DOI:10.15774/PPKE.ITK.2010.001

Neural Mechanisms Underlying Visual Short-Term Memory for Facial Attributes

Dissertation submitted for the degree of
Doctor of Philosophy

Éva M. Bankó
neurobiologist

Scientific advisor:
Prof. Zoltán Vidnyánszky, *Ph.D.*, *D.Sc.*

Faculty of Information Technology
Pázmány Péter Catholic University

MR Research Center
Szentágothai Knowledge Center

Budapest April 6, 2010

DOI:10.15774/PPKE.ITK.2010.001

"In contemplation, if a man begins with certainties, he shall end in doubts; but if he be content to begin with doubts, he shall end in certainty."

(Francis Bacon)

DOI:10.15774/PPKE.ITK.2010.001

Acknowledgments

> *"Knowledge is in the end based on acknowledgment."*
> *(Ludwig Wittgenstein)*

First and foremost, I would like to thank and acknowledge the support of my supervisor, *Zoltán Vidnyánszky.* Throughout my work he guided my steps with his great theoretical insight and pushed me even further with his enthusiasm not letting me settle for the average, for which I am most grateful. Hereby I also acknowledge the collaborative work and help of *Gyula Kovács* who have introduced me to human electrophysiology early on in my scientific career.

I am indebted to *Prof. Tamás Roska*, head of the doctoral school for providing me with a multidisciplinary milieu, which was truly decisive throughout my PhD studies. I would also like to express my gratitude to *Zsuzsa Vágó* and *Prof. Árpád Csurgay* for giving me a deeper understanding of mathematics and physics which is essential in my work. I say thanks to *Prof. György Karmos* and *Prof. József Hámori* for sharing their educational experiences with me during the courses I had the opportunity to help out with and for proving that being a good researcher and a good teacher at the same time are not mutually exclusive.

I am thankful to my close colleagues who contributed to my work *Viktor Gál*, *István Kóbor*, *Judit Körtvélyes* and *Lajos Kozák* for their help, valuable discussions and useful advices. I'm greatly honored to have been able to spend some time and gather experience in the Center for Neural Sciences in New York University and I thank *Prof. David Heeger* for the possibility and *Jonas Larsson* and *Tobias Donner* for their tutoring.

Very special thanks go to all my fellow PhD students and friends, especially to *Barnabás Hegyi*, *Gergely Sóos*, *Csaba Benedek*, *András Kiss*, *Frank Gollas*, *Norbert Bérci*, *Béla Weiss*, *Dániel Szolgay*, *Viktor Gyenes*, *Dávid Tisza*, *Antal Tátrai*, *László Füredi*, *József Veres*, *Ákos Tar* and *Giovanni Pazienza.* I so greatly appreciate their scientific help and fruitful discussions on various topics concerning my work. It is them who have contributed to my work in a unique way: lending me a helping hand with the computational aspects of my work from theoretical backgrounds to implementation. I also feel honored for their

DOI:10.15774/PPKE.ITK.2010.001

friendship and thankful for all the fun we had together. The same holds true for my best friend, *Erika Kovács* to whom I am grateful for supporting me through the ups and downs of scientific and personal life.

I acknowledge the kind help of *Anna Csókási*, *Lívia Adorján* and *Judit Tihanyi* and the rest of the *administrative* and *financial personnel* in all the administrative issues. I also thank the *reception personnel* for giving my days a good start with kind greetings and smiles.

Last but not least, I am most grateful to my *Mother* and *Father* who have helped me tremendously in every possible way so that I could focus on research. Their emotional support proved invaluable. I am also thankful to my two sisters, *Verka* for being proud of me and *Ági* for her support and patience during the years we have lived together. I am privileged to have such a family. *Thank you!*

DOI:10.15774/PPKE.ITK.2010.001

Contents

DOI:10.15774/PPKE.ITK.2010.001

Summary of abbreviations

Abbreviation	**Concept**
ANOVA	analysis of variance
BCa method	bootstrap bias-corrected and adjusted method
BOLD	blood oxigenation-level dependent
EEG	electroencephalogram
ERP	event-related potential
FDR	false discovery rate
FFA	fusiform face are
*f*MRI	functional magnetic resonance imaging
FWHM	full-width half-maximum
GFP	global field power
GLM	general-linear model
iFG	inferior frontal gyrus
HRF	hemodynamic response function
ISI	inter-stimulus interval
IT	inferior temporal cortex
ITI	inter-trial interval
JND	just noticeable difference
LO or LOC	lateral occipital cortex

DOI:10.15774/PPKE.ITK.2010.001

Abbreviation	**Concept**
LRP	lateralized readiness potential
MEG	magnetoencephalogram
MTL	medial-temporal lobe
PCA	principal component analysis
PFC	prefrontal cortex
pSTS	posterior superior temporal sulcus
RV	residual variance
RT	reaction time
SEM	standard error of the mean
SP	source pair
TE	time to echo
TFE	turbo field echo
TR	time to relaxation
VSTM	visual short-term memory
WM	working memory

DOI:10.15774/PPKE.ITK.2010.001

Chapter 1

Introduction

1.1 Motivations

Face processing is one of the most researched fields of cognitive neuroscience, since the majority of socially relevant information is conveyed by the face, rendering it as a stimulus of exquisite importance. On the other hand, the development of reliable computational face recognition algorithms is a key issue in computer vision, which has been the focus of extensive research effort. Nonetheless, we have yet to see a system that can be deployed effectively in an unconstrained setting with all of the possible variability in imaging parameters such as sensor noise, viewing distance, illumination and in facial expressions that has a huge impact on the configural layout of the particular face. The only system that does seem to work well in the face of all these challenges is the human visual system itself. Therefore, advancing the understanding of the strategies this biological system employs is not only invaluable for a better insight of how the brain works but also a first step towards eventually translating them into machine-based algorithms (for review see [10]).

1.2 Faces are special but in what way?

Despite the centuries of research that has gone into unveiling the mechanisms of face processing, the central question still remains: Is face perception carried out by domain-specific mechanisms, that is, by modules specialized for processing

DOI:10.15774/PPKE.ITK.2010.001

faces in particular [11, 12]? Or are faces handled by domain-general, fine-level discriminator mechanisms that can operate on nonface visual stimuli as well [13, 14]?

1.2.1 FFA: ***F***usiform ***F***ace ***A***rea

The proponents of the first, domain-specific view [11, 12] invoke several lines of evidence to support their hypothesis. Namely, psychophysical observations suggest s special mechanism for face as opposed to object processing in the temporal cortex, since face recognition is more disrupted by inversion (i.e. turning the stimulus upside down) than is object recognition (the well-known face inversion effect) [15, 16]. Also, accuracy at discriminating individual face parts is higher when the entire face is presented than when the parts are presented in isolation, whereas the same holistic advantage is not found for parts of houses or inverted faces [17]. Other strong support can be found in the neuropsychological literature that there is a double dissociation between face and object processing: patients of prosopagnosia are unable to recognize previously familiar faces, despite a largely preserved ability to recognize objects [18], whereas patients of object-agnosia are seriously impaired in recognizing non-face objects with the spared ability to recognize faces [19]. In prosopagnosic patents the brain lesion incorporates a well defined area in the middle fusiform gyrus termed fusiform face area (FFA) [20] either in the right hemisphere or bilaterally [21]. In accordance with this, numerous fMRI studies have shown higher activity in FFA for faces than scrambled faces or non-face objects [12, 20, 22]. Furthermore, a study showed that the FFA was the most likely source of the face-inversion effect [12] but see [23]. Similarly to the fMRI studies, selective responses to faces are reported using scalp ERPs [24, 25] and MEG [26], namely the N170/M170 component which is most prominent over posterior temporal sites and most likely originates in the FFA [27, 28, 29], but see: [23, 30].

1.2.2 FFA: ***F***lexible ***F***usiform ***A***rea

According to the other, domain-general view [13, 14], however, the specific responses obtained for faces is a result of the type of judgment we are required to make whenever viewing a face: differentiating that individual face from the rest (i.e. subordinate-level of categorization) and also the level of expertise with which we make these categorization judgments. These factors represent confounds when comparing the processing of non-face objects to face processing, since objects are generally categorized on a basic level - that is differentiating between e.g. a chair and a table. In fact, several experiment have shown that the FFA is more active for judgments requiring classification at a subordinate level compared to more categorical judgments for a large variety of objects, living or artifactual [31, 32]. Furthermore, faces represent a stimulus of high evolutionary relevance due to their role in social communication, therefore, every healthy individual can be regarded as a "face-expert". In accordance with this the FFA in the right hemisphere are both recruited when observers become experts in discriminating objects from a visually homogeneous category. This occurs both in bird and car experts with many years of experience [33] as well as in subjects trained for only 10 hours in the laboratory to be experts with novel objects called "Greebles" [34]. With expertise acquired, Greebles - similarly to faces - are processed more holistically and the behavioral measure of holistic processing correlates with the increase in FFA activity [35]. Moreover, similar inversion costs can be found in expert dog judges when recognizing dogs, much like the inversion costs all subjects show in recognizing faces [36]. Such expertise effects indicate a high degree of flexibility with regard to acceptable image geometries in the neural network of the FFA.

Despite the above, the two views are not mutually exclusive. Much evidence supporting the domain-general hypothesis is also consistent with the possibility that the increased response of the Fusiform Face Area with acquired expertise reflects distinct but physically interleaved neural populations within the FFA [33]. Indeed, a high-resolution fMRI study by Grill-Spector and colleagues [37] found that the FFA was not a homogeneous area of face selective neurons, but a rather heterogeneous, in that regions of high selectivity for faces were intermingled with regions of lower selectivity and the different regions were also activated by other object categories. So what is so special about FFA? A possible answer comes from a recent modeling study of Tong and colleagues [38], where they trained neural networks to discriminate either at a basic level (basic networks) or at a subordinate level (expert networks). These expert networks trained to discriminate within a visually homogeneous class developed transformations

DOI:10.15774/PPKE.ITK.2010.001

that magnify differences between similar objects, that is in order to distinguish them their representations were spread out within the elements of the network. This was in marked contrast to networks trained to simply categorize the objects because basic networks represent invariances among category members, and hence compress them into a small region of representational space. The transformation performed by expert networks (i.e. magnifying differences) generalizes to new categories, leading to faster learning. These simulations predict that FFA neurons will have highly variable responses across members of an expert category, which is in good agreement of the high variability of voxel activation found by Grill-Spector [37].

1.3 Models of face perception

1.3.1 Distributed neural system for face perception

The prevailing view of face processing is captured in a model proposed by Haxby and colleagues [39] where they propose two functionally and neurologically distinct pathways for the visual analysis of faces: one codes changeable facial properties (such as expression, lipspeech and eye gaze) and involves the inferior occipital gyri and superior temporal sulcus (STS), whereas the other codes invariant facial properties (such as identity) and involves the inferior occipital gyri and lateral fusiform gyrus. These cortical regions comprise the core system of their model which is completed by an extended system that aides in but is not entirely dedicated to face processing (Fig. 1.1). This model is in agreement with the ideas of Bruce and Young [40] who assumed separate functional routes for the recognition of facial identity and facial expression in a model of face recognition. Evidence for this dissociation comes from neuropsychological studies of prosopagnosic patients showing impairments in facial identity recognition but intact facial expression recognition [41, 42, 43, 44, 45, 46]. However, in most of these studies the cause of the identity impairments has not been established [41, 43, 45, 46] and therefore do not prove that this dissociation has a visuoperceptual origin. Single cell recordings from macaques constitute another pool of evidence [47, 48, 49, 50]. These studies have identified a number of face selective cells, most of which responded either to identity or facial expression.

DOI:10.15774/PPKE.ITK.2010.001

The former group of cells was mostly located in the cortex of superior temporal sulcus, while the latter were found predominantly in the inferior temporal gyrus. A smaller portion of the measured face selective neurons, however, responded to both of identity and expression or even showed an interaction between these features.

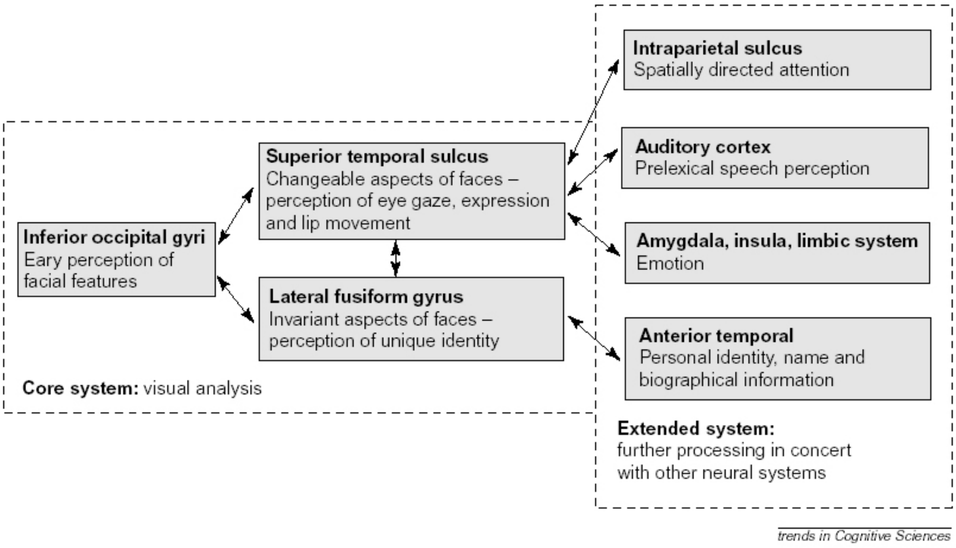

Figure 1.1: A model of the distributed human neural system for face perception. The model is divided into a core system, consisting of three regions of occipitotemporal visual extrastriate cortex, and an extended system, consisting of regions that are also parts of neural systems for other cognitive functions. Changeable and invariant aspects of the visual facial configuration have distinct representations in the core system. Interactions between these representations in the core system and regions in the extended system mediate processing of the spatial focus of anothers attention, speech-related mouth movements, facial expression and identity. [Taken from [39].]

This framework has remained the dominant account of face perception despite the large number of emerging evidence that questions the complete independence of facial identity and expression processing [51, 52, 53, 54, 55, 56]. Even though the central idea of some form of dissociation between these two facial cues is undeniable, these studies all show interaction and overlap between facial identity and emotion processing: in the case of the FFA a sensitivity for emotionally charged faces [51, 55], increased activation when attending to facial expression [54], release from adaptation with change in the facial expression of the adaptor and test faces [56]; and conversely in the case of posterior STS significant adaptation effects to keeping the identity constant across face pairs

[52, 56].

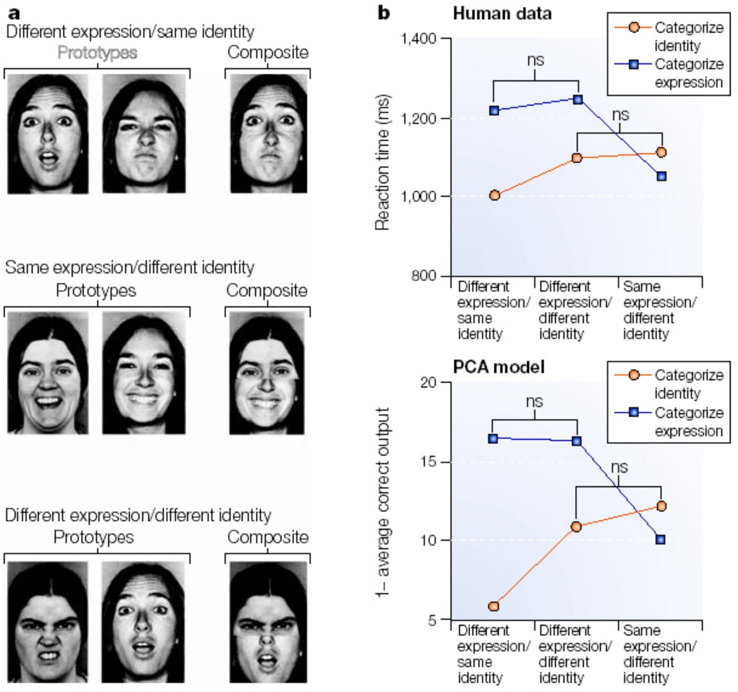

Figure 1.2: PCA model. Principal component analysis (PCA) is a form of linearized compact coding that seeks to explain the relationships among many variables in terms of a smaller set of principal components (PCs). As applied to faces, the pixel intensities of a standardized set of images are submitted to a PCA. Correlations among pixels are identified, and their coefficients (PCs) are extracted. The PCs can be thought of as dimensions that code facial information and can be used to code further faces. The particular advantage of techniques such as PCA is that they can reveal the statistical regularities that are inherent in the input with minimal assumptions. Panel *a* shows composite faces that were prepared by combining the top and bottom halves of two faces with different expressions posed by the same identity, the same expression posed by different identities, or different expressions posed by different identities. Reaction times for reporting the expression in one face half were slowed when the two halves showed different expressions (that is, different expression/same identity and different expression/different identity) relative to when the same expressions (posed by different identities) were used (that is, same expression/different identity); however, no further cost was found when the two halves contained different expressions and identities compared with when they contained different expressions and same identities (top graph in panel *b*). A corresponding effect was found when subjects were asked to report the identity of one face half. The bottom graph in panel *b* shows a simulation of this facial identityexpression dissociation using a PCA-based model. (ns, not significant; all other comparisons among the categorize identity or categorize expression levels were statistically reliable.)[Taken from [53].]

DOI:10.15774/PPKE.ITK.2010.001

1.3.2 Principal component analysis model of face perception

In the light of the above evidence, Calder and Young [53] suggested a relative rather than absolute segregation of identity and expression perception. They based this argument on findings from principle component analyses (PCA) showing that certain components were necessary for discriminating facial identity, others for discriminating facial expression, and yet others for discriminating either (Fig. 1.2). Their PCA model offers a different perspective, in that it shows that the independent perception of facial identity and expression can arise from an imagebased analysis of faces with no explicit mechanism for routing identity- or expression-relevant cues to different systems. The result is a single multidimensional framework in which facial identity and expression are coded by largely (although not completely) different sets of dimensions (PCs). Therefore, independent perception does not need to rely on totally separate visual codes for these facial cues. [53].

1.4 Visual short-term memory for faces

However, if faces represent a class of stimuli of special importance, it is not only the neural mechanisms underlying processing of facial attributes that needs to be fine-tuned. The same should hold true for higher cognitive processes dealing with faces. Memory seems to be an especially important mechanism among these, since in every social encounter efficient processing of the faces in itself is not enough if we cannot remember who the person that we encountered was. In accordance with this, Curby and Gauthier [57] found a visual short-term memory (VSTM) advantage for faces in that given sufficient encoding time more faces could be stored in VSTM than inverted faces or other complex non-face objects. Their experiments point towards the conclusion that the reason for this advantage is holistic processing, since faces are processed more holistically than objects or inverted faces. They recently found the same advantage for objects of expertise [58], which are known to be processed more holistically than objects with which one does not have expertise [35]. Furthermore, Freire and colleagues [59] have shown that the VSTM difference between upright and

DOI:10.15774/PPKE.ITK.2010.001

inverted faces was present only when the facial configural information has to be encoded and stored as opposed to the null effect of orientation in cases when only featural information changes. They also showed that there was no decay in discrimination performance of these gross featural/configural changes over time up to 10 seconds. The time-scale of VSTM capacity for realistic fine changes is not known, however, which has evolutionary significance in monitoring continuously changing facial features such as facial mimic conveying important emotional information.

1.5 Goals of the dissertation

In accordance with the above, the present dissertation focuses on the time-scale of VSTM capacity of different facial features which have different statistical probablity of changing over time (i.e. time-invariant and changeable facial features) by comparing facial memory across different delay intervals in a series of well controlled experiments. It also aims at uncovering the neural mechanism of short-term memory for these facial attributes by investigating the time-course of neural activation of encoding and retrieving the faces to be compared when separated by variable delay periods.

DOI:10.15774/PPKE.ITK.2010.001

Chapter 2

Characterization of short-term memory for facial attributes

2.1 Introduction

Facial emotional expressions are crucial components of human social interactions [60, 61, 62]. 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 [63, 64] and thus will be processed efficiently even in the presence of distractors [65] or in cases of poor visibility [66]. 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 could 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 [39, 53]. Invariant facial attributes do not require constant online monitoring during social interaction. After registering a person's

DOI:10.15774/PPKE.ITK.2010.001

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 [67]. Furthermore, it was shown that the processing of changeable and invariant facial attributes took place on specialized, to some extent independent functional processing routes [39, 53]. Functional neuroimaging results suggested that facial identity might be processed primarily in the inferior occipito-temporal regions, including the fusiform face area [20, 68], whereas processing of the information related to emotional expressions involved the superior temporal cortical regions [49, 69, 51, 70, 52, 71]. 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.

Our goal was to investigate in a series of experiments how efficiently humans could 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 affected facial emotion and facial identity discrimination. The logic of our approach was as follows: if there was 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 was lost during the process of memory encoding, maintenance or recall then increasing the delay between the faces to be compared should impair discrimination performance.

DOI:10.15774/PPKE.ITK.2010.001

2.2 EXP. 1: Short-term memory capacity for facial attributes

2.2.1 Motivations

Previous research investigating short term memory for basic visual dimensions (e.g. spatial frequency and orientation) using delayed discrimination tasks [72, 73, 74] found a significant increase in reaction times (RT) at delays longer than 3 s as compared to shorter delays. It has been proposed that increased RTs at longer delays might reflect the involvement of memory encoding and retrieval processes, which were absent at delays shorter than 3 s. To test whether increasing the delay lead 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 was a significant increase in RTs when the faces were separated by more than 3 s. Furthermore, it was also found that RTs saturated 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)(Fig. 2.2a).

2.2.2 Methods

Subjects. Ten subjects (6 females, mean age: 24 years) gave their informed and written consent to participate in Experiment 1, which was approved by the local ethics committee. 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 cov-

DOI:10.15774/PPKE.ITK.2010.001

ered 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) [75, 76, 77] by 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. (Fig.2.1a-d). 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 grey 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 (Fig.2.2b-c). 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 50-100% reference identity strength. (Fig.2.2d) The rational behind choosing the 75% instead of the 50% reference identity as the midpoint image for identity discrimination was to have test faces which clearly exhibit the reference identity as pilot experiments revealed that using test stim-

DOI:10.15774/PPKE.ITK.2010.001

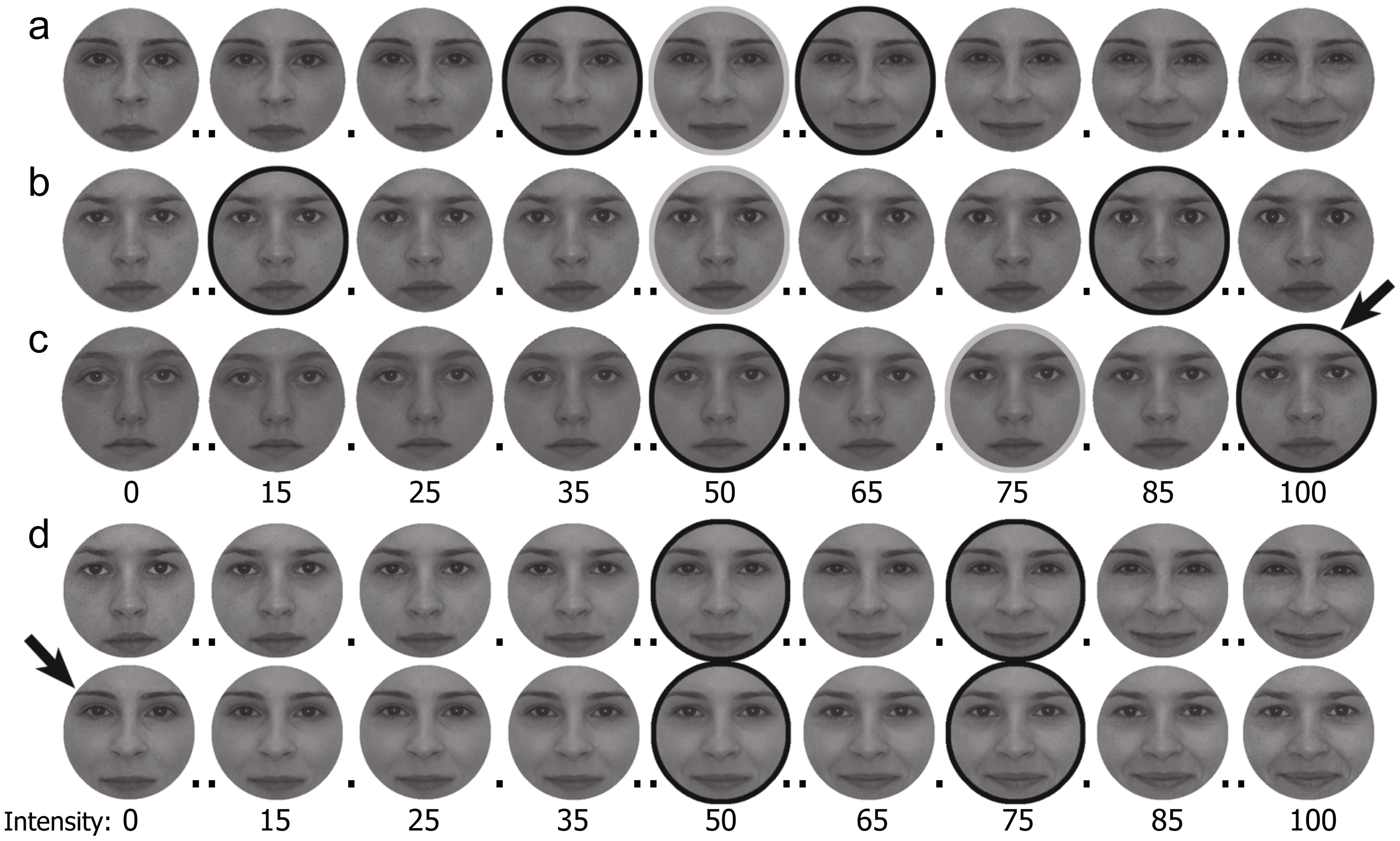

Figure 2.1: Exemplar (a) happy, (b) fearful and (c) identity morphed face sets used in Experiment 1-3. Grey circles indicate midpoint faces used as one constituent of each face pair while black circles show the two extremes of each set used as the other constituent in Experiment 1-3. Figure shows the original span of 101 faces, but the actual morph continua used were assigned to the [0 1] interval for analysis and display purposes. (d) Exemplar composite face set used in Experiment 4. Black circles indicate a typical face pair yielding 75% performance. The arrow indicates the reference face for identity discrimination for sets (c-d).

DOI:10.15774/PPKE.ITK.2010.001

uli 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 (Fig.2.2b-c).

Procedure. A trial consisted of 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 (Fig.2.2a). Subjects were given a maximum 2 s response window. The intertrial interval (ITI) was randomized between 400-600 ms. The fixation cross was present troughout the entire experiment. 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

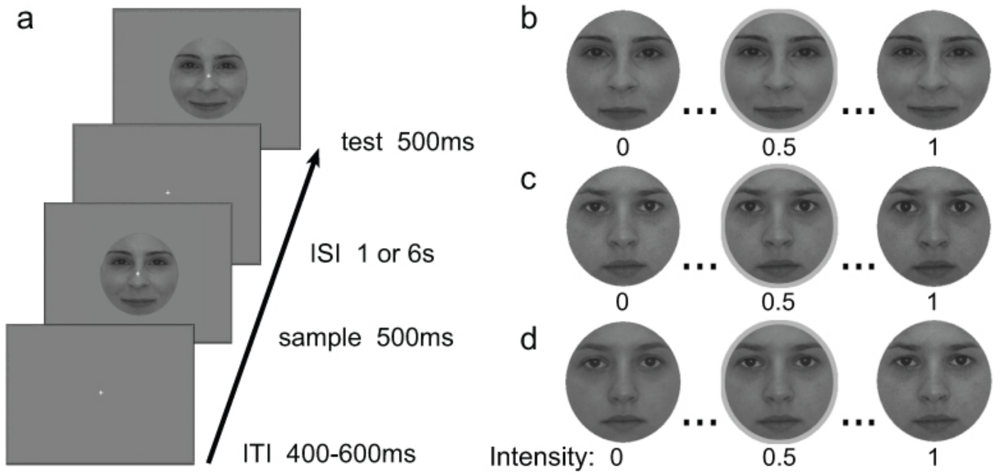

Figure 2.2: Experimental design and morphed face sets used in Experiment 1. (a) Stimulus sequence with a happiness discrimination trial. Stimulus sequence was similar for Experiment 2-4. Exemplar (b) happy, (c) fearful and (d) identity morphed face sets used in Experiment 1. Each face pair consisted of the midpoint face - indicated by grey circles - and one of eight predefined stimuli. 0 and 1 show the typical two extremes while the other six stimuli were evenly distributed in between.

DOI:10.15774/PPKE.ITK.2010.001

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 [78]. 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 [79]. Reaction times were calculated as the average of the reaction times for stimuli yielding 25% and 75% performance. 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.

2.2.3 Results

The pilot experiment revealed that increasing the delay between the face pairs lead 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 (Fig.2.3a). 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 results of the pilot experiment, in all three discrimination conditions reaction times were longer by approximately 150-200 ms in the LONG ISI (6s) than in the SHORT ISI (1s) conditions (Fig.2.3b), providing support for the involvement

DOI:10.15774/PPKE.ITK.2010.001

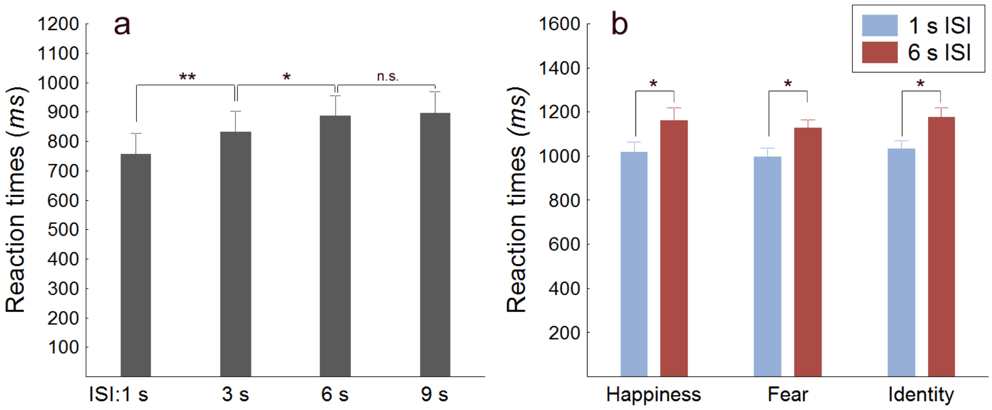

Figure 2.3: Reaction times for delayed emotion (happiness) discrimination measured during (a) a pilot experiment and (b) Experiment 1. Mean RTs were calculated from trials with face pairs yielding 25% and 75% percent performance. There was a significant difference in RT between 1-s and 6-s ISI, while RTs saturated around 6 s. Error bars indicate $\pm SEM$ ($N = 3$ and 10 for the pilot experiment and Experiment 1, respectively; $*p < 0.05$, $**p < 0.01$, $n.s.$ not significant).

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 ($F_{(2,18)} = 0.022, p = 0.978$).

However, alternative explanations regarding the RT difference are also possible since other mechanisms can be in play. The typical experimental setting for these experiments is blocking the different ISI conditions in separate runs and presenting only the two stimuli to be compared in each trial. Therefore, one of the caveats is the increased temporal uncertainty concerning the presentation of the second face stimulus which subjects are required to respond in the case of the 6-s ISI. Indeed, if temporal uncertainty is equated by introducing a temporal cue shortly before the second stimulus in both delay conditions this RT difference drops to 50 ms but still remains statistically significant (see Chapter 3 for more details). Nevertheless, blocking the different ISIs in separate runs further allows for differences in the state of the motor system such as adaptation increasing RTs in the case of the 1-s ISI; conversely, a possible decreased motor

DOI:10.15774/PPKE.ITK.2010.001

alertness can result in the increase of RTs in the case of the 6-s ISI condition rendering the meaning of the RT difference unclear. (For further discussion please refer to Appendix A.)

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 (Fig.2.4a-c, see also Fig.2.4d 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.

2.3 EXP. 2: The importance of configural processing

2.3.1 Motivations

The primary goal of the above study was to investigate how efficiently could humans recognize and monitor changes in different facial attributes. Therefore, it was crucial to show that performance in our facial attribute discrimination task was indeed based on high-level, face-specific attributes or attribute configurations as opposed to some intermediate or low level feature properties of the face images (e.g. local contour information, luminance intensity). For this reason, in Experiment 2 a follow-up study was conducted with the same three facial attribute discrimination conditions as in Experiment 1, only the same stimuli were also presented in an inverted position. Thus, taking away the configural

DOI:10.15774/PPKE.ITK.2010.001

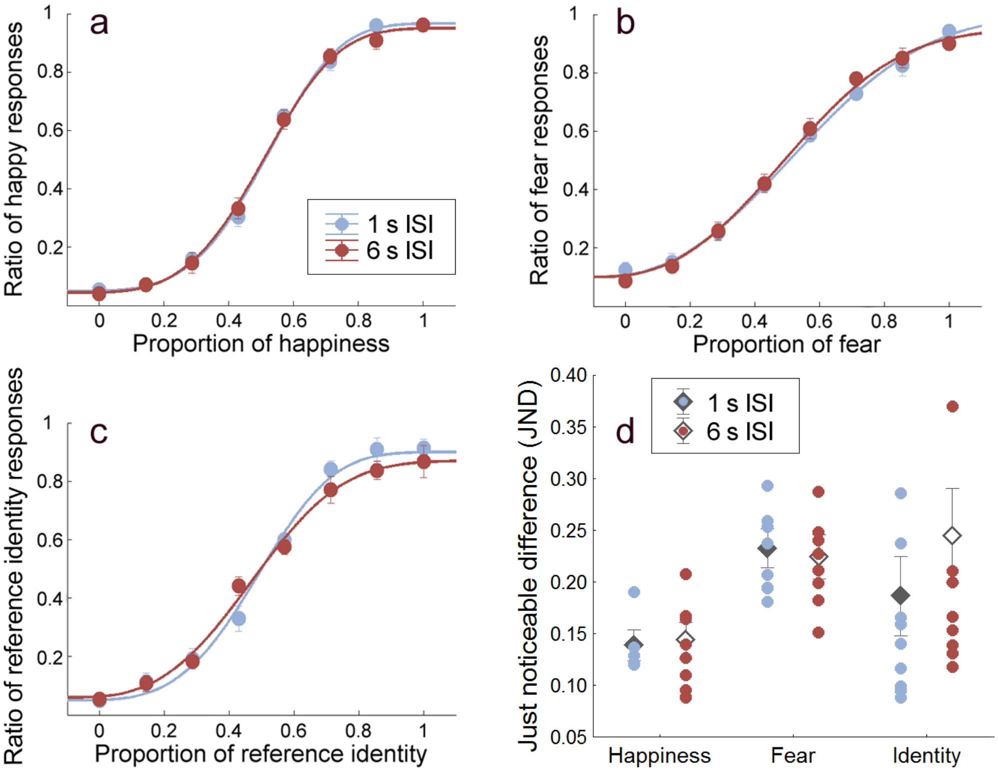

Figure 2.4: Effect of ISI on the performance of facial emotion and identity discrimination. Weibull psychometric functions fit onto (a) happiness, (b) fear and (c) identity discrimination performance. Introducing a 6 s delay (brown line) between sample and test faces had no effect on emotion discrimination 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 $\pm SEM$ ($N = 10$).

feature but leaving the low level features unaltered [80, 16].

2.3.2 Methods

Subjects. Six right-handed (all female, mean age 23 years) 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.

DOI:10.15774/PPKE.ITK.2010.001

Stimuli and Procedure The same face sets were used as in Experiment 1. Stimulus presentation was very similar to that of Experiment 1 with only minor differences: faces were presented centrally both in an upright and in an inverted position with using a SHORT interstimulus delay of 1 s only. Every other parameter and procedural detail was identical to Experiment 1.

Data analysis. Performance evaluation and reaction times calculation for Experiment 2 was done identically. All measurements were entered into a 3×2 repeated measures ANOVA with attribute (happy vs. fear vs. identity) and orientation (upright vs. inverted) as within subject factors. Post-hoc t-tests were done by Tukey HSD tests.

2.3.3 Results

Reaction times did not differ across attributes or orientations. ANOVA performed on the RT data showed no main effect of attributes (happiness $vs.$ fear $vs.$ identity, $F_{(2,10)} = 1.12, p = 0.364$), no significant main effect of orientation (upright vs. inverted, $F_{(1,5)} = 0.493, p = 0.514$), and no interaction between these variables ($F_{(2,10)} = 1.19, p = 0.343$).

Compared to the conditions with upright faces, the performance for inverted faces was degraded in all three facial attribute discrimination conditions increasing the JND values (Fig.2.5). The performance for each attribute also differed; higher JND values were obtained for fearful faces than for the other two attributes. The ANOVA results yielded a significant main effect of attributes (happiness $vs.$ fear $vs.$ identity, $F_{(2,10)} = 39.48, p < 0.0001$), a significant main effect of orientation ($F_{(1,5)} = 28.34, p = 0.003$) and no interaction between these variables ($F_{(2,10)} = 2.11, p = 0.172$). Since it has been shown that face inversion selectively affected processing of high-level, face-specific information [80, 16], these results provide support that facial attributes in our experiment with upright faces were compared based on at least partially face-specific featural and configural information rather than solely on the intermediate and low level feature properties of the face images.

DOI:10.15774/PPKE.ITK.2010.001

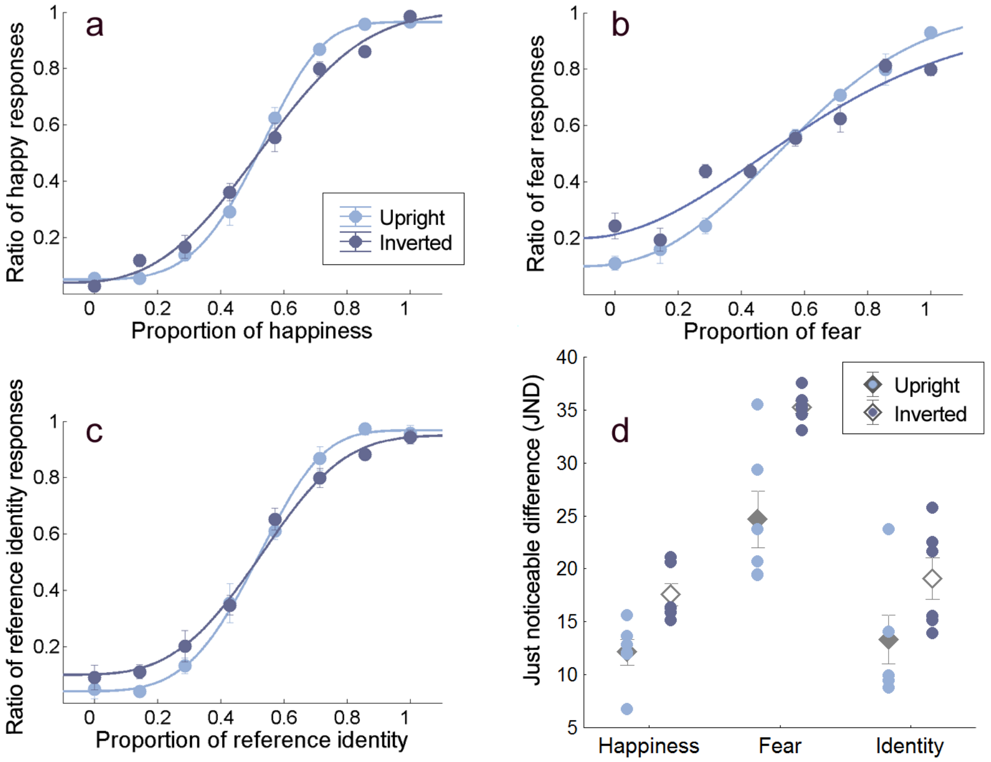

Figure 2.5: Effect of inversion on the performance of facial emotion and identity discrimination. Weibull psychometric functions fit onto (a) happiness, (b) fear and (c) identity discrimination performance. Performance in all conditions were degraded by inverting the face stimuli to be compared (dark blue line) as opposed to upright presentation (light blue line). The x axis denotes morph intensities of the constant stimuli. (d) Just noticeable differences (JNDs) obtained in Experiment 2. Diamonds represent mean JNDs in each condition while circles indicate individual data for upright (light blue) and inverted (dark blue) presentation. Error bars indicate $\pm SEM$ ($N = 6$).

2.4 EXP. 3: The role of learning in facial attribute discrimination and memory

2.4.1 Motivations

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

DOI:10.15774/PPKE.ITK.2010.001

term memory for facial emotions also extended to situations where the faces and the delayed discrimination task were 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 was based on the perceptual memory representation of the sample stimulus [74, 81] or it was 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 criterion-setting theory [82]. This is because in Experiment 3 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.

2.4.2 Methods

Subjects. Altogether 206 participants took part in Experiment 3, which was approved by the local ethics committee. 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 Procedure. In Experiment 3 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 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

DOI:10.15774/PPKE.ITK.2010.001

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 [83] 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 3 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 difference separately. [83]. 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 condition). RT data was 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.

2.4.3 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 (Fig.2.6a). 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 (Fig.2.6b). 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 difference, 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 large and small difference, 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 novel.

Since the face images used in Experiment 3 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 2.6b in Experiment 3 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

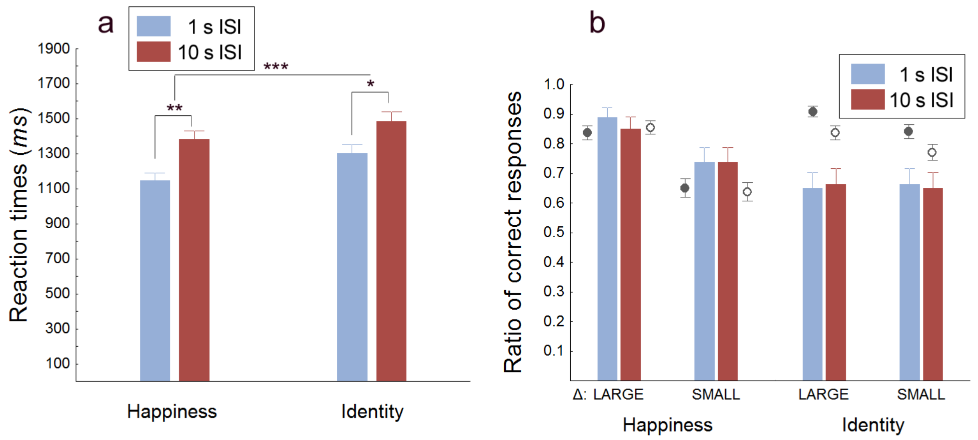

Figure 2.6: Reaction times and discrimination performance in Experiment 3. (a) There was a significant RT increase in the LONG compared to the SHORT ISI condition in the case of both attributes (Valid number of measurements: $N = 74$ and $N = 60$ for the SHORT and LONG ISI condition, respectively). (b) 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 Experiment 1 and Experiment 3, grey 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 Experiment 3 and 1, respectively; $*p < 0.05$, $**p < 0.01$, $***p = 0.001$).

DOI:10.15774/PPKE.ITK.2010.001

of practice in Experiment 1. On the other hand, overall identity discrimination performance in Experiment 3 was worse than in Experiment 1, suggesting that practice and familiarity of faces affected performance in the facial identity discrimination task but not in the facial emotion discrimination task.

2.5 EXP. 4: Activation of cortical areas involved in emotional processing

2.5.1 Motivations

To confirm that emotion discrimination in our short-term memory paradigm involved high-level processing of facial emotional attributes, we performed an *f*MRI experiment. Previous studies have shown that increased *f*MRI responses in the posterior superior temporal sulcus (pSTS) during tasks requiring perceptual responses to facial emotions compared to those to facial identity could be considered as a marker for processing of emotion-related facial information [84, 69, 51, 70, 52, 71]. Therefore, we conducted an *f*MRI experiment in which we compared *f*MRI 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 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 *f*MRI 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.

DOI:10.15774/PPKE.ITK.2010.001

2.5.2 Methods

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

Stimuli. Like in Experiment 3, 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 (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 (Fig.2.7a). 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 pseudorandomized 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 Co., 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.

DOI:10.15774/PPKE.ITK.2010.001

Procedure. The task remained identical to that of Experiment 1 and 2 but the experimental paradigm was slightly altered to be better suited for *f*MRI. A trial began with a task cue (0.5 deg) appearing just above fixation for 500 ms being either 'E' for emotion and '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 response window. The ISI varied between 5 and 8 seconds 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 X™ 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, Inc., Natick, MA) using the Psychtoolbox 2.54 [85, 86].

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 Eyetracking 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 dataset, 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 [52].

DOI:10.15774/PPKE.ITK.2010.001

fMRI imaging and analysis. *Data acquisition.* Data were collected at the MR Research Center of Szentágothai Knowledge Center, (Semmelweis University, Budapest, Hungary) on a 3T Philips Achieva (Best, The Netherlands) scanner equipped with an 8-channel SENSE headcoil. 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 3.5×3.5 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 = 24 s, TE = 30 ms, FA = 75°, FOV = 220 mm, 64×64 image matrix, 7 runs, duration of each run = 516 s).

Data analysis. Preprocessing and analysis of the imaging data was performed using BrainVoyager QX v1.91 (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 FWHM 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 $df_{(random)} = 9$.

DOI:10.15774/PPKE.ITK.2010.001

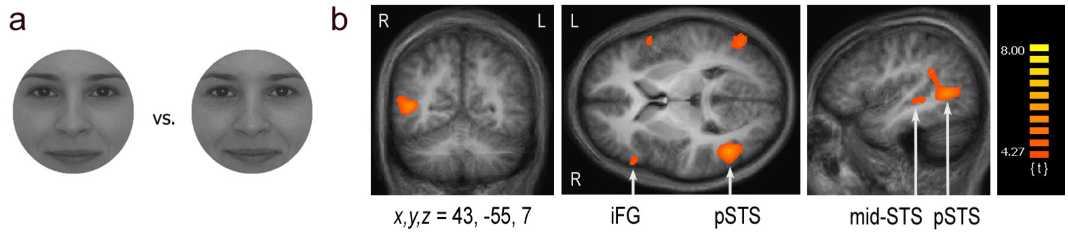

Figure 2.7: Stimuli and results of Experiment 4. (a) An exemplar face pair taken from the female composite face set which differs slightly along both the facial identity and emotion axis. (b) *f*MRI responses for sample faces. Emotion vs. identity contrast revealed significantly stronger *f*MRI responses 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 [87] and the AAL atlas provided with MRIcro [88]($N = 10$)

2.5.3 Results

Subjects' accuracy during scanning was slightly better in the identity than in the emotion discrimination task (mean± 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±SEM: 831 ± 66 ms and 869± 71 ms for emotion and identity, respectively; $t_{(9)} = -1.49, p = 0.168$).

To asses the difference between the neural processing of the face stimuli in the emotion and identity discrimination tasks, we contrasted *f*MRI 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 $x, y, z = 43, -55, 7; t = 6.18, p < 0.01_{FDR}$, Fig.2.7b). 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

DOI:10.15774/PPKE.ITK.2010.001

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, $x, y, z = 51, 26, 7; t = 4.73, p < 0.01_{FDR}$) and in left inferior frontal gyrus (opercularis); (Br. 44, $x, y, z = -51, 14, 8; t = 4.53p < 0.01_{FDR}$) and finally, in left insula (Br. 13, $x, y, z = -36, 8, 13; t = 4.65p < 0.01_{FDR}$). 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 and colleagues [71]. 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 and colleagues [89], 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

DOI:10.15774/PPKE.ITK.2010.001

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$\pm SEM$: $82.2 \pm 1.7\%$ and $81.9 \pm 2.0\%$ 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 sig-

DOI:10.15774/PPKE.ITK.2010.001

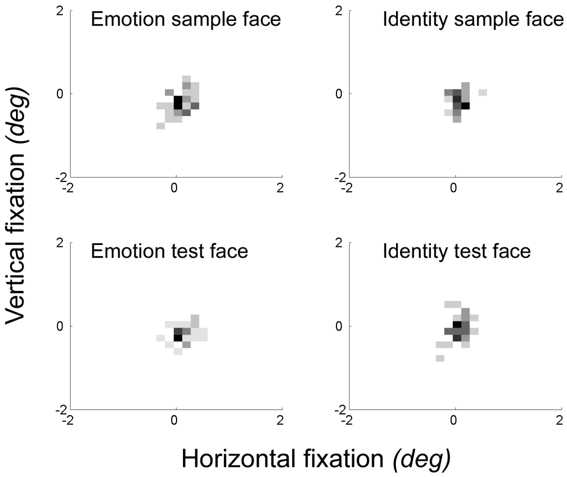

Figure 2.8: Representative fixation patterns of one subject during the practice session preceding Experiment 4, shown separately for emotion and identity discrimination trials recorded during sample and test face presentation. There was no difference between the fixation patterns for the two discrimination conditions neither during sample nor during test face presentation.

nificant 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 ($F_{(1,4)} = 0.040, p = 0.852$; See Fig.2.8 for a representative fixation pattern). 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 *f*MRI 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 [90] in response to emotion and identity discrimination tasks.)

2.6 Discussion

The results of the present series of experiments 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 seconds. More-

DOI:10.15774/PPKE.ITK.2010.001

over, Experiment 2 proved that this ability is not solely based on intermediate or low-level visual feature properties of the face images (e.g. local contour information, luminance intensity) but underpins the importance of high-level, face-specific attributes or attribute configurations, since performance was significantly reduced compared to the upright presentation when the face images were presented upside down. Furthermore, it was also found that discrimination of facial emotions was just as good when the observers perform the task only twice with novel faces as it was after extensive practice. Importantly, high-fidelity short term memory found in Experiment 3 cannot be accounted for by the criterion-setting theory proposed by Lages and Treisman [82] to explain results of delayed discrimination of basic visual dimensions [74]. 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 [82] 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 [39, 53], 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 have found that humans were 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

DOI:10.15774/PPKE.ITK.2010.001

anatomically segregated brain regions - as it was shown previously [39] 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 [91, 92, 93] that the visual association cortex, including the fusiform face area [20] and the lateral prefrontal cortex was 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 abstracted mnemonic information, while sensory areas represent specific features of the memoranda [93]. However, a recent fMRI study [71] 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 the 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 [39, 53]. 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 posses 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

DOI:10.15774/PPKE.ITK.2010.001

and receive prioritized processing [63, 64, 66]. 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 to 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 [67].

DOI:10.15774/PPKE.ITK.2010.001

Chapter 3

Retention interval affects VSTM processes for facial emotional expressions

3.1 Introduction

In the previous chapter we have shown that humans posses flawless, high-resolution visual short-term memory for emotional expressions [1]. 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 were impaired on visual working memory (WM) tasks only when information had to be stored for several seconds but no WM deficits were found in the same tasks when retention duration was very short, 1 s [94, 95]. Second, previous research investigating short-term memory processes using delayed discrimination of basic visual dimensions (e.g. spatial frequency

DOI:10.15774/PPKE.ITK.2010.001

and orientation: [72, 73, 83] as well as facial attributes: [1]) found a significant increase in reaction times (RT) at delays longer than 3 s as compared to 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 were absent at delays shorter than 3 s.

Previous research showed that short-term memory processes for facial attributes involved a distributed network of brain areas [91, 92, 93, 96, 94, 71] and that encoding and retrieval mechanisms were reflected in several ERP components evoked by the face stimuli [97, 98, 96, 99, 100, 101], including the early sensory P100 and N170 components as well as the late memory-related P3b wave complex. Surprisingly, however, the question whether the length of retention interval affected 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 processes for facial expressions (happiness) when the faces to be compared were separated by one or by six seconds. We recorded event related potentials (ERP) while participants performed the same delayed emotion discrimination task with a 1-s or a 6-s ISI and found that several ERP response components were strongly modulated by retention duration both during encoding and retrieval of facial emotional information.

3.2 Methods

Subjects. Nineteen (four left-handed, five females, mean age: 23 years) 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.

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) [75, 76, 77] by pairing a neutral and a happy picture of the same person, creating a set of intermediate facial

DOI:10.15774/PPKE.ITK.2010.001

expression images similarly to Experiment 1 in section 2.1 [1]. Four sets were created from black and white photos of four unknown actors (two females and two males). Stimuli (8 deg) were presented centrally (viewing distance of 60 cm) on a uniform grey 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 approximately 75% performance.

Procedure. Observers were shown face pairs, which were separated by a varying interstimulus interval (ISI). 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 long 6-s delay, since in an earlier pilot experiment we found that in a similar delayed-match-to-sample emotion discrimination task RTs saturate at 6-s delay, since no further increase in RTs was observed at delays longer than 6 s [1].

A trial consisted of 1600-1800 ms blank fixation followed by a temporal cue: a small circle surrounding the fixation cross, presented for 100 ms. 500 ms following the cue offset, the sample face was displayed for 300 ms. Subsequently, with either a short (1 s) or a long (6 s) delay showing only the fixation cross, the test face appeared for 300 ms. The test face was also preceded by a temporal cue, with a similar timing to that of the sample face (Fig.3.1). Introducing a cue before the test face was aimed at eliminating the temporal uncertainty present in the 6-s ISI condition, which - according to our pilot study - can confound reaction time measurements. Cueing of the sample face was also needed, in turn, for better comparison of the neural responses arising to the two face stimuli. Trials were initiated by the observers by pressing either of the response buttons with their right index or middle finger. The two ISIs were given in separate blocks, the order of which was randomized across subjects. Each participant completed four blocks (containing 256 trials altogether) of each ISI condition. Prior to the experiment each participant was given a practice session

DOI:10.15774/PPKE.ITK.2010.001

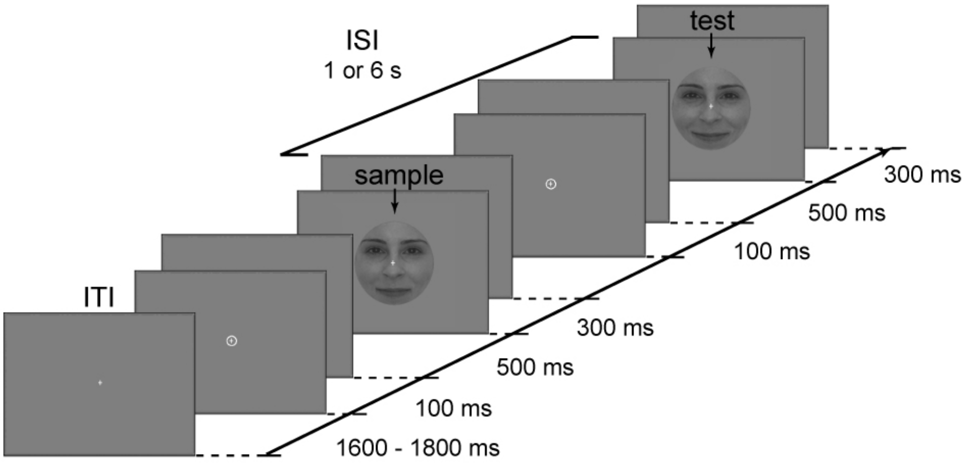

Figure 3.1: Delayed-match-to-sample paradigm. Two faces (sample and test) were displayed for 300 ms with either 1-s or 6-s interstimulus interval (ISI). Participants indicated by button press which of the two faces had a happier expression. Each face was preceded by a temporal cue to reduce timing uncertainty.

to familiarize themselves with the task and to determine the exact face pairs to be used. In all experiments the stimulus presentation was controlled by Matlab 7.1. (The MathWorks Inc, Natick, MA) using the Psychtoolbox 2.54 [85, 86].

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-s and 6-s ISI conditions were analyzed with paired t-tests.

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 in an EasyCap (Easycap GmbH, Herrsching-Breitbrunn, Germany) with four additional periocular 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 online and later digitally transformed to an average reference; the ground was placed on the

DOI:10.15774/PPKE.ITK.2010.001

nasion. All input impedance was kept below $10k\Omega$. Data were sampled at 1000 Hz with an analog bandpass 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 drifts, 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 $\pm 100\mu V$, or electro-oculogram activity exceeding $\pm 70\mu V$ were rejected. Data processing was done using BrainVision Analyzer (Brainproducts GmbH., Munich, Germany) and custom written Matlab software.

ERP Data Analysis. The trial-averaged EEG waveform - i.e. the event-related potential (ERP) - was computed including correct trials only. Data was segmented into 700 ms epochs separately for sample and test faces starting from 100 ms preceding the stimuli. 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, since it constitutes a single reference-free measure of response strength [102] and its local maxima reflect components in the event-related potential [102, 103]. The GFP was computed as the standard deviation of all electrodes at each time point, for each subject and condition and submitted to statistical evaluation.

Source Localization. An equivalent dipole source model was determined by the Brain Electric Source Analysis program (BESA 5.2, MEGIS Software GmbH, Gräfelfing, Germany) using dipolar inverse modeling and a four-shell ellipsoidal head model approximation registered to Talairach space for forward modeling. In principle, a source model is derived by fitting the source iteratively to the data until a minimum in residual variance (RV) is reached, i.e. the percentage of variance in the recorded potential distribution not accounted for by the source model is minimized. Beside the RV criterion, the energy criterion was also applied for fitting thus favoring source solutions with relatively low dipole strengths [104] to reduce the probability of interacting dipoles (i.e. adjacent dipoles with opposing high-amplitude potential fields). Regularization constant was set to 1. Symmetry constraints with respect to location were ap-

DOI:10.15774/PPKE.ITK.2010.001

plied to dipole pairs in order to limit the number of parameters estimated. No other constraints with respect to localization were used. Goodness of fit was estimated by the resultant RV.

Sources were localized on the overall grand mean ERPs averaged across subjects and conditions (2 ISI × 2 Task phase: sample and test) to avoid introducing a bias into later statistical comparisons between different conditions. Pairs of equivalent dipoles were fitted to time windows around the local maxima of the global field power (GFP) of the overall average ERP signal using the sequential multiple source fitting technique until 5% RV was reached. The exact length and borders of each window were determined by looking at the principal component decomposition of the residual waveforms and their 3D topographical maps, trying to maximize the amount of variance the first principal component explained. Each source pair was oriented at the maximum amplitude of the corresponding time window. The resulting inverse solution was then applied to the trial-averaged ERPs of each condition and each subject to determine the individual source waveforms which describe the amplitude (i.e. the strength) of the sources over time. These provided an image of brain function in terms of magnitude and timing of the underlying source currents.

Statistical Analysis. To test for significant differences between the ISI conditions for both sample and test phases, 90% confidence intervals for the GFP difference and the difference waves on selected electrodes and sources were determined using the bootstrap bias-corrected and adjusted (BCa) method [105]. This method provides confidence intervals without the necessity of prior assumptions about the distribution of the individual waveforms, e.g., a Gaussian distribution. The bootstrap statistics was conducted separately for sample and test stimuli. The difference between 6-s and 1-s ISI conditions was considered significant if the confidence interval of the difference source wave/GFP difference did not include zero [106, 96, 107].

Nonparametric tests were performed to further investigate the electrophysiological data which also do not assume Gaussian distribution of the data. Latencies of the late ERP peaks to test faces were tested with Wilcoxon matched pairs tests, the nonparametric equivalent of the paired t-test. Peak latencies were measured on the GFP waveforms at the corresponding maxima [108, 103].

DOI:10.15774/PPKE.ITK.2010.001

2 × 2 repeated measures analyses of variance were performed on ranked data of the mean intensities of PFC and MTL sources separately with ISI (1 s vs. 6 s) and task phase (sample vs. test face) as within-subject factors. To investigate the relationship between subjects' performance and source intensities or electrode amplitudes nonparametric Spearman correlations were performed on mean intensities averaged over time windows centered on the components of interest (120-190, 250-340 and 400-540 ms for N170, 325 ms peak and P3b, respectively - see Results section). Single electrodes were chosen for the correlation where the topographies of the respective sources showed their maximum and minimum (for source scalp topographies see Fig.3).

3.3 Results

3.3.1 Behavioral Results

Behavioral data showed that in accordance with our previous results [1] ISI did not affect delayed facial emotion discrimination accuracy. Subjects performed just as well in the 6-s ISI condition (mean$\pm SEM$: 80.0 ± 1.4%) as they did in the 1-s ISI (79.7 ± 1.7%) 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 (mean$\pm SEM$: 853 ± 38 ms) compared to the 1-s ISI (803 ± 36 ms) condition ($t_{(16)} = -2.69, p = 0.016$, Fig.A.1a). The possibility that enhanced temporal uncertainty caused the increased RT in the LONG ISI condition can be excluded because of the presence of temporal cues in our study.

3.3.2 ERP Results

Scalp topographies of the ERP activity peaks are shown in Figure 3.2a, while mean ERP waveforms of both ISI conditions that were time locked separately to sample and test faces are presented in Figure 3.2b and 3.2c, respectively. Similar ERP components were identified on the basis of their peak latencies and scalp topographies for both 1-s and 6-s ISI conditions. These components were also reflected in maxima of the global field power waveform (GFP; see Fig.3.4). The first component elicited by the stimuli was the occipital P100, which peaked around 105 ms on electrodes O1/O2 and had a central negative counterpart over Cz. Then a prominent occipito-temporal negative deflection, the N170 component proceeded, giving minimum over P9/P10 and maximum over FCz (aka. Vertex Positive Peak; VPP) [109] 160 ms after stimulus onset. The next to follow was the P220 component around 225 ms peaking over occipito-parietal PO3/PO4 electrodes, which showed an associated temporal negative deflection on T7/T8. This was followed by two peak activities with similar scalp distribution around 325 ms and 420 ms: apparent as a negative deflection on the central fronto-polar sites (Fpz) and a respective positive counterpart over electrodes PO3/PO4. Finally, there was a late ERP peak around 525 ms, which had a more complex scalp distribution having a large positive deflection with a centro-parietal maximum at Pz and CPz and multiple negative peaks: bilaterally at the occipito-temporal sites of P9/P10 and a frontal negativity around Fpz. This later peak corresponded to the P3b peak of the late sustained positive wave known as P300 [97].

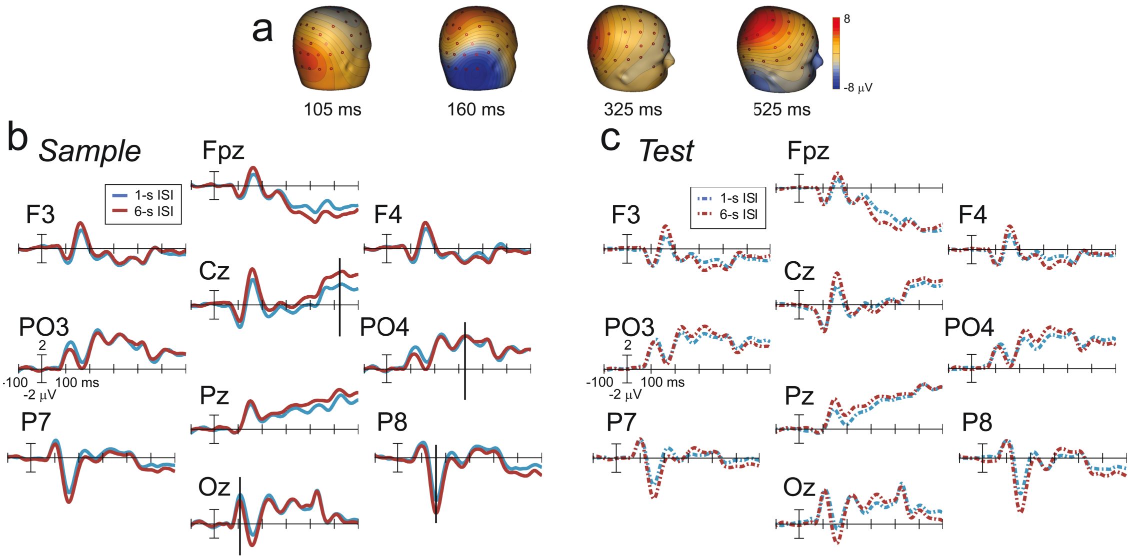

Figure 3.2: Grand average ERP responses and scalp topographies of components of interest. (a) Topographic maps of four ERP components averaged across the four conditions are shown on top with latencies indicated in parentheses, while ERPs are displayed separately for (b) sample faces (solid lines) and (c) test faces (dashed lines) with 1-s (blue) and 6-s(brown) ISI conditions overlaid. Vertical black lines indicate the timepoints corresponding to the topographic maps.

DOI:10.15774/PPKE.ITK.2010.001

	Fitting time	Name	Talairach coord.		
	ms		x	y	z
SP1	80-105	LO/pSTS	±40	−68	6
SP2	140-165	IT1	±36	−63	−11
SP3	165-195	IT2	±44	−48	−7
SP4	285-430	MTL	±17	−10	−14
SP5	440-540	PFC	±19	41	−2

Table 3.1: Fitting windows and Talairach coordinates for each source pair (SP). Rows follow fitting order.

3.3.3 Source location and activity

We used dipolar inverse modeling (BESA 5.2) to localize the generators of the observed averaged event-related potentials and to disentangle their specific contributions to the differential ERP responses between the two ISI conditions. Table 3.1 summarizes the detailed fitting information for each dipole pair of the equivalent dipole source model and their subsequent fitted location in Talairach coordinates. An inverse solution of five equivalent dipole pairs (Fig.3.3b) was able to explain 97.4% of the data variance throughout the 600 ms post-stimulus epoch. The five source pairs were localized in lateral occipital cortex, inferotemporal cortex (two source pairs), anterior medial temporal lobe and prefrontal cortex, and were termed LO/pSTS, IT1, IT2, MTL and PFC, respectively based on their corresponding anatomical region. Figure 3.3 shows the inverse solution and the derived source waveform and respective scalp topography of each source for sample and test faces separately (Fig.3.3a and 3.3c, respectively). Importantly, recent *f*MRI experiments have shown that these brain regions were part of the neural network underlying short-term memory for facial attributes [96, 71, 101],[1] and comparison of the source Talairach coordinates with activation maxima found in these studies demonstrated good correspondence.

3.3.4 Source fitting and description

The first dipole pair (LO/pSTS) was fitted between 75-105 ms post stimulus onset and was localized in lateral occipital cortex (LO, Fig.3.3). This was in

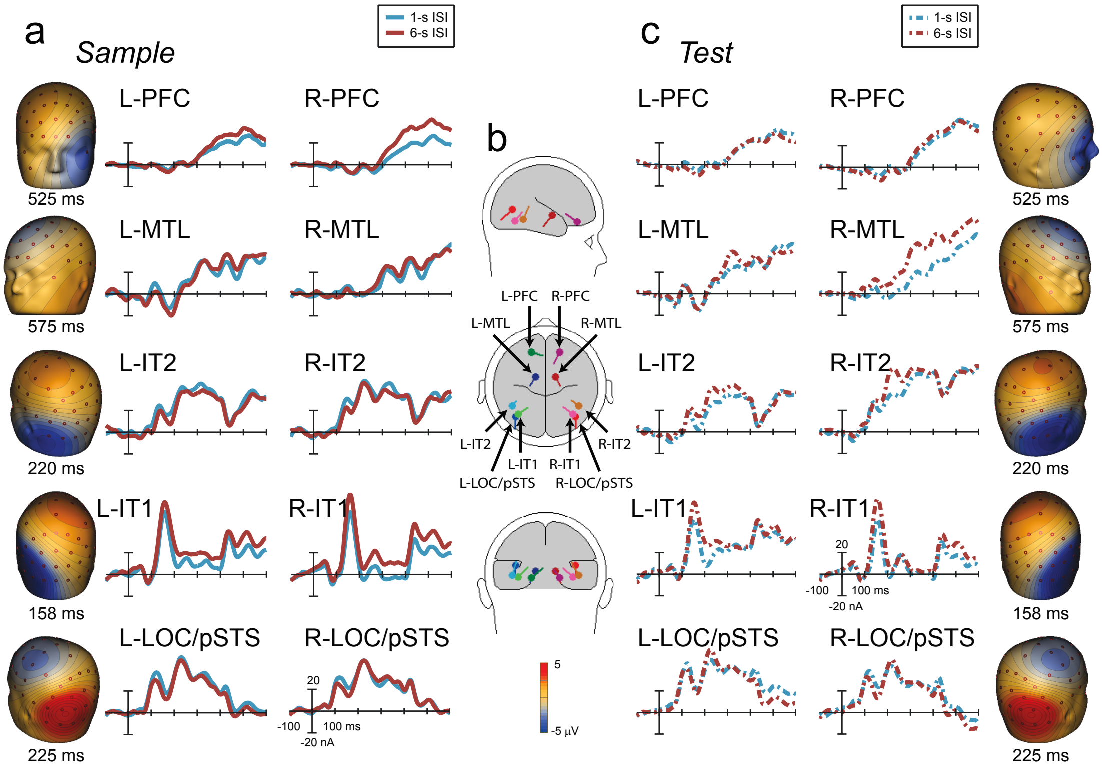

Figure 3.3: Location and time course of source activities. Source waveforms are displayed separately for (a) sample faces (solid lines) and (c) test faces (dashed lines) with 1-s (blue) and 6-s (brown) ISI conditions overlaid. Source waveforms are flanked by the topographical voltage maps of the respective sources generated with the forward model from the grand mean ERPs of all conditions taken at the time point of their maximum amplitude. Scalp topographies of left-side sources are shown on the left, right-side sources on the right. (b) The inverse solution is shown in the middle.

agreement with the previous results, showing that the main source of P100 component was localized in the lateral extrastriate cortex [110]. It is important to note, that even though the source contributing to the P100 component was localized in the LO, its activity might reflect neural processes of a larger cortical region surrounding the source location. Therefore, it is most likely that this source activity also integrated neural activity of an adjacent cortical region know to be involved in the processing of facial emotions, the posterior part of the superior temporal sulcus (pSTS; [84, 69, 51, 70, 52, 71],[1], which is located anterior and dorsal to the LO. Following the stimulus presentation this source pair was the first to activate, displaying an early transient activity

DOI:10.15774/PPKE.ITK.2010.001

peaking around 120 ms and a later transient peaking around 220 ms. Analysis of the scalp projections showed that the LO/pSTS source pair contributed to an occipital positivity (maximum on O1/O2) and a fronto-central negativity (strongest on FCz). The first transient activity accounted for the P100 component of the ERPs, while the second contributed to the P220 component, along with the IT2 source pair. Additionally, the LO/pSTS source pair also showed a late transient peak at 420 ms, which most likely represents the early peak of the stimulus offset response (120 ms after the offset of the face stimulus, i.e. 300+120 ms; for similar results see [96]).

The second and third source pairs were fitted in the time windows of 140-165 and 165-195 ms and both were localized in the inferotemporal cortex (IT1 and IT2, respectively). Despite their proximity, these source pairs greatly differed in their activation patterns, which is in agreement with previous findings [96]. The posterior IT1 source showed a robust positive transient reaching its maximum around 160 ms and a later, smaller activation peak around 440 ms, which again seems to reflect the offset response of the source pair and is in agreement with the findings of Bledowski and colleagues (2006). IT1 scalp projections were very similar to and largely responsible for the voltage map of the N170 ERP component, namely a bilateral occipito-temporal negativity (P9/P10) coupled with a mid-central positive deflection (Cz), which is in line with several other studies showing that the fusiform gyrus might be the primary source of the N170 component [27, 28, 29] but see: [23, 30]. This is also in good agreement with our current understanding of N170 as a component reflecting the processing step of face-specific, configural information [80, 24, 111] which takes place in the fusiform gyrus of the inferotemporal cortex [20, 16, 12]. Moreover, this early transient activity was somewhat larger in the right source, which supports the observation of right hemisphere dominance in face processing [22, 20, 111]. The anterior IT2 source, on the other hand, had a more sustained activation pattern, starting from 120 ms reaching its maximum by 220 ms post-stimulus onset. The scalp topography of this source pair closely resembled to that of IT1.

The last two source pairs were fitted from 285-430 and 440-540 ms and were localized in the anterior medial temporal lobe (MTL) and medial ventral prefrontal cortex (PFC), respectively. As opposed to the early and more transient activity of the LO/pSTS and IT source pairs, these source pairs displayed a

DOI:10.15774/PPKE.ITK.2010.001

more sustained activity slowly rising from approx. 250 ms post-stimulus onset. The MTL source pair was located in the anterior parahippocampal gyrus, which is in agreement with the results of fMRI studies showing activation in the parahippocampal/perirhinal cortex during fine discrimination and memory for complex objects, including faces [94, 71, 112]. In the case of the MTL source pair, there were two transient activity peaks superimposed on the sustained activity, which reached its maximum around 580 ms. The transient MTL source activity peaks and their scalp topography corresponded closely to the 325 ms and 420 ms peaks found on the scalp electrodes (Fig.3.3), namely an occipital positivity (Oz) and a frontal negativity (AFz). Based on its time course, one might suggest that the MTL source activity found in the current study using scalp EEG recordings might correspond to the memory related anterior medial temporal lobe activity described in a similar time window (also called AMTL-N400) in previous intracranial recording studies [113, 114, 115].

In the case of the PFC source pair, there was pronounced right hemisphere dominance. The peak activity of this source pair was observable around 525 ms and produced a mid-fronto-polar negativity (Fpz) and a centro-parietal positive deflection (CPz), which corresponded closely to the late positive peak found on the scalp electrodes. It appears that the PFC source is the major contributor to the P3b wave, which is a late sustained positive deflection peaking at CPz-Pz electrode sites, since both their onset and peak latency (270 ms and 525 ms, respectively) coincide. This is further supported by the results of a previous experiment investigating the effect of memory load [96], where they showed that WM load dependent modulation of the P3b complex originated from sources located in the ventrolateral prefrontal cortex. In addition to this, the PFC source pair is also likely to reflect orbitofrontal activity, which is implicated in emotional information processing and maintaining information in WM [116, 71].

3.3.5 Relationship between behavior and source activity

To further verify our source localization approach, we investigated the correlation between source activities and subjects' accuracy in the delayed facial emotion discrimination task. We reasoned that in case our source modeling was successful, the obtained source activities should reflect more accurately the

DOI:10.15774/PPKE.ITK.2010.001

neural activities of specific brain regions than the scalp potentials, which reflect a mixture of activity of many underlying cortical generators. If so, one might expect that subject's accuracy will correlate more closely with the source activities than with the scalp potentials.

We found positive correlation between subjects' performance and mean source activities in the right PFC and right MTL source (averaged in the 400-540 and 250-340 ms time windows, respectively) in both ISI conditions (see Table 3.2), whereas left PFC and left MTL mean source activities did not show any correlation with task performance. In the case of right PFC source activity, strong significant correlation was observed following the presentation of test stimuli (Spearman correlation: $r = 0.55, p = 0.023$ and $r = 0.77, p < 0.001$, for 1-s and 6-s ISI, respectively). In the case of sample stimuli, there was a marginally significant correlation between PFC source activity and task performance ($r = 0.48, p = 0.05$) only in the 6-s ISI, but not in the 1-s ISI condition ($r = 0.33, p = 0.20$). Importantly, correlation between subjects' performance and mean scalp electrode potentials measured on the CPz and Fpz electrodes in the 400-540 ms time window (representing the maximum and minimum of the late P3b peak activity, respectively) was less pronounced than the correlation

400-540 ms	Perf vs. S-1 s	Perf vs. T-1 s	Perf vs. S-6 s	Perf vs. T-6 s
R-PFC	0.33	**0.55***	**0.48***	**0.77*****
L-PFC	-0.29	0.16	0.07	0.24
Fpz	-0.24	-0.31	-0.42	**-0.76*****
Cpz	0.16	**0.48***	0.40	**0.48***
250-340 ms	Perf vs. S-1 s	Perf vs. T-1 s	Perf vs. S-6 s	Perf vs. T-6 s
R-MTL	**0.70****	**0.67****	**0.55***	**0.66****
L-MTL	-0.12	0.02	0.37	0.46
AFz	0.05	-0.07	-0.10	-0.41
Oz	-0.47	-0.47	-0.45	-0.46

Table 3.2: Spearman correlations (rho) between subjects' performance and the mean PFC/MTL source activities in the given time windows. For comparison, correlations between performance and the time courses of corresponding electrodes are also shown (S = sample face, T = test face; $*p < 0.05$; $**p < 0.01$; $***p < 0.001$)

DOI:10.15774/PPKE.ITK.2010.001

between task performance and source activity (Table 3.2).

We also found strong correlation between subjects' performance and mean right MTL source activity both in the case of sample ($r = 0.70, p = 0.001$ and $r = 0.55, p = 0.023$, for 1-s and 6 s-ISI, respectively) and test face stimuli ($r = 0.67, p = 0.003$ and $r = 0.66, p = 0.004$, for 1-s and 6-s ISI, respectively). This is in agreement with recent fMRI studies showing that medial temporal lobe areas might be involved both in working memory encoding and retrieval processes [71, 101]. Importantly, however, we found no significant correlation between task performance and mean scalp potentials measured on Oz and AFz electrodes in the 250-340 ms time window (representing the maximum and minimum of the 325 ms peak activity, respectively; Table 3.2). These findings provide support for our source modeling results and suggest that source activity might reflect the neural processes more accurately than scalp potentials, especially in the case of deep cortical generators.

3.3.6 Effect of ISI on the neural responses to sample faces

Bootstrap statistics revealed that mean GFP waveforms differed significantly between the 1 and 6 s ISI conditions (Fig.3.4a) in the case of sample faces. ISI affected GFP (larger GFP in 1-s ISI than in 6-s ISI condition) already in a very early time window, peaking around 110 ms, which corresponded to the P100 ERP component. In accordance with this, activity of the scalp electrodes and sources related to the P100 component was significantly higher in the 1-s ISI than in the 6-s ISI condition (Fig.3.4a). On the other hand, GFP was stronger in the 6-s ISI than in the 1-s ISI condition between 140-200 ms; in a time interval corresponding to the N170 ERP component. Activity of the scalp electrodes and sources related to the N170 component was modulated by ISI correspondingly, being larger in the 6-s ISI than in the 1-s ISI condition. Furthermore, GFP was also larger in the 6-s ISI than in the 1-s ISI condition in a later time interval, lasting from 375-600 ms, corresponding to the P3b component. Source localization results suggested that this late modulation of ERP responses by ISI might originate from the right ventral prefrontal cortex, since significantly higher source activity in the 6-s ISI compared with the 1-s ISI

DOI:10.15774/PPKE.ITK.2010.001

condition was found in right PFC source in the time window starting around 300 ms and extending until the end of the analyzed 600 ms epochs.

3.3.7 Effect of ISI on the neural responses to test faces

Bootstrap statistics revealed significantly higher GFP in the 6-s ISI than in the 1-s ISI condition in a time interval corresponding to the N170 ERP component (Fig.3.4b) in the case of test faces. Source modeling suggested that modulation of these early components by ISI might be due to strongly reduced transient inferior-temporal cortical neural responses in the 1-s ISI compared to the 6-s ISI condition, which was reflected in the IT1 source activity. Furthermore, significantly higher GFP in the 6-s ISI than in the 1-s ISI condition was also found in a later time window, peaking around 325 ms. Source localization results suggested that this later modulation of ERP responses by ISI might primarily originate from the right anterior MTL (Fig.3.4b). This source showed a sustained activity difference: right MTL activity was significantly higher in the 6-s ISI than in the 1-s ISI condition in the time windows of 200-600 ms. This activity difference translated onto a scalp topography with a frontal negative (peaking at AFz) and an occipital positive deflection (peaking at Oz), which closely matched the ERP scalp map difference starting from 275 ms. Moreover, this ERP difference first peaked around 325 ms, which coincided with the first transient of the MTL source pair.

However, it is important to note that due to the short inter-stimulus interval, 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 ($< 1s$) presentation of an adaptor face stimulus could modulate the neural response to the subsequently presented test face stimulus, leading to reduced N170 [77] as well as M170 amplitude [117], 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 ($< 1s$) [117] we could not exclude the possibility that early sensory components of the ERP responses to the test faces, in particular the N170 amplitudes were affected by adaptation processes in the 1-s ISI condition. Thus, the observed difference in the amplitudes of the

DOI:10.15774/PPKE.ITK.2010.001

early ERP response components between the 1-s and 6-s ISI conditions might be at least partly due to the adaptation effects present in the former case.

Examination of the latencies of late ERP peaks found in response to the

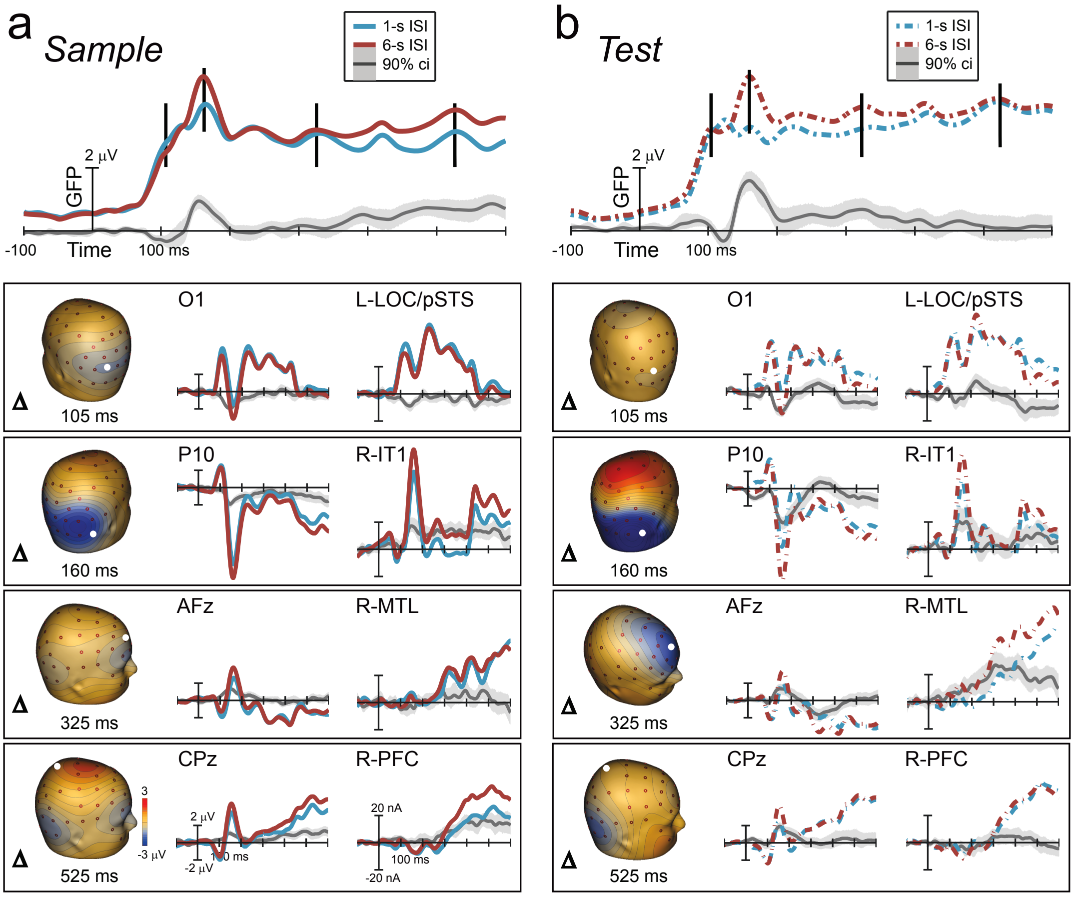

Figure 3.4: Delay effects observed in memory (a) encoding and (b) retrieval. The prefrontal cortical (PFC) source showed a significant modulation by delay length during encoding, while in the case of retrieval the medial temporal lobe (MTL) source was significantly more active in the long delay condition compared with the short delay. Mean global field power (GFP) waveforms at 1-s and 6-s delays in the case of sample (solid blue and brown lines, respectively) and test faces (dashed blue and brown lines, respectively). GFP maxima of interest are marked with black lines and the corresponding topographic difference voltage maps of the 6-s vs. 1-s ISI conditions are depicted underneath with the ERPs on representative electrodes. Alongside are the time courses of possible cortical generators. Electrode locations are marked on scalp topographies with white dots. On all time courses gray bands indicate 90% BCa bootstrap derived confidence intervals of the 6-s vs. 1-s delay differences (dark gray lines).

DOI:10.15774/PPKE.ITK.2010.001

presentation of test faces failed to reveal any difference between the 1-s and 6-s ISI conditions (Wilcoxon matched pairs test: $Z_{(17)} = 1.73, p = 0.083$ (with opposite trend to that of the RT increase) and $Z_{(17)} = 0.28, p = 0.78$ for the 325 ms peak and P3b, respectively). These results were surprising, since consistent with earlier findings, our behavioral results showed a significant increase in RTs in the 6-s ISI compared to the 1-s ISI condition. Further research is required to uncover the neural processes leading to longer RTs in conditions where sample and test faces are separated by several seconds (See AppendixA).

3.3.8 Modulation of PFC and MTL source activity by ISI

An important finding of the current study is that right PFC and right MTL source activities were modulated differently by ISI: PFC, but not MTL source activity was affected by ISI during encoding, whereas MTL, but not PFC source activity was modulated by ISI during retrieval (Fig.3.5). In the case of the right PFC source the ISI effect was apparent as a decreased activity for sample faces in the 1-s ISI condition as compared to the other conditions (Fig.3.5a). This was supported by the results of a repeated measures ANOVA - computed on the

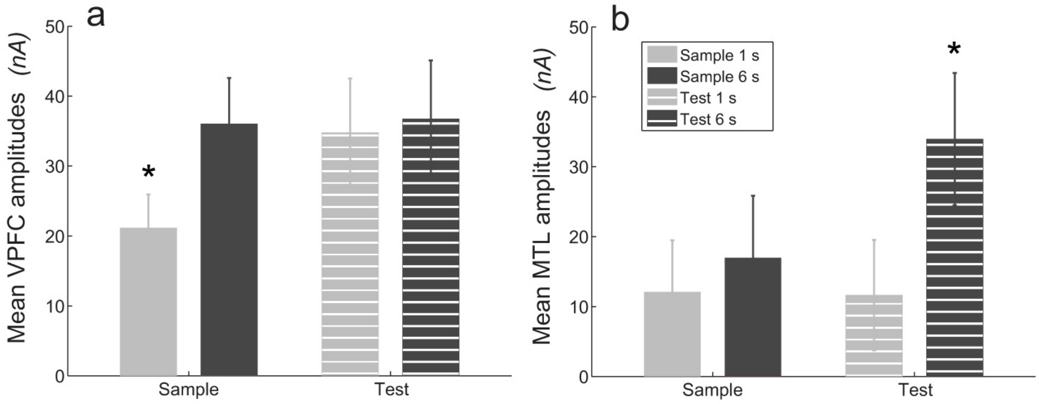

Figure 3.5: Mean PFC (a) and MTL (b) source intensities for sample and test faces in both ISI conditions averaged over a time window of 400-540 and 250-340 ms, respectively. PFC source activity was significantly reduced during encoding in the case of sample faces relative to the other cases, while MTL source activity was significantly enhanced during retrieval following the long delay of 6 s. S = sample face, T = test face. Error bars indicate $\pm SEM$. Asterisks indicate significant difference.

DOI:10.15774/PPKE.ITK.2010.001

ranked data of the right PFC source mean intensities averaged over a 400-540 ms time bin - showing a significant main effect of ISI ($F_{(1,16)} = 9.21, p = 0.008$), no main effect of task phase ($F_{(1,16)} = 0.54, p = 0.47$) and a significant interaction between these variables ($F_{(1,16)} = 7.34, p = 0.015$).

Right MTL source activity on the other hand, exhibited a different pattern of ISI modulation: MTL source activity in the case of the test faces in the 6-s ISI condition was higher than in the other conditions (Fig.3.5b). This was also supported by the results of the ANOVA on MTL mean amplitudes averaged over a time bin of 250-340 ms showing a significant main effect of ISI ($F_{(1,16)} = 6.43, p = 0.022$), a significant main effect of task phase ($F_{(1,16)} = 5.24, p = 0.036$) and a significant interaction between these factors ($F_{(1,16)} = 5.80, p = 0.028$). These results suggest that the observed significant difference in the MTL source activity between the 6-s and 1-s ISI conditions in the case of test faces is due to the increased MTL source activity in response to test faces in the 6-s ISI condition, rather than due to a decrease of this source activity in the 1-s ISI condition. This implies that the difference in MTL source activities between the 1-s and 6-s ISI conditions cannot be explained based on adaptation processes, which would lead to reduced neural responses in the case of test faces in the 1-s ISI condition compared to the other conditions. Furthermore, in the case of test faces there was no correlation between the magnitude of retention duration effect (i.e. the mean intensity difference in the time window of 250-340 ms) found on the right MTL source activity and the magnitude of retention interval effects found on the activity of visual cortical sources (R-IT1 and R-IT2) corresponding to the N170 component measured in the time window of 120-190 ms (Spearman correlation: $r = -0.03, p = 0.91$ and $r = 0.18, p = 0.48$ for R-IT1 and R-IT2, respectively). This speaks against the possibility that the same adaptation processes underlie the modulation of these two components or that N170 adaptation effects might be simply carried over to the later 325 peak. Based on these results we argue that the observed difference in the MTL source activity between the 1-s and 6-s ISI conditions in the case of test faces reflects modulation of the retrieval processes by retention interval.

3.4 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-s and 6-s ISI conditions, showing that early sensory processing was 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 alpha band), the amplitude and the inter-trial phase synchronization of which is known to affect the P100 component (for review see [118, 119]). Furthermore, using a delayed matching to sample task it was shown [99, 100] that increasing working memory load - i.e. the number of complex visual objects or faces that need to be compared - lead to increased P100 amplitudes in the ERP responses to the sample stimuli. In addition, the study by Morgan et al. [100] revealed that WM load also modulated 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 amplitude in the ERP responses to the sample face stimuli and reduced N170 amplitudes for 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 [99, 100] 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. 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 350 ms and peaking at 525 ms - corresponding to the later peak of the previously described memory-related positive ERP component, known as P3b [97, 98, 96, 100]. Our source localization

DOI:10.15774/PPKE.ITK.2010.001

results suggested that the effects of retention duration on the late P3b component might primarily originate from the prefrontal cortex. This is in agreement with previous findings showing that modulation of the P3b component by WM load originated from the prefrontal cortex [96]. 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 [96, 120, 100]. Our results revealed that the P3b component was reduced in responses to sample faces in the 1-s delay condition as compared to the P3b evoked by the sample faces in 6-s delay condition as well as compared to the P3b evoked by the test stimuli. Furthermore it was also found that subjects' performance in the delayed emotion discrimination task correlated with the right prefrontal source activity in response to the test stimuli as well as to the sample stimuli at 6-s delay but not with the prefrontal source activity evoked by the sample stimuli at 1-s delay. A possible explanation of these results might be that at 1-s delay representation of the sample face information is maintained online via persistent delay period activity (for review see [81]) of a neural network involving the visual cortical areas specialized for face processing and thus processing of sample faces does not involve memory search and matching processes reflected in the P3b component. On the other hand, at 6-s delay online 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 delays would involve memory search and matching processes and thus will lead to increased P3b component.

We found significant modulation of the ERP responses by retention duration also in the case of test face stimuli: ERP responses were significantly stronger at 6-s delay than at 1-s delay in a time window peaking at 325 ms. Source localization suggested that the modulation of this component by storage duration might primarily originate from the anterior medial temporal lobe, where source activity was strongly increased in the 6-s ISI condition. We also found that the right MTL source activity preceding the 325 ms peak strongly correlated with subjects' performance in the delayed happiness discrimination task both at 1-s and 6-s delays. Importantly, we found no correlation between task performance and scalp potentials on corresponding electrodes in the same time window, sug-

DOI:10.15774/PPKE.ITK.2010.001

gesting that our source modeling was successful and the obtained source activities provided a more accurate estimate of neural activity than scalp potentials. Furthermore, involvement of anterior MTL in face processing is supported by several lines of evidence coming form previous studies using intracranial ERP recordings [113, 114, 115] as well as fMRI [94, 89, 71, 112]. In particular, it was found that tasks requiring detailed processing of facial attributes invoked neural processes of several anterior MTL regions, including the hippocampus, the amygdala, the perirhinal cortex and the temporal pole. Most relevant for the current study, intracranial ERP recordings investigating the time course of facial recognition revealed a cascade of MTL activity in the 200- to 600-ms time frame (for review see [115]). The first component, also called the face AMTL-N400 [113, 114], consisted of more than one peaks in the 200-400 ms time window and originated from the perirhinal cortex and the temporal pole. It was accompanied by a slow positive hippocampal component, starting approximately in the same time but peaking later than 400 ms. Thus, based on these intracranial ERP results one might suggest that the MTL source activity found in the current study reflects the combined activity of several anterior MTL regions present in the 200-600 ms time window after stimulus onset. If so, the results of the current study suggest that the retrieval phase of delayed discrimination of facial emotions at 6-s delay might be based on MTL processes to a larger extent than that of at 1-s delay. Importantly, such stronger involvement of the MTL in retrieval processes with increased retention duration would be consistent with the results of previous neuropsychological research on brain lesioned patients, showing 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 [94, 95].

DOI:10.15774/PPKE.ITK.2010.001

Chapter 4

Conclusions and possible applications

The findings of the above series of studies revealed 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 depression, autism and schizophrenia [121, 122, 123].

Furthermore, our results also showed that retention interval affected short-term memory processes for facial emotions. We found that in a delayed emotion discrimination task both encoding and retrieval processes differed when the faces to be compared were separated by several seconds from that when they were presented with a very short, 1-s delay. 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 both encoding and retrieval of facial emotional information reflect changes in mnemonic processes as a function of storage duration and cannot be accounted for by differences in sensory processing demands or overall task difficulty across the conditions with different retention intervals. The results of

DOI:10.15774/PPKE.ITK.2010.001

the present study thus provide the first evidence that different neural encoding and retrieval processes underlie flawless, high-resolution short-term memory for facial emotional expressions 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 [124]) - should be revised to include retention interval as an important factor affecting neural processes of memory encoding.

DOI:10.15774/PPKE.ITK.2010.001

Chapter 5

Summary

5.1 New scientific results

1. Thesis: I have characterized the efficiency of short-term memory storage of different facial attributes, namely emotional expression and identity in a facial attribute discrimination task revealing a high-fidelity storage for both. Furthermore, I have proved that this storage is based on holistic processing of faces focused on the given attribute as opposed to the mere processing of local features.

Published in [1], [3].

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 contrast, there are facial attributes - such as identity or gender - that on the short and intermediate timescale are invariant [39, 53]. Therefore, invariant facial attributes do not require constant online monitoring during social interaction. Using a two interval forced choice facial attribute discrimination task we measured how increasing the delay between the subsequently presented face stimuli affected facial emotion and facial identity discrimination.

DOI:10.15774/PPKE.ITK.2010.001

1.1. I have shown that people posses a high-fidelity visual short-term memory for facial emotional expression and identity, since 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, in contrast to my prediction, I found no significant effect of increasing the delay between the sample and the test face in the case of facial identity discrimination.

Observers performed delayed discrimination of three different facial attributes: happiness, fear and identity. In all three discrimination conditions reaction times were longer by approximately 150-200 ms in the 6-s delay than in the 1-s delay conditions providing support for the involvement of short term memory processes in delayed facial attribute discrimination in the case of 6-s delay condition. Increasing the delay between the face images to be compared had only a small non-significant effect on observers' performance in the identity discrimination condition revealed by a slight increase in the just noticeable difference (JND) value between the two faces. On the other hand, discrimination of facial emotions was not affected by the delay (Fig.2.4). These results suggest that fine-grained information about both facial emotions and identity can be stored with high precision, without any loss in VSTM.

1.2. Furthermore, I have corroborated my findings on a large sample size of 160 subjects revealing flawless short-term memory for both facial emotions and facial identity also when the discrimination task was performed with novel faces. I have also shown that practice and familiarity of faces affected performance in the facial identity discrimination task but not in the facial emotion discrimination task, which did not require learning.

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 we conducted an experiment, where 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 a 1-s delay while for the other half of participants the delay was 10 s. The results revealed that sub-

DOI:10.15774/PPKE.ITK.2010.001

jects' emotion and identity discrimination performance was not affected by the delay between the face stimuli to be compared, even though the faces were novel (Fig.2.6). It also excludes the possibility that the present attribute discrimination 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 criterion-setting theory [82]. Instead it is based on the perceptual memory representation of the sample stimulus similarly to other delayed discrimination tasks [74, 81].

> *1.3. I have confirmed that the discrimination performance depended on holistic facial processing and could not be based solely on the processing of local features. In accordance with this I have proved that discrimination of fine-grained emotional expressions involved processing of high-level facial emotional attributes.*

It was crucial to show that performance in our facial attribute discrimination task was indeed based on high-level, face-specific attributes or attribute configurations as opposed to some intermediate or low level feature properties of the face images (e.g. local contour information, luminance intensity). In an experiment with the face stimuli presented in an inverted position where the configural feature was taken away leaving the low level features unaltered [80, 16] we found a significant drop in performance for all three attributes (i.e. increased JND values), which was most pronounced for upside-down fear discrimination, therefore proving holistic processing (Fig.2.5).

Furthermore, we have confirmed that the emotion discrimination in our short-term memory paradigm involved high-level processing of facial emotional attributes by performing an fMRI experiment and contrasting trials where subjects made decisions based on the emotional content of the stimuli with those based on identity content. We found no brain regions where activation was higher in the identity compared to the emotion discrimination condition. However, our analysis revealed significantly higher activations in the case of emotion compared to identity discrimination in the right posterior superior temporal sulcus (pSTS) among others (Fig.2.7). This is in agreement with previous studies showing that increased fMRI responses in the pSTS during tasks requiring perceptual responses to facial emotions compared to those to facial identity can

DOI:10.15774/PPKE.ITK.2010.001

be considered as a marker for processing of emotion-related facial information [69, 71].

2. Thesis: I have shown that different neural mechanisms underlay high-fidelity short-term memory for emotional expressions depending on whether information had to be stored for one or for several seconds, which has been an unresolved question. This result was not confounded by differences in sensory processing demands or overall task difficulty which otherwise might offer alternative explanations to the above findings.

Published in [2], [3].

Previous research has implicated that during VSTM tasks encoding and retrieval processes were changing depending on how long the information has to be stored in VSTM. Patients with medial temporal lobe (MTL) lesions were impaired on VSTM tasks only when information had to be stored for several seconds but no VSTM deficits were found in the same tasks when retention duration was very short, 1 s [94, 95]. Moreover, studies investigating delayed discrimination of basic visual dimensions [72, 73, 83] found a significant increase in reaction times (RT) at delays longer than 3 s as compared to shorter, 1-s delays. Therefore, the goal of the present study was to directly compare short-term memory processes for facial expressions (happiness) when the faces to be compared were separated by one or by six seconds. We recorded event related potentials (ERP) while participants performed the same delayed emotion discrimination task with a 1-s or a 6-s ISI and found that several ERP response components are strongly modulated by retention duration both during encoding and retrieval of facial emotional information.

> *2.1. During encoding I found significant differences in the early P100 and N170 components of the event-related potential (ERP) responses between the delay conditions in a delayed emotion discrimination task. Furthermore, the memory-related late P3b ERP component was significantly reduced in the 1-s delay condition, which based on source-modeling results primarily originated from the right prefrontal cortex.*

DOI:10.15774/PPKE.ITK.2010.001

Delay affected global field power (GFP) activity in the encoding phase already in a very early time window, peaking around 110 ms, corresponding to the P100 ERP component: activity of the scalp electrodes and sources related to the P100 component was significantly higher in the 1-s delay than in the 6-s delay condition (Fig.3.4a). On the other hand, GFP was higher in the 6-s delay than in the 1-s delay condition between 140-200 ms, in a time interval corresponding to the N170 ERP component. Activity of the scalp electrodes and sources related to N170 was modulated by delay correspondingly, being larger at 6 s than at 1 s delay.

Furthermore, GFP was also larger in the 6-s delay than in the 1-s delay condition in a later time interval, lasting from 375-600 ms, corresponding to the P3b component. This modulation pattern was consistent only with one source in our source model, namely with right ventral prefrontal cortex (PFC), which suggests PFC as the primary generator of this signal difference. This was also consistent with other findings [71] showing that modulation of the P3b component by memory load originated from the prefrontal cortex.

> *2.2. During retrieval, on the other hand, I found significantly stronger anterior medial temporal lobe (MTL) source activity peaking at 325 ms in the 6-s compared to the 1-s delay condition, which is evidence for MTL - possibly hippocampus - involvement during memory retrieval if faces are stored for several seconds.*

Bootstrap statistics revealed significantly higher GFP in the 6-s delay than in the 1-s delay condition in a time interval corresponding to the N170 ERP component (Fig.3.4b) in the case of the retrieval phase. Source modeling suggested this modulation might have been due to a strongly reduced transient inferior-temporal cortical neural response at 1-s delay as a possible contamination by adaptation effect. However, significantly higher GFP in the 6-s delay than in the 1-s delay condition was also found in a later time window, peaking around 325 ms. Source localization results suggested that this later delay-related modulation might have primarily originated from the right anterior MTL. The involvement of anterior MTL structures in face processing is supported by several lines of evidence coming form previous studies using intracranial ERP recordings [113, 114] as well as *f*MRI [94, 71, 112]. The results

DOI:10.15774/PPKE.ITK.2010.001

of the current study suggest that the retrieval phase of delayed discrimination of facial emotions at 6-s delay might be based on MTL processes to a larger extent than that of at 1-s delay, which is in line with several neuropsychologic studies [94, 95].

DOI:10.15774/PPKE.ITK.2010.001

Appendix A

Possible source of reaction time increase

A.1 Lateralized Readiness Potential analysis

In the study in chapter 3 we reported longer reaction times in the 6-s ISI condition compared with the 1-s ISI condition, which is in agreement with our previous experiments [1] described in chapter 2 and several other experiments using delayed discrimination tasks to investigate visual short-term memory for basic visual dimensions [72, 73, 74]. However, we did not find any difference between the 1-s and 6-s ISI conditions in the latencies of late ERP peaks found in response to the presentation of test faces questioning that the source of the RT difference stems from sensory or decision-related processes. On the other hand, this raises the possibility that the RT difference might be due to differences in the motor processes between the two conditions.

To tease apart the possible contribution of sensory and motor-related processes to the RT increase, we calculated the lateralized readiness potential (LRP), which provides a particularly suited psychophysiological measure to localize experimental effects on RT (for a review see [125, 126]). The LRP is the difference signal of electrodes contralateral versus ipsilateral to the responding hand and its onset is believed to reflect the moment in time when response specific processes become centrally active (IN: [127]). This onset can be measured relative to the onset of the stimulus which elicits the response (S-locked)

DOI:10.15774/PPKE.ITK.2010.001

or to the onset of the overt response (R-locked). The interval between stimulus onset and the S-locked LRP onset - denoted as S-LRP interval - estimates the duration of RT processes that occur before the onset of response activation. By contrast, the R-LRP interval - the interval between the R-locked LRP onset and the onset of the overt response - estimates the duration of the RT process that occur after the beginning of response activation. By examining which of these two intervals is affected by an experimental manipulation, one can determine whether the manipulation affects the duration of stimulus processing and response selection or response preparation and motor execution [128, 129, 130]. In general, any manipulation that affects only sensory processes will cause a delay in the onset of the stimulus-locked LRP, but no change in the onset of the response-locked LRP. On the contrary, any manipulation that affects the preparation of responses will cause a change in the onset of the response-locked LRP, but no change in the onset of the stimulus-locked LRP. The manipulation effect will be referred to as "stimulus-locked effect" in the former and "response-locked effect" in the latter case.

A.2 Methods

The lateralized readiness potential (LRP) measures lateralized motor preparatory activity and is traditionally computed by subtracting the activity of the electrode over motor cortex ipsilateral to the responding hand from the electrode contralateral to it: $(C3-C4)_{right\ hand\ movement}$ [125]. This makes explicit the lateralized negativity that precedes a movement. In conventional LRP paradigms this is done for both hands and then averaged to eliminate any non-motor related asymmetrical activity. Since in our paradigm subjects were required to respond with their right hand only, the latter step is omitted. Segments starting from -200 until 1000 ms relative to test faces were averaged to obtain the stimulus-locked waveforms while segments from -1000 to 200 ms relative to the overt response were averaged to obtain the response-locked LRPs. In both cases baseline correction was based on the first 200 ms of each segment. To identify the onset of the LRP in the average waveform a regression-based method was used which defines the onset of the LRP as the "break-point" between two intersecting straight lines that are fit to the LRP waveform [131, 132]. In general, one of the lines is fit to the putative preonset segment of the LRP, whereas another line is fit to the falling segment of the LRP dip. In both S- and R-locked LRPs the preonset segment started from test face presentation until the observable bend in the trace and was restricted to positive slopes. (Unlike in our study, LRP curves are traditionally depicted as positive slope deflections, since conventional ERP curves show negative up; therefore if restricted, the slope of the preonset segment is limited to negative slopes to avoid biasing the onset time towards later latencies.) The second segment was fit from the time of the

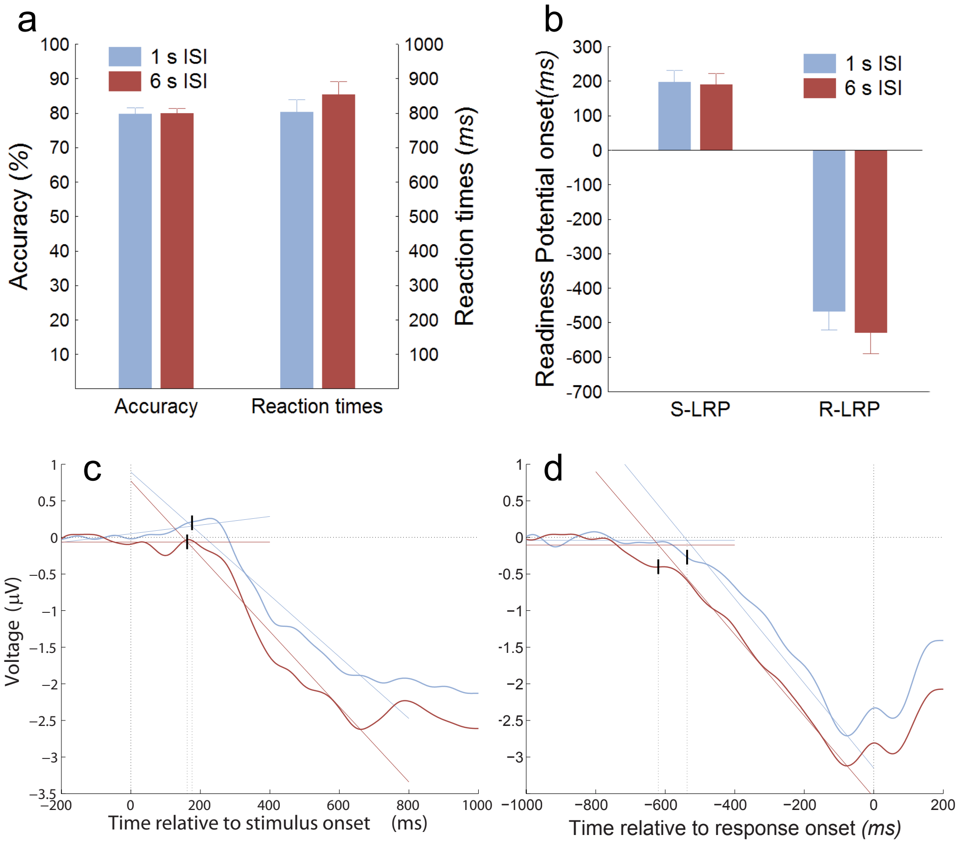

Figure A.1: Behavioral measures and results of the Lateralized Readiness Potential (LRP) analysis. (a) Unlike accuracy which was similar in both ISI conditions, reaction times were significantly longer in the LONG ISI condition. However, this difference most likely did not reflect memory processes, since (b) the results of the LRP analysis revealed a significant response-locked effect. (c) The onset times of stimulus-locked LRPs of the two ISI conditions did not differ, while (d) the onset of the response-locked LRP in the LONG ISI was significantly earlier than the respective onset in the SHORT ISI condition. Their difference was comparable to that of measured in the RT.

bend until the peak and was locked at the time and height of the peak. The time of the intersection was taken as the estimate of the LRP onset latency. The lines were fit to individual average waveforms using least-square optimization in Matlab and onset latencies were compared using a paired t-test.

A.3 Results

Figure A.1c-d shows the stimulus- and response-locked grand average LRPs for the two ISI condition along with the accuracy and RT measures (A.1a-b). Statistical evaluation of the onset of individual LRPs showed a response-locked effect: a difference in LRP onset between the ISI conditions was observable only in the case of R-locked LRPs, which was also reflected on the grand average LRP waveforms. The R-LRP interval in the LONG ISI condition (mean$\pm SEM$: 529 ± 60 ms) was significantly longer compared to the SHORT ISI condition (mean$\pm SEM$: 466 ± 55 ms) ($t_{(16)} = 2.16, p < 0.05$), while the S-LRP interval was unaffected by the ISI ($t_{(16)} = 0.34, p = 0.74$; mean$\pm SEM$:$198 \pm 33$ and 191 ± 31 ms for the SHORT and LONG ISI condition, respectively; Fig.A.1b). This indicates that the source of RT increase between conditions most likely lies in response preparation processes and is unrelated to possible differences in sensory processing.

DOI:10.15774/PPKE.ITK.2010.001

References

The author's journal publications

[1] **É. M. Bankó**, V. Gál, and Z. Vidnyánszky, "Flawless visual short-term memory for facial emotional expressions," *Journal of Vision*, vol. 9, no. 1, pp. 12.1–13, 2009.

[2] **É. M. Bankó** and Z. Vidnyánszky, "Retention interval affects visual short-term memory encoding," *Journal of Neurophysiology*, vol. 103, no. 3, pp. 1425–1430, 2010.

The author's conference publications

[3] **É. Bankó** and Z. Vidnyányszky, "High-fidelity short-term memory for facial expressions," *Neuroscience Meeting Planner.*, Atlanta, GA, USA: Society for Neuroscience, 2006. Online. 548.27/X2.

The author's other journal and conference publications

[4] V. Gál, I. Kóbor, **É. M. Bankó**, L. R. Kozák, J. Serences, and Z. Vidnyánszky, "Electrophysiological correlates of learning-induced modulation of visual motion processing in humans," *Frontiers in Human Neuroscience*, vol. 3, p. 69, 2010.

DOI:10.15774/PPKE.ITK.2010.001

[5] V. Gál, L. R. Kozák, I. Kóbor, **É. M. Bankó**, J. Serences, and Z. Vidnyánszky, “Learning to filter out visual distractors,” *European Journal of Neuroscience*, vol. 29, no. 8, pp. 1723–1731, 2009.

[6] G. Kovács, M. Zimmer, **É. M. Bankó**, I. Harza, A. Antal, and Z. Vidnyánszky, “Electrophysiological correlates of visual adaptation to faces and body-parts in humans,” *Cerebral Cortex*, vol. 16, no. 5, pp. 742–753, 2006.

[7] **É. Bankó**, J. Körtvélyes, V. Gál, K. Nagy, G. Kovács, and Z. Vidnyánszky, “Dissociating the neural processes associated with perceptual processing demands and decision difficulty,” Program No. 805.13. 2009 *Neuroscience Meeting Planner*, Chicago, IL, USA: Society for Neuroscience, 2009. Online.

[8] J. Körtvélyes, **É. Bankó**, V. Gál, P. Domsa, J. Németh, and Z. Vidnyánszky, “Neural correlates of fovea-related impairment of visual object processing in amblyopia,” *The Association for Research in Vision and Opthalmology Annual Meeting (ARVO)*, Fort Lauderdale, FL, USA, 2009. Poster No. E-3820.

[9] **É. Bankó**, G. Kovács, D. Melcher, and Z. Vidnyánszky, “Hemifield-contingent face aftereffects,” *Perception ECVP Abstract Supplement*, La plainCoruña, Spain, p. 34(S):167, 2005.

Publications cited in the dissertation

[10] P. Sinha, B. Balas, Y. Ostrovsky, and R. Russell, “Face recognition by humans: Nineteen results all computer vision researchers should know about,” *Proceedings of the IEEE*, vol. 94, no. 11, pp. 1948–1962, 2006.

[11] N. Kanwisher, “Domain specificity in face perception,” *Nature Neuroscience*, vol. 3, pp. 759–763, Aug. 2000.

[12] N. Kanwisher and G. Yovel, “The fusiform face area: a cortical region specialized for the perception of faces,” *Philosophical Transactions of the*

DOI:10.15774/PPKE.ITK.2010.001

Royal Society of London. Series B, Biological Sciences, vol. 361, pp. 2109–28, Dec. 2006.

[13] M. J. Tarr and I. Gauthier, “FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise,” *Nature Neuroscience*, vol. 3, pp. 764–769, Aug. 2000.

[14] C. M. Bukach, I. Gauthier, and M. J. Tarr, “Beyond faces and modularity: the power of an expertise framework,” *Trends in Cognitive Sciences*, vol. 10, pp. 159–166, Apr. 2006.

[15] R. K. Yin, “Looking at upside-down faces,” *Journal of Experimental Psychology*, vol. 81, no. 1, pp. 141–145, 1969.

[16] G. Yovel and N. Kanwisher, “The neural basis of the behavioral face-inversion effect,” *Current Biology*, vol. 15, pp. 2256–62, Dec. 2005.

[17] J. W. Tanaka and M. J. Farah, “Parts and wholes in face recognition,” *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, vol. 46, pp. 225–245, May 1993.

[18] M. J. Farah, K. L. Levinson, and K. L. Klein, “Face perception and within-category discrimination in prosopagnosia,” *Neuropsychologia*, vol. 33, pp. 661–674, June 1995.

[19] M. Moscovitch, G. Winocur, and M. Behrmann, “What is special about face recognition? nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition,” *Journal of Cognitive Neuroscience*, vol. 9, pp. 555–604, Oct. 1997.

[20] N. Kanwisher, J. McDermott, and M. M. Chun, “The fusiform face area: a module in human extrastriate cortex specialized for face perception,” *The Journal of Neuroscience*, vol. 17, pp. 4302–11, June 1997.

[21] E. D. Renzi, D. Perani, G. A. Carlesimo, M. C. Silveri, and F. Fazio, “Prosopagnosia can be associated with damage confined to the right hemisphere–an MRI and PET study and a review of the literature,” *Neuropsychologia*, vol. 32, pp. 893–902, Aug. 1994.

DOI:10.15774/PPKE.ITK.2010.001

[22] A. Puce, T. Allison, J. C. Gore, and G. McCarthy, “Face-sensitive regions in human extrastriate cortex studied by functional MRI,” *Journal of Neurophysiology*, vol. 74, pp. 1192–9, Sept. 1995.

[23] R. J. Itier and M. J. Taylor, “Source analysis of the N170 to faces and objects,” *Neuroreport*, vol. 15, pp. 1261–5, June 2004.

[24] R. J. Itier and M. J. Taylor, “N170 or N1? spatiotemporal differences between object and face processing using ERPs,” *Cerebral Cortex*, vol. 14, pp. 132–42, Feb. 2004.

[25] S. Bentin, T. Allison, A. Puce, E. Perez, and G. McCarthy, “Electrophysiological studies of face perception in humans,” *Journal of Cognitive Neuroscience*, vol. 8, pp. 551–565, Nov. 1996.

[26] J. Liu, M. Higuchi, A. Marantz, and N. Kanwisher, “The selectivity of the occipitotemporal M170 for faces,” *Neuroreport*, vol. 11, pp. 337–341, Feb. 2000.

[27] K. Linkenkaer-Hansen, J. M. Palva, M. Sams, J. K. Hietanen, H. J. Aronen, and R. J. Ilmoniemi, “Face-selective processing in human extrastriate cortex around 120 ms after stimulus onset revealed by magneto- and electroencephalography,” *Neuroscience Letters*, vol. 253, pp. 147–50, Sept. 1998.

[28] R. Caldara, F. Jermann, G. L. Arango, and M. V. der Linden, “Is the N400 category-specific? a face and language processing study,” *Neuroreport*, vol. 15, pp. 2589–93, Dec. 2004.

[29] I. Deffke, T. Sander, J. Heidenreich, W. Sommer, G. Curio, L. Trahms, and A. Lueschow, “MEG/EEG sources of the 170-ms response to faces are co-localized in the fusiform gyrus,” *NeuroImage*, vol. 35, pp. 1495–1501, May 2007.

[30] S. R. Schweinberger, E. C. Pickering, I. Jentzsch, A. M. Burton, and J. M. Kaufmann, “Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions,” *Brain Research. Cognitive Brain Research*, vol. 14, pp. 398–409, Nov. 2002.

DOI:10.15774/PPKE.ITK.2010.001

[31] I. Gauthier, A. W. Anderson, M. J. Tarr, P. Skudlarski, and J. C. Gore, “Levels of categorization in visual recognition studied using functional magnetic resonance imaging,” *Current Biology*, vol. 7, pp. 645–651, Sept. 1997.

[32] I. Gauthier, M. J. Tarr, J. Moylan, A. W. Anderson, P. Skudlarski, and J. C. Gore, “Does visual subordinate-level categorisation engage the functionally defined fusiform face area?,” *Cognitive Neuropsychology*, vol. 17, pp. 143–164, Feb. 2000.

[33] I. Gauthier, P. Skudlarski, J. C. Gore, and A. W. Anderson, “Expertise for cars and birds recruits brain areas involved in face recognition,” *Nature Neuroscience*, vol. 3, pp. 191–197, Feb. 2000.

[34] I. Gauthier, M. J. Tarr, A. W. Anderson, P. Skudlarski, and J. C. Gore, “Activation of the middle fusiform ’face area’ increases with expertise in recognizing novel objects,” *Nature Neuroscience*, vol. 2, pp. 568–73, June 1999.

[35] I. Gauthier and M. J. Tarr, “Unraveling mechanisms for expert object recognition: bridging brain activity and behavior,” *Journal of Experimental Psychology. Human Perception and Performance*, vol. 28, pp. 431–446, Apr. 2002.

[36] R. Diamond and S. Carey, “Why faces are and are not special: an effect of expertise,” *Journal of Experimental Psychology. General*, vol. 115, pp. 107–117, June 1986.

[37] K. Grill-Spector, R. Sayres, and D. Ress, “High-resolution imaging reveals highly selective nonface clusters in the fusiform face area,” *Nature Neuroscience*, vol. 9, pp. 1177–1185, Sept. 2006.

[38] M. H. Tong, C. A. Joyce, and G. W. Cottrell, “Why is the fusiform face area recruited for novel categories of expertise? a neurocomputational investigation,” *Brain Research*, vol. 1202, pp. 14–24, Apr. 2008.

DOI:10.15774/PPKE.ITK.2010.001

[39] J. V. Haxby, E. A. Hoffman, and M. I. Gobbini, "The distributed human neural system for face perception," *Trends in Cognitive Sciences*, vol. 4, pp. 223–233, June 2000.

[40] V. Bruce and A. Young, "Understanding face recognition," *British Journal of Psychology*, vol. 77, pp. 305–327, Aug. 1986.

[41] E. C. Shuttleworth, V. Syring, and N. Allen, "Further observations on the nature of prosopagnosia," *Brain and Cognition*, vol. 1, pp. 307–322, July 1982.

[42] R. Bruyer, C. Laterre, X. Seron, P. Feyereisen, E. Strypstein, E. Pierrard, and D. Rectem, "A case of prosopagnosia with some preserved covert remembrance of familiar faces," *Brain and Cognition*, vol. 2, pp. 257–284, July 1983.

[43] N. L. Etcoff, "Selective attention to facial identity and facial emotion," *Neuropsychologia*, vol. 22, no. 3, pp. 281–295, 1984.

[44] D. Tranel, A. R. Damasio, and H. Damasio, "Intact recognition of facial expression, gender, and age in patients with impaired recognition of face identity," *Neurology*, vol. 38, pp. 690–696, May 1988.

[45] F. M. Parry, A. W. Young, J. S. Saul, and A. Moss, "Dissociable face processing impairments after brain injury," *Journal of Clinical and Experimental Neuropsychology*, vol. 13, pp. 545–558, July 1991.

[46] A. W. Young, F. Newcombe, E. H. de Haan, M. Small, and D. C. Hay, "Face perception after brain injury. selective impairments affecting identity and expression," *Brain*, vol. 116 (Pt 4), pp. 941–959, Aug. 1993.

[47] D. I. Perrett, P. A. Smith, D. D. Potter, A. J. Mistlin, A. S. Head, A. D. Milner, and M. A. Jeeves, "Neurones responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity and relation to perception," *Human Neurobiology*, vol. 3, no. 4, pp. 197–208, 1984.

DOI:10.15774/PPKE.ITK.2010.001

[48] G. C. Baylis, E. T. Rolls, and C. M. Leonard, "Selectivity between faces in the responses of a population of neurons in the cortex in the superior temporal sulcus of the monkey," *Brain Research*, vol. 342, pp. 91–102, Sept. 1985.

[49] M. E. Hasselmo, E. T. Rolls, and G. C. Baylis, "The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey," *Behavioural Brain Research*, vol. 32, pp. 203–18, Apr. 1989.

[50] D. I. Perrett, J. K. Hietanen, M. W. Oram, and P. J. Benson, "Organization and functions of cells responsive to faces in the temporal cortex," *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, vol. 335, no. 1273, pp. 23–30, 1992.

[51] P. Vuilleumier, J. L. Armony, J. Driver, and R. J. Dolan, "Effects of attention and emotion on face processing in the human brain - an Event-Related fMRI study," *Neuron*, vol. 30, pp. 829–841, June 2001.

[52] J. S. Winston, R. N. Henson, M. R. Fine-Goulden, and R. J. Dolan, "fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception," *Journal of Neurophysiology*, vol. 92, pp. 1830–1839, Apr. 2004.

[53] A. J. Calder and A. W. Young, "Understanding the recognition of facial identity and facial expression," *Nature Reviews. Neuroscience*, vol. 6, pp. 641–51, Aug. 2005.

[54] T. Ganel, K. F. Valyear, Y. Goshen-Gottstein, and M. A. Goodale, "The involvement of the "fusiform face area" in processing facial expression," *Neuropsychologia*, vol. 43, no. 11, pp. 1645–1654, 2005.

[55] A. Ishai, C. F. Schmidt, and P. Boesiger, "Face perception is mediated by a distributed cortical network," *Brain Research Bulletin*, vol. 67, pp. 87–93, Sept. 2005.

DOI:10.15774/PPKE.ITK.2010.001

[56] C. J. Fox, S. Y. Moon, G. Iaria, and J. J. S. Barton, “The correlates of subjective perception of identity and expression in the face network: an fMRI adaptation study,” *NeuroImage*, vol. 44, no. 2, pp. 569–580, 2009.

[57] K. M. Curby and I. Gauthier, “A visual short-term memory advantage for faces,” *Psychonomic Bulletin & Review*, vol. 14, pp. 620–628, Aug. 2007.

[58] K. M. Curby, K. Glazek, and I. Gauthier, “A visual short-term memory advantage for objects of expertise,” *Journal of Experimental Psychology. Human Perception and Performance*, vol. 35, pp. 94–107, Feb. 2009.

[59] A. Freire, K. Lee, and L. A. Symons, “The face-inversion effect as a deficit in the encoding of configural information: direct evidence,” *Perception*, vol. 29, no. 2, pp. 159–170, 2000.

[60] P. Ekman, *Darwin and Facial Expression: A Century of Research in Review.* New York: Academic Press, 1973.

[61] C. E. Izard, *Human Emotions.* New York: Plenum, 1 ed., June 1977.

[62] A. J. Fridlund, *Human Facial Expression.* New York: Academic Press, 1994.

[63] A. Ohman, A. Flykt, and F. Esteves, “Emotion drives attention: detecting the snake in the grass,” *Journal of Experimental Psychology. General*, vol. 130, pp. 466–78, Sept. 2001.

[64] J. D. Eastwood, D. Smilek, and P. M. Merikle, “Negative facial expression captures attention and disrupts performance,” *Perception & Psychophysics*, vol. 65, pp. 352–8, Apr. 2003.

[65] N. Lucas and P. Vuilleumier, “Effects of emotional and non-emotional cues on visual search in neglect patients: evidence for distinct sources of attentional guidance,” *Neuropsychologia*, vol. 46, pp. 1401–14, Apr. 2008.

[66] T. Lee, R. J. Dolan, and H. D. Critchley, “Controlling emotional expression: behavioral and neural correlates of nonimitative emotional responses,” *Cerebral Cortex*, vol. 18, no. 1, pp. 104–13, 2008.

DOI:10.15774/PPKE.ITK.2010.001

[67] D. J. Simons and D. T. Levin, "Failure to detect changes to people during a real-world interaction.," *Psychonomic Bulletin & Review*, vol. 5, pp. 644–649, Dec. 1998.

[68] J. V. Haxby, M. I. Gobbini, M. L. Furey, A. Ishai, J. L. Schouten, and P. Pietrini, "Distributed and overlapping representations of faces and objects in ventral temporal cortex," *Science*, vol. 293, pp. 2425–2430, Sept. 2001.

[69] J. Narumoto, T. Okada, N. Sadato, K. Fukui, and Y. Yonekura, "Attention to emotion modulates fMRI activity in human right superior temporal sulcus," *Brain Research. Cognitive Brain Research*, vol. 12, pp. 225–31, Oct. 2001.

[70] T. J. Andrews and M. P. Ewbank, "Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe," *NeuroImage*, vol. 23, pp. 905–13, Nov. 2004.

[71] M. L. LoPresti, K. Schon, M. D. Tricarico, J. D. Swisher, K. A. Celone, and C. E. Stern, "Working memory for social cues recruits orbitofrontal cortex and amygdala: a functional magnetic resonance imaging study of delayed matching to sample for emotional expressions," *The Journal of Neuroscience*, vol. 28, pp. 3718–28, Apr. 2008.

[72] S. Magnussen, E. Idas, and S. H. Myhre, "Representation of orientation and spatial frequency in perception and memory: a choice reaction-time analysis," *Journal of Experimental Psychology. Human Perception and Performance*, vol. 24, pp. 707–18, June 1998.

[73] I. Reinvang, S. Magnussen, M. W. Greenlee, and P. G. Larsson, "Electrophysiological localization of brain regions involved in perceptual memory," *Experimental Brain Research*, vol. 123, pp. 481–4, Dec. 1998.

[74] S. Magnussen, "Low-level memory processes in vision," *Trends in Neurosciences*, vol. 23, pp. 247–51, June 2000.

DOI:10.15774/PPKE.ITK.2010.001

[75] G. Kovács, M. Zimmer, I. Harza, A. Antal, and Z. Vidnyánszky, "Position-specificity of facial adaptation," *Neuroreport*, vol. 16, pp. 1945–9, Nov. 2005.

[76] G. Kovács, M. Zimmer, E. Bankó, I. Harza, A. Antal, and Z. Vidnyánszky, "Electrophysiological correlates of visual adaptation to faces and body parts in humans," *Cerebral Cortex*, vol. 16, pp. 742–53, May 2006.

[77] G. Kovács, M. Zimmer, I. Harza, and Z. Vidnyánszky, "Adaptation duration affects the spatial selectivity of facial aftereffects," *Vision Research*, vol. 47, pp. 3141–9, Nov. 2007.

[78] F. A. Wichmann and N. J. Hill, "The psychometric function: I. fitting, sampling, and goodness of fit," *Perception & psychophysics*, vol. 63, pp. 1293–313, Nov. 2001.

[79] B. Lee and J. Harris, "Contrast transfer characteristics of visual short-term memory," *Vision Research*, vol. 36, pp. 2159–2166, July 1996.

[80] B. Rossion and I. Gauthier, "How does the brain process upright and inverted faces?," *Behavioral and Cognitive Neuroscience Reviews*, vol. 1, pp. 63–75, Mar. 2002.

[81] T. Pasternak and M. W. Greenlee, "Working memory in primate sensory systems," *Nature Reviews. Neuroscience*, vol. 6, pp. 97–107, Feb. 2005.

[82] M. Lages and M. Treisman, "Spatial frequency discrimination: visual long-term memory or criterion setting?," *Vision research*, vol. 38, pp. 557–72, Feb. 1998.

[83] S. Magnussen, M. W. Green, P. M. Aslaksen, and O. O. Kildebo, "High-fidelity perceptual long-term memory revisited–and confirmed," *Psychological science*, vol. 14, no. 1, pp. 74–6, 2003.

[84] M. E. Hasselmo, E. T. Rolls, and G. C. Baylis, "The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey," *Behavioural Brain Research*, vol. 32, pp. 203–18, Apr. 1989.

[85] D. G. Pelli, "The VideoToolbox software for visual psychophysics: transforming numbers into movies," *Spatial Vision*, vol. 10, pp. 437–442, 1997.

[86] D. H. Brainard, "The psychophysics toolbox," *Spatial vision*, vol. 10, no. 4, pp. 433–436, 1997.

[87] J. L. Lancaster, M. G. Woldorff, L. M. Parsons, M. Liotti, C. S. Freitas, L. Rainey, P. V. Kochunov, D. Nickerson, S. A. Mikiten, and P. T. Fox, "Automated talairach atlas labels for functional brain mapping," *Human Brain Mapping*, vol. 10, pp. 120–31, July 2000.

[88] C. Rorden and M. Brett, "Stereotaxic display of brain lesions," *Behavioural Neurology*, vol. 12, no. 4, pp. 191–200, 2000.

[89] S. G. Costafreda, M. J. Brammer, A. S. David, and C. H. Y. Fu, "Predictors of amygdala activation during the processing of emotional stimuli: a meta-analysis of 385 PET and fMRI studies," *Brain Research Reviews*, vol. 58, pp. 57–70, June 2008.

[90] C. Pierrot-Deseilligny, D. Milea, and R. M. Müri, "Eye movement control by the cerebral cortex," *Current Opinion in Neurology*, vol. 17, pp. 17–25, Feb. 2004.

[91] T. J. Druzgal and M. D'Esposito, "Dissecting contributions of prefrontal cortex and fusiform face area to face working memory," *Journal of Cognitive Neuroscience*, vol. 15, pp. 771–84, Aug. 2003.

[92] B. R. Postle, T. J. Druzgal, and M. D'Esposito, "Seeking the neural substrates of visual working memory storage," *Cortex*, vol. 39, no. 4-5, pp. 927–46, 2003.

[93] J. H. Yoon, C. E. Curtis, and M. D'Esposito, "Differential effects of distraction during working memory on delay-period activity in the prefrontal cortex and the visual association cortex," *NeuroImage*, vol. 29, pp. 1117–26, Feb. 2006.

[94] E. A. Nichols, Y. Kao, M. Verfaellie, and J. D. Gabrieli, "Working memory and long-term memory for faces: Evidence from fMRI and global amnesia

DOI:10.15774/PPKE.ITK.2010.001

for involvement of the medial temporal lobes," *Hippocampus*, vol. 16, no. 7, pp. 604–616, 2006.

[95] I. R. Olson, K. Page, K. S. Moore, A. Chatterjee, and M. Verfaellie, "Working memory for conjunctions relies on the medial temporal lobe," *The Journal of Neuroscience*, vol. 26, pp. 4596–4601, Apr. 2006.

[96] C. Bledowski, K. C. Kadosh, M. Wibral, B. Rahm, R. A. Bittner, K. Hoechstetter, M. Scherg, K. Maurer, R. Goebel, and D. E. J. Linden, "Mental chronometry of working memory retrieval: a combined functional magnetic resonance imaging and event-related potentials approach," *The Journal of Neuroscience*, vol. 26, no. 3, pp. 821–9, 2006.

[97] M. Fabiani, D. Karis, and E. Donchin, "P300 and recall in an incidental memory paradigm," *Psychophysiology*, vol. 23, pp. 298–308, May 1986.

[98] L. K. McEvoy, M. E. Smith, and A. Gevins, "Dynamic cortical networks of verbal and spatial working memory: effects of memory load and task practice," *Cerebral Cortex*, vol. 8, pp. 563–574, Nov. 1998.

[99] C. Haenschel, R. A. Bittner, F. Haertling, A. Rotarska-Jagiela, K. Maurer, W. Singer, and D. E. J. Linden, "Contribution of impaired early-stage visual processing to working memory dysfunction in adolescents with schizophrenia: a study with event-related potentials and functional magnetic resonance imaging," *Archives of General Psychiatry*, vol. 64, pp. 1229–1240, Nov. 2007.

[100] H. M. Morgan, C. Klein, S. G. Boehm, K. L. Shapiro, and D. E. J. Linden, "Working memory load for faces modulates P300, N170, and N250r," *Journal of Cognitive Neuroscience*, vol. 20, pp. 989–1002, June 2008.

[101] S. E. Prince, N. A. Dennis, and R. Cabeza, "Encoding and retrieving faces and places: distinguishing process- and stimulus-specific differences in brain activity," *Neuropsychologia*, vol. 47, pp. 2282–2289, Sept. 2009.

[102] D. Lehmann and W. Skrandies, "Reference-free identification of components of checkerboard-evoked multichannel potential fields," *Electroen-*

DOI:10.15774/PPKE.ITK.2010.001

cephalography and Clinical Neurophysiology, vol. 48, pp. 609–21, June 1980.

[103] H. L. Hamburger and M. A. G. v.d. Burgt, "Global field power measurement versus classical method in the determination of the latency of evoked potential components," *Brain Topography*, vol. 3, pp. 391–396, Mar. 1991.

[104] M. Scherg and P. Berg, "Use of prior knowledge in brain electromagnetic source analysis," *Brain Topography*, vol. 4, no. 2, pp. 143–50, 1991.

[105] B. Efron and R. J. Tibshirani, *An introduction to the bootstrap.* New York: Chapman and Hall, 1993.

[106] K. Hoechstetter, A. Rupp, A. Stancák, H. M. Meinck, C. Stippich, P. Berg, and M. Scherg, "Interaction of tactile input in the human primary and secondary somatosensory cortex–a magnetoencephalographic study," *NeuroImage*, vol. 14, pp. 759–67, Sept. 2001.

[107] A. Strobel, S. Debener, B. Sorger, J. C. Peters, C. Kranczioch, K. Hoechstetter, A. K. Engel, B. Brocke, and R. Goebel, "Novelty and target processing during an auditory novelty oddball: a simultaneous event-related potential and functional magnetic resonance imaging study," *NeuroImage*, vol. 40, pp. 869–83, Apr. 2008.

[108] W. Skrandies, "EEG/EP: new techniques," *Brain Topography*, vol. 5, pp. 347–350, June 1993.

[109] D. A. Jeffreys, "A face-responsive potential recorded from the human scalp," *Experimental Brain Research*, vol. 78, no. 1, pp. 193–202, 1989.

[110] F. D. Russo, A. Martínez, M. I. Sereno, S. Pitzalis, and S. A. Hillyard, "Cortical sources of the early components of the visual evoked potential," *Human Brain Mapping*, vol. 15, pp. 95–111, Feb. 2002.

[111] S. M. Letourneau and T. V. Mitchell, "Behavioral and ERP measures of holistic face processing in a composite task," *Brain and Cognition*, vol. 67, pp. 234–45, July 2008.

DOI:10.15774/PPKE.ITK.2010.001

[112] A. C. H. Lee, V. L. Scahill, and K. S. Graham, "Activating the medial temporal lobe during oddity judgment for faces and scenes," *Cerebral Cortex*, vol. 18, pp. 683–696, Mar. 2008.

[113] P. Trautner, T. Dietl, M. Staedtgen, A. Mecklinger, T. Grunwald, C. E. Elger, and M. Kurthen, "Recognition of famous faces in the medial temporal lobe: an invasive ERP study," *Neurology*, vol. 63, pp. 1203–1208, Oct. 2004.

[114] T. Dietl, P. Trautner, M. Staedtgen, M. Vannuchi, A. Mecklinger, T. Grunwald, H. Clusmann, C. E. Elger, and M. Kurthen, "Processing of famous faces and medial temporal lobe event-related potentials: a depth electrode study," *NeuroImage*, vol. 25, pp. 401–407, Apr. 2005.

[115] E. J. Barbeau, M. J. Taylor, J. Regis, P. Marquis, P. Chauvel, and C. Liegeois-Chauvel, "Spatio temporal dynamics of face recognition," *Cerebral Cortex*, vol. 18, pp. 997–1009, May 2008.

[116] P. Wright, D. Albarracin, R. D. Brown, H. Li, G. He, and Y. Liu, "Dissociated responses in the amygdala and orbitofrontal cortex to bottom-up and top-down components of emotional evaluation," *NeuroImage*, vol. 39, no. 2, pp. 894–902, 2008.

[117] A. Harris and K. Nakayama, "Rapid face-selective adaptation of an early extrastriate component in MEG," *Cerebral Cortex*, vol. 17, no. 1, pp. 63–70, 2007.

[118] W. Klimesch, R. Freunberger, P. Sauseng, and W. Gruber, "A short review of slow phase synchronization and memory: evidence for control processes in different memory systems?," *Brain Research*, vol. 1235, pp. 31–44, Oct. 2008.

[119] R. Freunberger, R. Fellinger, P. Sauseng, W. Gruber, and W. Klimesch, "Dissociation between phase-locked and nonphase-locked alpha oscillations in a working memory task," *Human Brain Mapping*, Apr. 2009.

[120] J. Polich, "Updating p300: an integrative theory of p3a and p3b," *Clinical Neurophysiology*, vol. 118, pp. 2128–2148, Oct. 2007.

DOI:10.15774/PPKE.ITK.2010.001

[121] N. Sasson, N. Tsuchiya, R. Hurley, S. M. Couture, D. L. Penn, R. Adolphs, and J. Piven, "Orienting to social stimuli differentiates social cognitive impairment in autism and schizophrenia," *Neuropsychologia*, vol. 45, pp. 2580–8, June 2007.

[122] K. Humphreys, N. Minshew, G. L. Leonard, and M. Behrmann, "A fine-grained analysis of facial expression processing in high-functioning adults with autism," *Neuropsychologia*, vol. 45, pp. 685–95, Mar. 2007.

[123] M. H. Kosmidis, V. P. Bozikas, M. Giannakou, D. Anezoulaki, B. D. Fantie, and A. Karavatos, "Impaired emotion perception in schizophrenia: a differential deficit," *Psychiatry Research*, vol. 149, no. 1-3, pp. 279–84, 2007.

[124] J. Jonides, R. L. Lewis, D. E. Nee, C. A. Lustig, M. G. Berman, and K. S. Moore, "The mind and brain of short-term memory," *Annual Review of Psychology*, vol. 59, pp. 193–224, 2008.

[125] M. G. Coles, "Modern mind-brain reading: psychophysiology, physiology, and cognition," *Psychophysiology*, vol. 26, pp. 251–269, May 1989.

[126] M. Eimer, "The lateralized readiness potential as an on-line measure of central response activation processes," *Behavior Research Methods, Intruments & Computers*, vol. 30, no. 1, pp. 146–156, 1998.

[127] J. R. Jennings and M. G. H. Coles, *Handbook of Cognitive Psychophysiology: Central and Autonomic Nervous System Approaches.* Chichester, UK: Wiley, 1991.

[128] A. Osman, C. M. Moore, and R. Ulrich, "Bisecting RT with lateralized readiness potentials: Precue effects after LRP onset," *Acta Psychologica*, vol. 90, pp. 111–127, Nov. 1995.

[129] H. Müller-Gethmann, R. Ulrich, and G. Rinkenauer, "Locus of the effect of temporal preparation: evidence from the lateralized readiness potential," *Psychophysiology*, vol. 40, pp. 597–611, July 2003.

DOI:10.15774/PPKE.ITK.2010.001

[130] S. Hsieh, "The lateralized readiness potential and P300 of stimulus-set switching," *International Journal of Psychophysiology*, vol. 60, pp. 284–91, June 2006.

[131] P. Schwarzenau, M. Falkenstein, J. Hoormann, and J. Hohnsbein, "A new method for the estimation of the onset of the lateralized readiness potential (LRP)," *Behavior Research Methods, Intruments & Computers*, vol. 30, no. 1, pp. 110–117, 1998.

[132] J. T. Mordkoff and P. J. Gianaros, "Detecting the onset of the lateralized readiness potential: a comparison of available methods and procedures," *Psychophysiology*, vol. 37, pp. 347–60, May 2000.