Retention Interval Affects Visual Short-Term Memory Encoding

Éva M. Bankó1,2 and Zoltán Vidnyánszky1,2,3
1Neurobionics Research Group, Hungarian Academy of Sciences—Pázmány Péter Catholic University—Semmelweis University; 2Faculty of Information Technology, Pázmány Péter Catholic University; and 3MR Research Center, Szentágothai J. Knowledge Center—Semmelweis University, Budapest, Hungary

Submitted 22 September 2009; accepted in final form 11 January 2010

Bankó ÉM, Vidnyánszky Z. Retention interval affects visual short-term memory encoding. J Neurophysiol 103: 1425–1430, 2010. First published January 13, 2010; doi:10.1152/jn.00868.2009. Humans can efficiently store fine-detailed facial emotional information in visual short-term memory for several seconds. However, an unresolved question is whether the same neural mechanisms underlie high-fidelity short-term memory for emotional expressions at different retention intervals. Here we show that retention interval affects the neural processes of short-term memory encoding using a delayed facial emotion discrimination task. The early sensory P100 component of the event-related potentials (ERP) was larger in the 1-s interstimulus interval (ISI) condition than in the 6-s ISI condition, whereas the face-specific N170 component was larger in the longer ISI condition. Furthermore, the memory-related late P3b component of the ERP responses was also modulated by retention interval: it was reduced in the 1-s ISI as compared with the 6-s condition. The present findings cannot be explained based on differences in sensory processing demands or overall task difficulty because there was no difference in the stimulus information and subjects’ performance between the two different ISI conditions. These results reveal that encoding processes underlying high-precision short-term memory for facial emotional expressions are modulated depending on whether information has to be stored for one or for several seconds.

INTRODUCTION

Facial emotions are important cues of human social interactions (Ekman 1973; Fridlund 1994; Izard 1977). Emotional expressions are continuously changing, thus should be monitored, memorized, and compared from time to time during social intercourse. Accordingly, recently we have shown that humans possess flawless, high-precision visual short-term memory for emotional expressions (Bankó et al. 2009). It was found that discrimination of emotional expressions is just as efficient when the faces to be compared are separated by several seconds as when they are presented with a very short 1-s delay. However, an important, unresolved question that remains to be explored is whether high-fidelity short-term memory for facial emotional expressions is based on the same neural mechanisms at different retention durations or, alternatively, encoding and retrieval processes are changing depending on how long emotional information has to be stored in visual short-term memory. This later possibility is supported by two lines of experimental results. First, neuropsychological research revealed that patients with medial temporal lobe (MTL) lesions are impaired on visual working memory (WM) tasks only when information has to be stored for several seconds, and no WM deficits were found in the same tasks when retention duration was very short, 1 s (Nichols et al. 2006; Olson et al. 2006). Second, previous research investigating short-term memory processes using delayed discrimination of basic visual dimensions (e.g., spatial frequency) (Magnussen 2000; Magnussen et al. 1998; Reinvang et al. 1998) as well as facial attributes (Bankó et al. 2009) found a significant increase in reaction times (RTs) at delays >3 s as compared with shorter 1-s delays. Based on these findings it was proposed that increased RTs at longer delays might reflect the involvement of memory encoding and retrieval processes, which are absent at delays <3 s. Previous research showed that short-term memory processes for facial attributes involve a distributed network of brain areas (Bledowski et al. 2006; Druzgal and D’Esposito 2003; LoPresti et al. 2008; Nichols et al. 2006; Postle et al. 2003; Yoon et al. 2006) and that encoding and retrieval mechanisms are reflected in several ERP components evoked by the face stimuli (Bledowski et al. 2006; Fabiani et al. 1986; Haenschel et al. 2007; McEvoy et al. 1998; Morgan et al. 2008; Prince et al. 2009), including the early sensory P100 and N170 components as well as the late memory-related P3b wave complex. However, the question whether the length of retention interval affects neural processes of short-term memory for facial attributes has not been investigated before. The goal of the present study was to directly compare short-term memory encoding processes for facial expressions (i.e., happiness) when the faces to be compared (sample and test faces) are separated by a 1-s or a several-second retention interval. We recorded event-related potentials (ERP) while participants performed the same delayed emotion discrimination task with a 1- or 6-s interstimulus interval (ISI). We focused our analysis on the processing of sample faces reflecting encoding processes. The reason for this is that in the 1-s ISI condition processing of sample faces could result in adaptation effects on the processing of test faces, which would be reflected in the ERP responses. Previous results suggested that even a brief (<1 s) presentation of an adaptor face stimulus can modulate the neural response to the subsequently presented test face stimulus, leading to reduced N170 (Kovács et al. 2007) as well as M170 amplitude (Harris and Nakayama 2007), which is an MEG component corresponding to the N170 component measured with EEG. Even though it was suggested that such rapid adaptation effects might persist only for a very short time (~1 s) (Harris and Nakayama 2007), we cannot exclude the possibility that early sensory components of the ERP responses to the test faces in the 1-s ISI condition are affected by adaptation. Therefore here we present and analyze only the ERP results obtained in response to the sample faces. Our findings revealed that retention interval affects both the early sensory and the late
P3b components of the ERP responses during encoding of facial emotional information.

METHODS

Subjects

Nineteen (4 left-handed, 5 females, mean age: 23 yr) subjects gave their informed and written consent to participate in the study, which was approved by the local ethics committee. None of them had any history of neurological or ophthalmologic diseases and all had normal or corrected-to-normal visual acuity. Two participants were excluded from further analysis due to excessive blink artifacts leaving not enough segments to analyze.

Visual stimuli

Stimuli consisted of front view pictures of faces with gradually changing emotional expressions of happiness. Faces were cropped and covered with a circular mask. Test stimuli of varying emotional intensity were generated with a morphing algorithm (Winmorph 3.01) (Kovács et al. 2005–2007) by pairing a neutral and a happy picture of the same person, creating a set of intermediate facial expression images (Bankö et al. 2009). Four sets were created from black and white photos of four unknown actors (2 females and 2 males). Stimuli (8°) were presented centrally (viewing distance of 60 cm) on a uniform gray background. Four stimulus pairs were chosen from each face set. The difference in emotional expression between the pairs was individually adjusted based on a practice session to yield ~75% performance.

Procedure

Observers were shown face pairs, which were separated by a varying ISIs. They performed a delayed-match-to-sample emotion discrimination task whereby they were required to report which of two successively presented faces, termed sample and test, had the happier facial expression. Emotion discrimination was measured by a two-interval forced-choice procedure. Presentation order of the two faces of each pair was counterbalanced within a block, one of the pair appearing as sample image for half of the trials and as test image for the other half. Two ISIs were used for testing: a short 1-s and a longer 6-s delay because in an earlier pilot experiment we found that in the other half. Two ISIs were used for testing: a short 1-s and a longer 6-s delay because in an earlier pilot experiment we found that in

FIG. 1. Delayed-match-to-sample paradigm. Two faces (sample and test) were displayed for 300 ms with either a 1- or 6-s interstimulus interval (ISI). Participants indicated by button press which of the 2 faces had a happier expression. Each face was preceded by a temporal cue to reduce timing uncertainty.

Data analysis

Responses and reaction times were recorded during the experiment. For further analysis, only RTs for correct response trials were used. Both measures were averaged across all face pairs and the average accuracy and RT of each subject for the 1- and 6-s ISI conditions were analyzed with paired t-test.

EEG acquisition and processing

EEG data were acquired using a BrainAmp MR (Brainproducts GmbH., Munich, Germany) amplifier from 60 Ag/AgCl scalp electrodes placed according to the extended 10–20 international electrode system and mounted on an EasyCap (Easycap GmbH, Herrsching-Breitbrunn, Germany) with four additional pericoronal electrodes placed at the outer canthi of the eyes and above and below the right eye for the purpose of recording the electrooculogram. All channels were referenced to the nose on-line and later digitally transformed to an average reference; the ground was placed on the nasion. All input impedance was kept <10 kΩ. Data were sampled at 1,000 Hz with an analog band-pass of 0.016–250 Hz. Subsequently, a digital 0.1-Hz, 12-db/octave zero-phase Butterworth high-pass filter was used to remove DC driffs, and a 50-Hz notch filter was applied to minimize line-noise artifacts. Finally, a 48-db/octave low-pass filter with cutoff frequency of 30 Hz was applied. Trials that contained voltage fluctuations exceeding ±100 μV, or electro-oculogram activity exceeding ±70 μV were rejected. Data processing was done using BrainVision Analyzer (Brainproducts GmbH., Munich, Germany) and custom-written Matlab software. Scalp maps were visualized with the help of Brain Electric Source Analysis program (BESA 5.2, MEGIS Software GmbH, Gräfelfing, Germany).

ERP data analysis

The trial-averaged EEG waveform—i.e., the ERP—was computed including correct trials only. Data were segmented into 700-ms epochs separately for sample and test faces starting from 100 ms preceding the stimulus. Segments were baseline corrected over the 100-ms prestimulus window, artifact rejected and averaged to obtain the ERP waveforms for each subject for each condition. To quantify differences between the ISI conditions, the global field power (GFP) was determined because it constitutes a single reference-free measure of response strength (Lehmann and Skrandies 1980), and its local maxima reflect components in the ERP (Hamburger and v.d. Burgt 1991; Lehmann and Skrandies 1980). The GFP was computed as the SD of the strongest activity at the time of the identified peaks (P100: occipital cluster—O1, Oz, O2, O9, Iz, O10; N170: temporal cluster—...
P9, P7, P8, P10, PO9, PO7, PO8, P10; P3b: parietal—C1, Cz, C2, CPz, P1, Pz, P2) were pooled into clusters and ERPs were analyzed on these clusters separately.

Statistical analysis

To test for significant differences between the ISI conditions for sample faces, 90% confidence intervals for the GFP difference and the difference waves on the electrode clusters were determined using the bootstrap bias-corrected and adjusted (BCa) method (Efron and Tibshirani 1993). This method provides confidence intervals without the necessity of prior assumptions about the distribution of the individual waveforms, e.g., a Gaussian distribution. The difference between 6- and 1-s ISI conditions was considered significant if the confidence interval of the electrode cluster/GFP difference did not include zero (Bledowski et al. 2006; Hoechstetter et al. 2001; Strobel et al. 2008).

RESULTS

Behavioral results

Behavioral data showed that in accordance with our previous results (Bankó et al. 2009), ISI did not affect delayed facial emotion discrimination accuracy. Subjects performed just as well in the 6-s ISI (mean ± SD: 80.0 ± 5.9%) as they did in the 1-s ISI (79.7 ± 7.2%) condition, as it is shown by the lack of significant difference between the two conditions [paired t-test: t(16) = −0.26, P = 0.80]. Reaction times for correct trials, on the other hand, were significantly longer in the 6-s ISI (853 ± 158 ms) compared with the 1-s ISI (803 ± 149 ms) condition [t(16) = −2.69, P = 0.016]. The possibility that enhanced temporal uncertainty caused the increased RT in the 6-s ISI condition can be excluded because of the presence of temporal cues in our study.

ERP results

Mean ERP waveforms of both ISI conditions that were time locked to the onset of the sample faces are presented in Fig. 2. Similar ERP components can be identified on the basis of their latency and scalp topography. The scalp topography of the P100, N170, and P3b components averaged over the 2 ISI conditions are shown in Fig. 3. A difference map shows the strongest activity at the time of the P100, N170, and P3b components. On all time courses gray bands indicate 90% BCa bootstrap derived confidence intervals of the 6- vs. 1-s delay differences (thin dark gray lines). Scalp topography of the P100, N170, and P3b components averaged over the 2 ISI conditions are shown in Fig. 3. B—D, top, respectively. Middle: the ERP waveforms of the clusters of occipital, temporal and parietal electrodes showing the strongest activity at the time of the P100, N170, and P3b components. Bottom: scalp distribution of the voltage difference between 1- and 6-s ISI conditions (6-s ISI - 1-s ISI) is shown for the 3 different ERP components. On all time courses gray bands indicate 90% BCa bootstrap derived confidence intervals of the 6- vs. 1-s delay differences (thin dark gray lines).

FIG. 2. Grand average event-related potential (ERP) responses. ERPs evoked by the sample faces in the 1- and 6-s ISI conditions are shown in dark and light gray, respectively.

FIG. 3. The effects of retention interval on the ERP responses during encoding. A: mean global field power (GFP) waveforms in the 1- and 6-s ISI conditions (dark and light gray lines, respectively). GFP maxima showing significant differences between the 2 delay conditions are marked with black lines. Scalp topography of the P100, N170, and P3b components averaged over the 2 ISI conditions are shown B—D, top, respectively. Middle: the ERP waveforms of the clusters of occipital, temporal and parietal electrodes showing the strongest activity at the time of the P100, N170, and P3b components. Bottom: scalp distribution of the voltage difference between 1- and 6-s ISI conditions (6-s ISI - 1-s ISI) is shown for the 3 different ERP components. On all time courses gray bands indicate 90% BCa bootstrap derived confidence intervals of the 6- vs. 1-s delay differences (thin dark gray lines).
Effect of ISI on the neural responses during encoding

Bootstrap statistics revealed that mean GFP waveforms differed significantly between the 1- and 6-s ISI conditions (Fig. 3A) in the case of sample faces. ISI affected GFP (larger GFP in 1- than in 6-s ISI condition) already in a very early time window, peaking around 110 ms, which corresponds to the P100 ERP component. In accordance with this, we found a significantly larger P100 component on electrodes over the occipital cortex in the 1-s ISI than in the 6-s ISI condition (Fig. 3B). On the other hand, GFP was stronger in the 6- than in the 1-s ISI condition between 140 and 200 ms in the time interval of the N170 ERP component. Correspondingly, a significantly larger N170 component was found on electrodes over the temporal cortex in the 6- than in the 1-s ISI condition (Fig. 3C). Furthermore, GFP was also larger in the 6-s ISI than in the 1-s ISI condition in a later time interval lasting from 375 to 600 ms, corresponding to the P3b component. In accordance with this, activity of the parietal electrodes was significantly more positive in this time interval in the 6-s ISI than in the 1-s ISI condition (Fig. 3D). Interestingly, the scalp distribution of the voltage difference between 1- and 6-s ISI conditions in the case of the P3b component did not match the scalp distribution of the P3b component itself, the former displaying maximum over fronto-central sites as opposed to the parietal maximum found in the case of the P3b peak both in the 1- and 6-s conditions (Fig. 3D).

DISCUSSION

Using a delayed facial emotion discrimination task, we found significant differences in the early P100 and N170 components of the ERP responses to the sample faces between the 1- and 6-s ISI conditions, showing that early sensory processing is modulated by retention interval during encoding. Such an early onset of the memory-related modulation of neural responses during encoding appears to be in agreement with several lines of recent findings. It was found that memory processes modulate low-frequency oscillatory activity (in particular in the alpha band), the amplitude and the inter-trial phase synchronization of which is known to affect the P100 component (for review, see Freunberger et al. 2009; Klimesch et al. 2008). Furthermore, using a delayed-match-to-sample task, it was shown (Haenschel et al. 2007; Morgan et al. 2008) that increasing working memory load—i.e., the number of complex visual objects or faces that need to be compared—leads to increased P100 amplitudes in the ERP responses to the sample stimuli. In addition, the study by Morgan et al. (2008) revealed that WM load also modulates the amplitude of the N170 component both during encoding and retrieval. Using a face matching task with a 1-s ISI, it was found that increasing the number of faces that should be memorized during encoding resulted in larger N170 amplitudes in the ERP responses to the sample face stimuli and reduced N170 amplitudes to the test stimuli. However, it is not known whether modulation of the P100 and N170 amplitudes in response to the sample stimuli by WM load in these studies (Haenschel et al. 2007; Morgan et al. 2008) was due to the enhanced sensory processing demands posed by the increased number of objects/faces presented during the encoding stage or due to modulation of the encoding processes by WM load. The importance of the present results is that they provide the first evidence that encoding processes might affect P100 and N170 amplitudes even when the stimulus information is kept constant, which is in line with findings of Sreenivasan and colleagues (2007), who showed that the N170 amplitude to noise stimuli during maintenance is affected by the type of stimulus being maintained. However, the exact mechanisms underlying modulation of the early stages of facial information processing by short-term memory retention duration remains to be explored.

ERP responses to the sample faces were modulated by retention duration also in a later time window—starting from around 350 ms and peaking at ~525 ms—corresponding to the later peak of the previously described memory-related positive ERP component, known as P3b (Bledowski et al. 2006; Fabiani et al. 1986; McEvoy et al. 1998; Morgan et al. 2008). The scalp distribution of the difference waveform of the P3b component between the 1- and 6-s conditions showed a fronto-central maximum. In fact, previously it was shown that modulation of the P3b component by WM load originates from the prefrontal cortex (Bledowski et al. 2006). The late P3b component was proposed to reflect interaction between the prefrontal cortex and the temporal and parietal regions underlying memory search processes and matching of the representation of the incoming stimuli with stored memory representations (Bledowski et al. 2006; Morgan et al. 2008; Polich 2007). Our results revealed that the P3b component is reduced in responses to sample faces in the 1-s delay condition as compared with the P3b evoked by the sample faces in 6-s delay condition. A possible explanation of these results might be that 1-s delay representation of the sample face information is maintained on-line via persistent delay period activity (for review, see Pasternak and Greenlee 2005) of a neural network involving the visual cortical areas specialized for face processing. On the other hand, at 6-s delay on-line maintenance of the sensory representations of sample face information might not be an efficient strategy due to the increased probability of distraction and interference. Therefore encoding at longer retention intervals would rely to a larger extent on memory search and matching processes and thus will lead to increased P3b component as compared with when information has to be stored for a very short time.

In the current study, we investigated short-term memory processes for facial emotion attributes and showed that retention interval modulates neural processes during encoding. An important question for further research is whether short-term encoding processes are similarly affected by retention interval in the case of other facial attributes as well as in the case of nonfacial information (i.e., other object categories). Previous research showed that processing of changeable facial attributes, including emotional expressions takes place on a specialized network of brain areas involving the superior temporal cortex (Hasselselm et al. 1989; Haxby et al. 2000; LoPresti et al. 2008; Narumoto et al. 2001). It was also found that short-term memory for emotional faces is more efficient than...
for neutral faces (Langeslag et al. 2009) and that the representational advantage for emotional faces might be reflected already on the early components of the ERP responses (Langeslag et al. 2009). However, in our previous study (Bankó et al. 2009), we found that similarly to facial emotions, delayed facial identity discrimination performance in the case of neutral faces also remains constant despite the increase in delay between the stimuli to be compared, which implies that retention interval might affect short-term memory processes of changeable and invariant facial attributes similarly. Thus it remains to be explored whether retention interval effects described in the current study are specific to the short-term encoding of facial emotional expressions or would generalize to encoding processes of all facial attributes. On the other hand, previous research showed that retention interval might have a stronger behavioral effect in the case of low level visual features (Magnussen et al. 1998; Vogels and Orban 1986) as compared with that found in the case of facial attributes (Bankó et al. 2009). Increasing the retention interval in a delayed discrimination paradigm led to performance decay in the case of contrast, orientation and color, but not in the case of spatial frequency and motion information (for a review, see: Magnussen 2000; Pasternak and Greenlee 2005). Therefore one might expect that neural effects of retention interval might be also different in the case of these features.

Taken together, the results of the present study showed that retention interval affects short-term memory encoding processes for facial emotions. We found that in a delayed emotion discrimination task, encoding processes differ when the faces to be compared are separated by several seconds from that when they are presented with a very short, 1-s delay. This is in agreement with previous behavioral (Bankó et al. 2009; Magnussen 2000; Magnussen et al. 1998; Reinvang et al. 1998) and neuropsychological (Nichols et al. 2006; Olson et al. 2006) research suggesting that memory processes at very short (1 s) retention intervals might differ from those involved at longer storage durations. Importantly, in the current study, there was no difference in stimulus information and subjects’ discrimination performance between the two different ISI conditions. This implies that our findings, showing strong modulation of ERP responses by retention interval during encoding of facial emotional information, reflect changes in mnemonic processes as a function of storage duration and cannot be accounted by differences in sensory processing demands or overall task difficulty across the conditions with different retention intervals. The results of the present study thus provide the first evidence that neural encoding processes underlying flawless, high-precision short-term memory for facial emotional expressions are modulated depending on whether information has to be stored for one or several seconds. Our findings also imply that models of short-term memory—which treat storage of sensory information over a period of time ranging from one up to several seconds as a unitary process (for review, see Jonides et al. 2008)—should be revised to include retention interval as an important factor affecting neural processes of memory encoding.

ACKNOWLEDGMENTS

The authors thank I. Kőbor for help with running the experiments and G. Kovács for insightful comments.

GRANTS

This work was supported by Hungarian Scientific Research Fund Grant T048949 and by a grant from the Bolyai Fellowship to Z. Vidnyánszky.

REFERENCES


J Neurophysiol • VOL 103 • MARCH 2010 • www.jn.org


