer, M., Harza, I., Antal, A., & Vidnyánszky, Z. (2005). Position-specificity of facial adaptation. *Neuroreport, 16,* 1945–1949. [PubMed]

Kovács, G., Zimmer, M., Harza, I., & Vidnyánszky, Z. (2007). Adaptation duration affects the spatial selec-tivity of facial aftereffects. *Vision Research, 47,* 3141–3149. [PubMed]

Lages, M., & Treisman, M. (1998). Spatial frequency discrimination: Visual long-term memory or criterion setting? *Vision Research, 38,* 557–572. [PubMed]

Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., et al. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping, 10,* 120–131. [PubMed]

Lee, B., & Harris, J. (1996). Contrast transfer character-istics of visual short-term memory. *Vision Research, 36,* 2159–2166. [PubMed]

Lee, T. W., Dolan, R. J., & Critchley, H. D. (2008). Controlling emotional expression: Behavioral and neural correlates of nonimitative emotional responses. *Cerebral Cortex, 18,* 104–113. [PubMed] [Article]

LoPresti, M. L., Schon, K., Tricarico, M. D., Swisher, J. D., Celone, K. A., & Stern, C. E. (2008). Working memory for social cues recruits orbito-frontal cortex and amygdala: A functional magnetic resonance imaging study of delayed matching to sample for emotional expressions. *Journal of Neuro-science, 28,* 3718–3728. [PubMed] [Article]

Lucas, N., & Vuilleumier, P. (2008). Effects of emotional and non-emotional cues on visual search in neglect patients: Evidence for distinct sources of attentional guidance. *Neuropsychologia, 46,* 1401–1414. [PubMed]

Magnussen, S. (2000). Low-level memory processes in vision. *Trends in Neurosciences, 23,* 247–251. [PubMed]

Magnussen, S., Greenlee, M. W., Aslaksen, P. M., & Kildebo, O. O. (2003). High-fidelity perceptual long-term memory revisited and confirmed. *Psychological Science, 14,* 74–76. [PubMed]

Magnussen, S., Idås, E., & Myhre, S. H. (1998). Representation of orientation and spatial frequency in perception and memory: A choice reaction-time analysis. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 707–718. [PubMed]

Narumoto, J., Okada, T., Sadato, N., Fukui, K., & Yonekura, Y. (2001). Attention to emotion modulates fMRI activity in human right superior temporal sulcus. *Brain Research: Cognitive Brain Research, 12,* 225–231. [PubMed]

Ohman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General, 130,* 466–478. [PubMed]

Pasternak, T., & Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nature Reviews, Neuroscience, 6,* 97–107. [PubMed]

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442. [PubMed]

Pierrot-Deseilligny, C., Milea, D., & Müri, R. M. (2004). Eye movement control by the cerebral cortex. *Current Opinion in Neurology, 17,* 17–25. [PubMed]

Postle, B. R., Druzgal, T. J., & D'Esposito, M. (2003). Seeking the neural substrates of visual working memory storage. *Cortex, 39,* 927–946. [PubMed]

Reinvang, I., Magnussen, S., Greenlee, M. W., & Larsson, P. G. (1998). Electrophysiological localization of brain regions involved in perceptual memory. *Exper-imental Brain Research, 123,* 481–484. [PubMed]

Rorden, C., & Brett, M. (2000). Stereotaxic display of brain lesions. *Behavioural Neurology, 12,* 191–200. [PubMed]

Sasson, N., Tsuchiya, N., Hurley, R., Couture, S. M., Penn, D. L., Adolphs, R., et al. (2007). Orienting to social stimuli differentiates social cognitive impair-ment in autism and schizophrenia. *Neuropsychologia, 45,* 2580–2588. [PubMed] [Article]

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

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: An event-related fMRI study. *Neuron, 30,* 829–841. [PubMed] [Article]

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics, 63,* 1293–1313. [PubMed]

Winston, J. S., Henson, R. N., Fine-Goulden, M. R., & Dolan, R. J. (2004). fMRI-adaptation reveals dissoci-able neural representations of identity and expression in face perception. *Journal of Neurophysiology, 92,* 1830–1839. [PubMed] [Article]

Yoon, J. H., Curtis, C. E., & D'Esposito, M. (2006). Differential effects of distraction during working memory on delay-period activity in the prefrontal cortex and the visual association cortex. *Neuroimage, 29,* 1117–1126. [PubMed]