# The role of contexts in face processing

Righart, R. G. R.

*Publication date:*
2007

*Document Version*
Publisher's PDF, also known as Version of record

# The Role of Contexts in Face Processing

## Behavioral and ERP studies

# THE ROLE
# OF
# CONTEXTS IN FACE PROCESSING

## BEHAVIORAL AND ERP STUDIES

PROEFSCHRIFT

Ter verkrijging van de graad van doctor
aan de Universiteit van Tilburg,
op gezag van de rector magnificus,
prof. dr. F.A. van der Duyn Schouten,
in het openbaar te verdedigen ten overstaan van
een door het college voor promoties aangewezen commissie in de aula van de
Universiteit op vrijdag 15 juni 2007 om 16.15 uur

door

**Ruthger Gerdinus Reinout Righart**
geboren op 27 juli 1977 te Brouwershaven

Promotor: Prof. dr. B.L.M.F. de Gelder

Promotiecommissie: Prof. dr. P. Bertelson
Dr. G.J.M. van Boxtel
Prof. dr. N.H. Frijda
Prof. dr. E.H.F. de Haan
Prof. dr. K.R. Ridderinkhof




ISBN/EAN: 978-90-5335-123-9

# Table of Contents

# Chapter 1

# Introduction

Humans readily recognize and understand others by seeing their faces and facial expressions. Numerous studies in cognitive neuroscience have investigated how humans recognize facial identity and facial expressions (reviewed by Adolphs, 2002; Bruce & Young, 1986; Haxby, Hoffman, & Gobbini, 2000), and what factors affect the recognition of identity and expression. Factors like the color properties of the face (Yip & Sinha, 2002), the viewing angle of the perceived face (Hill, Schyns, & Akamatsu, 1997; Kowatari et al., 2004), and the intensity of the perceived facial expression (Palermo & Coltheart, 2004) crucially define how accurate individuals recognize others.

Although behavioral and neuroimaging studies have progressed our understanding about face recognition, an important factor that may influence how we recognize others has often been overlooked in experimental research. This factor is the context in which the face is perceived. In daily life, we rarely see faces without any context, like the photo in our passport.

We make sense of others and their reactions by understanding the environmental context that evoked those reactions (Frijda & Tcherkassof, 1997). For instance, facial expressions of fear are evoked rapidly when someone sees that a house is on fire. These expressions are understood within a split second, partly because of the context that evoked it. People act by calling 911 or try to help the victims that are helplessly caught in the fire. Thus, the context indicates what emotional expressions are appropriate and whether action is required.

Another example in which contexts define face processing is when an emotional event in which the person is perceived is highly arousing. This highly arousing event may affect recognition memory for the identity of the person. For example, in the case of a bank robbery, victims are often unable to report the identity of the criminal (see for instance eyewitness reports, Christianson, 1992).

## 1.1 Why should face recognition be studied in context?

There are at least three reasons why face recognition should be studied as a function of context information. First, faces are continuously seen in physically different environments (e.g., luminance, shade), and therefore context processing has a crucial role in face perception. For example, when viewing conditions are degraded, perhaps as a consequence of imperfect illumination, or when the distance to the person's face is increased, missing information may be filled-in (see Cox, Meyers, & Sinha, 2004). These are elementary processes which have been studied by using simple objects that are accompanied by different context conditions (Albright & Stoner, 2002; Lamme & Roelfsema, 2000), but these processes have as yet been ignored in most face recognition studies.

Second, contexts do not only provide meaningful information about the physical environment, but in addition they have a meaningful association with faces. Most of us are certainly highly accurate in recognizing others. However, even if you are highly accurate in face recognition, recognition may sometimes fail. Almost everyone has experienced the embarrassment when being unable to recognize someone. This failure may partly relate to the context. For instance, a butcher is often recognized easily when he is in his butchery, but when he is seen in the bus, the unusual context may delay recognition. Individuals often try to resolve such a failure by thinking about the context where the person was initially encountered (Mandler, 1980; Young, Hay, & Ellis, 1985). Thus, the semantic relatedness between a face and the context may define how we recognize others. A similar role has been proposed for the contexts in which non-face objects are perceived (Bar, 2004; Davenport & Potter, 2004).

Third, and most pertinent to the research in this dissertation, the emotional relevance of a context in which a face is seen may crucially define how facial expressions are recognized, and how a face is memorized. Facial expressions usually do not appear in a vacuum, but are part of the person's interaction with the environment (Frijda & Tcherkassof, 1997). The environmental context may inform the observer about the emotional or social signals others usually express. To date, recognition of facial expressions has in most experiments been studied separately from the context in which the facial expression originated. However, focussing on facial expressions without the context is of questionable generality to natural situations (Russell, 1994).

## 1.2 A definition of context

Numerous studies in cognitive neuroscience have investigated context effects. As a consequence, different definitions are in use. Context is defined as "the interrelated conditions in which something exists or occurs" by Merriam-Webster dictionary. Many studies have used the definition "context" to indicate the effect circumstances have on the processing of a target stimulus. In fact, in operational definition, different types of contexts have been used across experiments (Table 1.1). Wallbott (1988) has distinguished between a "person scenario approach" and a "context picture approach", in which in the first definition context information is presented verbally, while in the second definition information is presented visually. In present thesis, context is operationalized as the situation in which a face is perceived. We have used natural scenes that are non-emotional or emotional (e.g., fearful, disgusting, and happy) in which a face is positioned centrally. The face and scene stimulus always had similar onset- and offset times (for an overview of studies that have used similar designs, see Table 1.1; Example face-context stimulus, Figure 1.1).

The designation "emotional scene" is used throughout the text. The emotion that is indicated is the emotion that is evoked in the observer. For example, an emotional scene may express anger, by showing for example images of war, but this may still evoke fear in the observer. Thus, in this example the designation "fearful scene" would be used, given the emotion it evoked in participants. The scenes that were used have been validated for emotional arousal and valence as will be described further in each chapter.

**Table 1.1 Overview of Context Studies**

List of studies reporting context effects on face, object or word recognition.

*Design*
* Studies using natural photographic scenes
o Studies using similar onset times for target and context
x Studies using emotional arousing target and/or context

| *Authors* | *Design* | *Target* | *Context* | *Measurement* |
|---|---|---|---|---|
| Galli et al. (2006) | x | Faces | Emotional newspaper messages | ERP: face recognition memory |
| Hannula et al. (2006) | * | Faces | Scenes | ERP: face recognition memory |
| Mobbs et al. (2006) | * x | Facial expressions | Emotional scenes | fMRI: recognition facial expression |
| Smith et al. (2006) | * x | Objects | Emotional scenes | fMRI: object recognition memory |
| Sterpenich et al. (2006) | * o x | Faces | Emotional scenes | fMRI: face recognition memory |
| Erk et al. (2005) | * x | Words | Emotional scenes | fMRI: word memory |
| Meeren et al. (2005) | o x | Facial expressions | Bodily expressions | ERP: recognition facial expressions |
| Cox et al. (2004) | o | Faces | Bodies | fMRI: repetition memory (one-back) |
| Davenport & Potter (2004) | * o | Persons and objects | Scenes | Beh: object recognition |
| Goh et al. (2004) | * o | Objects | Scenes | fMRI: passive viewing task |
| Kim et al. (2004) | x | Facial expressions | Verbal emotional message | fMRI: recognition facial expression |
| Rousselet et al. (2004) | * o | Faces | Scenes | ERP: detection human or animal face |
| Smith et al. (2004) | * x | Objects | Emotional scenes | fMRI: object recognition memory |
| Smith et al. (2004) | * x | Objects | Emotional scenes | ERP: object recognition memory |
| Yovel & Paller (2004) | | Faces | Verbal occupation | ERP: face recognition memory |
| Yi et al. (2004) | * o | Faces | Scenes | fMRI: face repetition |
| Tsivilis et al. (2003) | * o | Objects | Scenes | fMRI: object recognition memory |
| Erk et al. (2003) | * x | Words | Emotional scenes | fMRI: word memory |
| Bar (2003) | * | Objects | Objects and scenes | fMRI: image recognition |
| Ganis & Kutas (2003) | * | Objects | Scenes | ERP: object recognition |
| Mathis et al. (2002) | o | Words | Objects and scenes | Beh: word discrimination |
| Rainis (2001) | * o x | Facial expressions | Emotional scenes | Beh: face recognition memory |
| Tsivilis (2001) | * o | Objects | Scenes | ERP: object recognition memory |
| Guillaume & Tiberghien (2001) | * o x | Faces | Scenes | ERP: detect face repetition |
| Henderson et al. (1999) | o | Objects | Line-drawings | Beh / eye movement: object memory |
| Hollingworth & Henderson (1998) | o | Objects | Line-drawings | Beh: object detection |
| Dolan et al. (1997) | * o | Faces and objects | Impoverished two-tone pictures | Beh: perceptual learning |
| Carroll & Russell (1996) | x | Facial expressions | Verbal emotional storyline | Beh: recognition facial expression |
| de Houwer & Hermans (1994) | * o x | Emotional words | Line-drawings | Beh: affect decision |
| Boyce et al. (1992) | o | Objects | Line-drawings | Beh: naming target objects |
| Feenan et al. (1990) | o | Pictures and words | Line-drawings / Names | Beh: recognition memory |
| Murphy et al. (1989) | o | Objects | Line-drawings | Beh: object categorization |
| Boyce et al. (1989) | o | Objects | Line-drawings | Beh: object detection |
| Wallbott (1988) Exp1 | x | Facial expressions | Verbal emotional storyline | Beh: recognition facial expression |
| Wallbott (1988) Exp2 | * o x | Facial expressions | Emotional scene | Beh: recognition facial expression |
| Beales & Parkin (1984) | * o | Faces | Scenes | Beh: face recognition memory |
| Memon & Bruce (1983) | * o | Faces | Scenes | Beh: face recognition memory |
| Davies & Milne (1982) | * o | Faces | Scenes | Beh: face recognition memory |
| Biederman et al. (1982) | o | Objects | Scenes | Beh: target detection |
| Friedman (1979) | o | Objects | Line-drawings | Beh / eye movement: memory |
| Palmer (1975) | o | Object | Line-drawings | Beh: object recognition |
| Frijda (1961) | x | Facial expressions | Verbal emotional storyline | Beh: describe person's feelings |

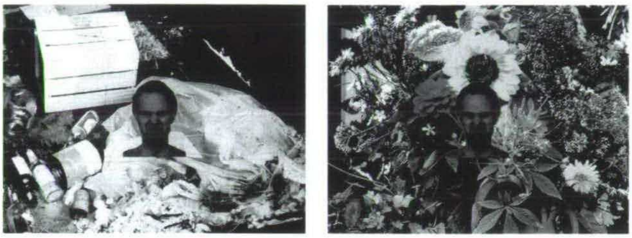


**Figure 1.1 Stimulus example.** Example of a face-context compound used in the behavioral experiments reported in chapter 3. Facial expressions of fear, happiness and disgust were paired with emotional scenes of fear, happiness and disgust. These pairs could constitute congruent emotions, for instance (a) a facial expression of disgust in a context of garbage or incongruent emotions, (b) the same expression shown among flowers.

Before introducing the research questions, behavioral and event-related potential studies of face recognition will be reviewed briefly for faces that were studied in isolation from context scenes.

## 1.3 Behavioral studies of face recognition

Behavioral experiments (accuracies and response times[1]) have contributed importantly to the study of face recognition for several reasons. First, behavioral methods provide a direct quantitative measure of performance. The error patterns and response times across experimental conditions are crucial in determining how well individuals can recognize faces and facial expressions. By means of response times, it can be inferred for example, that it took 30 ms longer to respond to condition A as compared to condition B. A disadvantage, however, is that it is not clear how response times relate to different stages of perception (Luck, 2006). For example, it is not clear how stages of visual encoding, response initiation and response execution contribute to the eventual response time (but see extensive experimental work by Donders (1868) and Sternberg (1969) about the use of RTs in the study of stages of information processing). Second, behavioral methods may provide complementary information for the results obtained by neuroimaging methods (e.g., event-related potentials (ERP), magnetoencephalography (MEG), functional Magnetic Resonance Imaging (fMRI)). For example, in studies that modulated task parameters, like task load, it has been shown that measures of performance related to the activation in certain brain regions (see Lavie, 2005; Yi, Woodman, Widders, Marois, & Chun, 2004). Third, the concepts that are studied by neuroimaging methods are greatly determined by the outcome of behavioral studies. In this thesis, behavioral experiments have been used to investigate the recognition of

[1] The more generic term "response time" is used throughout the text (see Luce, 1986 for a distinction between "response time" and "reaction time").

facial expressions (chapter 3) and recognition memory for faces (chapter 6), and are an important starting point for interpreting the ERP results (see chapter 4 and 6).

### *1.3.1 Recognition of facial expressions*

How individuals recognize facial expressions of emotions has been studied by emotion recognition tasks. In these tasks, participants are instructed to judge the emotion by using free-choice formats or forced-choice formats. In free-choice formats, participants are instructed to describe each presented facial expression, whatever comes to mind. Alternatively, in forced-choice formats, participants are instructed to select for each presented facial expression an emotion from predefined emotion labels. Although a disadvantage of forced-choice formats is that it forces participants to select an available emotion label (Frijda & Tcherkassof, 1997; Russell, 1994), a great advantage is that it relies less on individual vocabulary skills. Moreover, in addition to accuracy measures, response times can be measured (see chapter 3), and responses that are acquired by a forced-choice format can be coupled easily with event-related data in ERP and fMRI studies. Trials can be sorted according to the behavioral responses that were given by the participant. For instance, in studying the recognition of facial expressions, ERPs can be analyzed for each facial expression category separately (chapter 4).

Many studies that have employed forced-choice designs have shown better and faster categorization of facial expressions of happiness as compared to negative facial expressions (e.g., anger, disgust, fear, and sadness) (Calder, Young, Keane, & Dean, 2000; Grimshaw, Bulman-Fleming, & Ngo, 2004; Kirita & Endo, 1995; Kirouac & Dore, 1983; Leppänen, Tenhunen, & Hietanen, 2003; McKelvie, 1995; Palermo & Coltheart, 2004). It is not clear as yet why a response benefit is obtained for facial expressions of happiness. First, performance may depend on the number of alternatives that is available, and probably more important, what alternatives are available. Although several studies have found that performance depends on which choice alternatives are available (Russell, 1994; Tanaka-Matsumi, Attivissimo, Nelson, & D'Urso, 1995), a happy face advantage appears to be found invariably.

Second, the emotional intensity of the presented facial expressions may contribute differently to how facial expressions are recognized (Russell, 1997). A number of studies have found that the emotional intensity increases performance for negative facial expressions (i.e., angry, disgust, fearful, sad, surprise), whereas this increase was not observed for happy facial expressions (Hess, Blairy, & Kleck, 1997; Palermo & Coltheart, 2004). It should however be noted that accuracies for facial expressions of happiness are often at ceiling level, which may explain why emotional intensity did not affect recognition.

Third, low-level features of facial expressions have been considered as a factor that may contribute to the response advantage that has been observed for facial expressions of happiness (e.g., smiling mouth, teeth). This has been tested by using schematic stimuli, which do not contain the low-level features that are displayed by photographic stimuli. Similar response advantages were obtained for happy facial expressions when these schematic stimuli were used, which suggests that the response advantage for happiness is related to emotion rather than low-level features (Kirita & Endo, 1995). As will be seen, in addition to the response advantage that was found for facial expressions of happiness, context emotions contribute to how facial expressions are recognized (chapter 3), which suggests that it is a unique effect by emotion, not low-level features.

Last, as there are more categories for negative facial expressions, it may be more difficult to categorize a negative expression because they share a greater number of configural features (Leppänen et al., 2003).

The observed response advantage for facial expressions of happiness is counterintuitive with the idea that negative emotions are processed faster than positive emotions (LeDoux, 1996). Importantly, the idea expressed by LeDoux (1996) is that the brain responds rapidly to significant situations, and that the initial response that is elicited by significant stimuli does not require conscious awareness of the stimulus or conscious control of the response. Therefore, many factors in between this initial response of the brain and the ultimate manual response may have affected the reported response times. One of these factors may be the task which participants perform.

Other behavioral experiments have shown that results differ when participants have to detect facial expressions instead of categorizing them. The primary result from these experiments is that faster responses have been found when participants are required to detect negative facial expressions. For instance, in a visual search paradigm by Hansen and Hansen (1988), participants were required to discover the presence or absence of a discrepant face in a 3x3 matrix of nine faces. It was found that angry faces were detected faster among distracters of neutral faces and happy faces as compared to the detection of happy and neutral faces among distracters of angry faces. Many studies have confirmed these results (Eastwood, Smilek, & Merikle, 2003; Fox et al., 2000; Öhman, Lundqvist, & Esteves, 2001). The faster responses are not confined to negative facial expressions (Öhman, Lundqvist et al., 2001), but were also found for negative scenes (Blanchette, 2006; Öhman, Flykt, & Esteves, 2001). The differences that are found between detection tasks and categorization tasks may rely on attentional processes, as negative emotions may not only capture attention, but also hold attention in tasks that require a more elaborate discrimination between emotions (Fox, Russo, Bowles, & Dutton, 2001).

Recognition tasks for facial expressions have also been used in neurological and psychiatric populations. Schizophrenia patients show deficits in recognizing facial expressions (Feinberg, Rifkin, Schaffer, & Walker, 1986; Kohler et al., 2003). Research has shown correlations between their deficits to recognize facial expressions and their ability to socially interact (Mueser et al., 1996). In chapter 5, recognition of facial expressions in schizophrenia will be studied for facial expressions that are accompanied by emotional salient contexts.

### *1.3.2. Recognition memory for faces*

Memory for faces has been studied by using repetition paradigms or study-test paradigms. In a repetition paradigm, study- and test items are intermixed in a single experimental block (Itier & Taylor, 2004a; Schweinberger, 1995; Schweinberger, Pickering, Jentzsch, Burton, & Kaufmann, 2002). In a study-test paradigm, faces are studied and tested in separate blocks (Jenkins, Lavie, & Driver, 2005; Johansson, Mecklinger, & Treese, 2004; Reinitz, Morrissey, & Demb, 1994; Yovel & Paller, 2004), which may be advantageous for providing different instructions to each phase (Bower & Karlin, 1974; Jenkins et al., 2005; Memon & Bruce, 1983; Patterson & Baddeley, 1977).

For analyses, in face recognition memory studies, hits (i.e., the proportion of faces that are correctly recognized as old) and correct rejections (i.e., the proportion of faces that are correctly recognized as new) are calculated as a function of particular conditions, for instance the emotional context in which faces had been studied (chapter 6). Recognition memory is not only determined by the proportion of hits and correct rejections, but may also depend on the response bias (Snodgrass & Corwin, 1988), because a tendency to make an old response may introduce a parallel increase in the false alarm rate. In order to analyse response biases, misses (the proportion of faces that are incorrectly classified as old) and false alarm responses (the proportion of faces that are incorrectly classified as new) are analyzed.

Important response-bias effects have been observed for emotional stimuli. False alarm rates were found to be higher for emotional negative words as compared to neutral words (Maratos, Allan, & Rugg, 2000; McNeely, Dywan, & Segalowitz, 2004; Windmann & Kutas, 2001). For faces, false alarm rates were increased for facial expressions of emotions, particularly negative facial expressions, as compared to neutral facial expressions (Johansson et al., 2004). For this reason, it is crucial to correct for the response bias that is induced by emotional items. In this thesis, face recognition memory was tested as a function of context emotion in which the face had been encoded (chapter 6). The two-high threshold model of Snodgrass & Corwin (1988) was used to analyze the discrimination accuracy (*Pr*) and response bias (*Br*), which has proven its value in recognition memory studies. In the measure of discrimination accuracy (*Pr*), the proportion of hits is corrected by the proportions of false alarms (*p*(hit)-*p*(false alarm)), in order to prevent that accuracy performance is in fact based on a tendency to respond that an item was old. In addition, a measure for response bias (*Br*) is calculated by calculating the probability of saying "old" to an item when the individual is in fact in an uncertain state (*p*(false alarm) divided by *p*(1-*Pr*)).

Recognition memory may depend on factors like the depth of encoding (Bower & Karlin, 1974; Memon & Bruce, 1983; Patterson & Baddeley, 1977). It is assumed that deeper processing of stimuli increases memory performance (Craik & Lockhart, 1972). The effect of presenting faces in an emotional salient context during the study-phase, as these contexts may affect the depth of face encoding, was tested to investigate recognition memory for these faces in the test-phase (chapter 6).

## 1.4 Event-related potential (ERP) studies of face recognition

A shortcoming of behavioral experiments is that response times are indirect measures for face processing, because motor-related movements[2] (the response initiation and response execution) add undesirably to the real time that is needed for processing. Event-related potentials (ERPs) provide a more direct time-course estimate of how the brain processes stimuli. Ultimately, the method may show how neural correlates of face encoding relate to performance on face recognition tasks.

ERPs have been an important method for studying face processing. ERPs are extracted from the ongoing electrocephalogram (EEG) by the averaging method. The EEG is recorded by electrodes that are attached to the scalp. The EEG represents neural activity in the outer mantle of the brain, the cerebral cortex (gyri and sulci). Usually, tens or hundreds of trials of face stimuli are presented in an ERP experiment, in order to be able to average the EEG for a sufficient number of trials. After averaging, an ERP waveform is acquired for each condition (i.e. event related activity) and electrode position.

An advantage of ERPs over behavioral measures is that no explicit response to the target stimulus or conditions is required to investigate how the brain responds to that stimulus (Luck, 2006). For instance, we have investigated ERPs to facial expressions (fearful, neutral) accompanied by context emotions (fearful, neutral), while the participants were instructed to decide on the orientation of the faces (chapter 2). Thus, no explicit response to the emotional content of the stimulus was required. To compare the effect of task conditions across experiments (see Luck, 2006), a categorization task for facial expressions was used in chapter 4. Similarly to chapter 2, the faces were accompanied by contexts.

ERPs provide an upperbound estimation of the time-course at which information processing differs among two or more conditions (Rugg & Coles, 1995). An upper bound

[2] Several studies suggest that emotion processing does not differ on the level of response execution. Leppänen, Tenhunen, & Hietanen (2003) showed by using the lateralized readiness potential (LRP) that there was no difference in response execution times between happy, disgusted and angry facial expressions. Another study found similar response times to facial expressions by using a voice-key as compared to previous studies that have used motor response times (Palermo & Coltheart, 2004).

estimation indicates the time point at which conditions start to differ significantly, which can be determined by difference potentials (Guthrie & Buchwald, 1991; for an example see Thorpe, Fize, & Marlot, 1996). However, this does not necessarily imply that brain activity does not differ before this time point. Brain activity may still differ between conditions before this time point, because not all brain activity can be detected by distant electrodes on the scalp (e.g., a configuration of neurons resulting in closed fields, for instance the amygdala, Coles, Gratton, & Fabiani, 1990).

Another method of analyzing ERP data is selecting ERP components, and identifying peak latencies and amplitudes. An ERP component is based on the synchronous firing of thousands of similarly oriented neurons (Luck, 2006; Rugg & Coles, 1995). ERP components that are associated with face perception are the P1 and N170/VPP, as it has been found that the amplitudes for these components are larger to faces than non-face objects (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Itier & Taylor, 2004b). Most experimental studies have focussed on the N170 component. The N170 component is defined by a negative peak deflection at around 170 ms after stimulus onset, and is the first negative peak (i.e., N1) that follows the P1 in time. The N170 has maximal negative amplitudes on electrodes positioned on the occipito-temporal scalp. The VPP has maximal positive amplitudes on electrodes positioned on the vertex (Bötzel, Schulze, & Stodieck, 1995; Cauquil, Edmonds, & Taylor, 2000; Itier & Taylor, 2004b; Jeffreys, 1989; Joyce & Rossion, 2005). It has been proposed that the N170 and VPP may reflect the same neural generators (Jeffreys, 1989; Joyce & Rossion, 2005; but see for contrasting views Bötzel et al., 1995; George, Evans, Fiori, Davidoff, & Renault, 1996).

The P1 precedes the N170 components in time (although it is unclear whether the N170 amplitude relates to the appearance of the P1). The P1 has its maximal positive amplitude at occipital sites and occurs at around 100 ms after stimulus-onset. Relative to the N170/VPP component, few studies have been concerned with the P1 component, probably because the P1 has often been related to spatial attention and physical features like color, motion and shape (Hillyard & Anllo-Vento, 1998), features that are not easily controlled for in the comparison of faces and non-face objects. In addition, the participants' state of arousal may influence the P1 as well, as has been found in a task requiring discrimination of colored target letters (Vogel & Luck, 2000). Nonetheless, it has been found in several studies that the P1 is increased for faces over non-face objects (Herrmann, Ehlis, Ellgring, & Fallgatter, 2005; Itier & Taylor, 2004b).

A caveat of the ERP method is that it is rather difficult to trace the neural generators from which the scalp measured ERPs arise. The electrical field of the human brain is distorted because of the inhomogeneous properties of the brain. Brain tissue, the skull, cerebrospinal fluid and the scalp smear out the electrical field that is measured on the scalp (Nunez &

Srinivasan, 2006). A number of researchers have attempted to find the generators of the N170 component, and have reported that the N170 is related to activity in the superior temporal sulcus (Henson et al., 2003; Itier & Taylor, 2004c), while others have reported that the N170 may have a source in the fusiform gyrus (Herrmann, Ehlis, Muehlberger, & Fallgatter, 2005; Pizzagalli et al., 2002). Importantly, a correlation has been found between the N170 amplitude and the fMRI BOLD response in the fusiform gyrus (Iidaka, Matsumoto, Haneda, Okada, & Sadato, 2006), although this does not exclude generators in other cortical areas as well. In the general discussion, the ERP results will be discussed in reference to ERP source analyses and the fMRI literature, in order to arrive at hypotheses for further neuroimaging research.

*1.4.1. Recognition of facial expressions*

An important question in ERP research is how ERP components relate to perceptual and cognitive processes (Coles et al., 1990; Luck, 2006). The early ERP components may relate to perceptual encoding, but it is unclear as yet how the magnitude of the P1 and the N170 amplitude relate to face encoding. Attention, structural encoding and emotion perception, or a combination of one or more of these, are perceptual processes that have been related to these ERP components and will be discussed further below.

First, the P1 and N170 amplitude may relate to attention that is allocated to face stimuli. Faces are socially relevant stimuli in our environment and attract attention. ERP experiments have shown for simple basic stimuli that visual selective attention modifies sensory inputs at an early stage of processing. When participants were validly cued to a target, the P1 and N1 amplitude were increased (Mangun, 1995). However, the N1 attention effect appears to be found only when participants have to discriminate among stimuli (Vogel & Luck, 2000). A similar effect may be observed for faces (Campanella et al., 2000), as spatial attention for faces appears to affect the N170 amplitude (Eimer, Holmes, & McGlone, 2003), although others have not observed attentional effects (Cauquil et al., 2000). If individuals are more attentive to facial stimuli as compared to non-face stimuli, possibly because of the social relevance of faces, this may explain the increased N170 amplitudes for faces.

Second, the functional significance of the N170 component has been related to structural encoding of faces (Bentin et al., 1996; Eimer, 2000), which relates to the encoding of the physical invariants of the face. A model by Bruce & Young (1986) implies that (1) structural encoding of faces can be distinguished from non-face objects, and that (2) facial identity and facial expression are processed by relatively parallell routes, and that (3) semantic analyses occurs at a later stage (e.g., identity specific semantic representations). Many studies confirm that the N170 amplitude is larger to faces than non-face objects (Bentin et al., 1996). The insensitivity of the N170 amplitude to familiar faces has been posed as an argument that the N170 is related to structural encoding (Bentin & Deouell, 2000). Consistent with this view, other studies have observed that the N170 amplitude is not influenced when facial identities were shown repeatedly (Henson et al., 2003; Schweinberger et al., 2002). In contrast, other studies have found that the N170 amplitude is decreased when facial identities were shown repeatedly (Heisz, Watter, & Shedden, 2006; Itier & Taylor, 2004a), which poses an argument against the position that the N170 relates to structural encoding.

Third, the N170 component has been related to emotion perception, which is in many respects an argument against the structural encoding hypothesis of the N170. Although a number of studies show that the N170 is not sensitive to facial expressions of emotion (Eimer & Holmes, 2002; Eimer et al., 2003; Holmes, Vuilleumier, & Eimer, 2003), others have found that the N170 amplitude varies across facial expressions of emotion, and more generally, for social stimuli. As is displayed in Table 1.2, many studies have observed an effect for perceiving facial expressions and eye gaze direction. It is unclear as yet what has caused these different results across studies, but difference in experimental designs (e.g., task conditions) or method of analyses (e.g., mean amplitudes versus peak amplitudes) may account for this variance.

The table shows that at least three studies have observed increased N170 amplitudes for fearful faces as compared to neutral faces. A number of studies have in addition shown that the N170 is sensitive to social signals like mouth movement and gaze direction (Puce, Smith, & Allison, 2000), facial motion (Puce et al., 2003), eyes (Bentin et al., 1996; Itier, Latinus, & Taylor, 2006; Taylor, Edmonds, McCarthy, & Allison, 2001), biological motion (Jokisch, Daum, Suchan, & Troje, 2005), and expressional change (Miyoshi, Katayama, & Morotomi, 2004). Effects of emotion have also been observed on the P1 (Batty & Taylor, 2003; Meeren, van Heijnsbergen, & de Gelder, 2005).

A study with schizophrenia patients has shown that the N170 lacked sensitivity to facial expressions of fear and happiness (Campanella, Montedoro, Streel, Verbanck, & Rosier, 2006). These results were significantly different from the results obtained with healthy control participants, and suggest impairment on stages of face encoding (see also Herrmann, Ellgring, & Fallgatter, 2004; Onitsuka et al., 2006).

The ERP results altogether suggest that emotions modify the P1 and/or N170 amplitude, and define importantly the experimental hypotheses that were studied for facial expressions in emotional salient contexts in this thesis.

**Table 1.2**

Event-related potential (ERP) effects for emotional expressions on the P1 and N1/N170.

The top part shows significant effects on P1 and/or N170, the bottom part shows non-significant effect.

**Legend**
x = no information given
$x^{ma}$ = no latency data because of mean amplitudes
$x^{s}$ = sample-by-sample analysis (e.g., t-test analysis of each time-point)
> = for P1 amplitude means more positive, for N1/N170 amplitude means more negative
ns = not significant, $p > .05$
FE = facial expression
B = bodily expression
G = gaze
aFE = affective facial features
dFE = dynamic facial expression
FM = facial movement

Facial expressions: A=Anger; D=Disgust F=Fear; H=Happiness; Sa=Sad; Su=Surprise; N=Neutral

| *Authors* | *Subjects* | *Stimuli* | *Task* | *Time* | *P1 latency* | *P1 amplitude* | *N1/N170 latency* | *N1/N170 amplitude* |
|---|---|---|---|---|---|---|---|---|
| Williams et al. (2006) | N=219 | FE | Attend to faces | 500 ms | ns | F,H < N | ns | F,H > N |
| Caharel et al. (2005) | N=28 | FE | 1.Familiarity (yes/no). | 1000 ms | x | x | $x^{ma}$ | D>H>N |
| Caharel et al. (2005) | N=28 | FE | 2.Facial expression (neutral/smile/disgust) | 1000 ms | x | x | $x^{ma}$ | D>H>N. No interaction task |
| Meeren et al. (2005) | N=12 | FE and B | Facial expression (A, F) | 200 ms | ns | Incongruent > Congruent FE-B | ns | ns: F,A. |
| Klucharev & Sams (2004) | N=11 | FE and G | Indicate repetition of gaze direction | 300 ms | $x^{s}$ | Straight > right averted and H > A. | $x^{s}$ | ns |
| Miyoshi et al. (2004) | N=10 | dFE | Attend to face stimulus | 250 ms | x | x | ns | N170 for smile faces increased when viewing expressional change |
| Stekelenburg & de Gelder (2004) | N=12 | FE | Orientation-decision task | 500 ms | x | x | ns | left hemisphere: F > N |
| Batty and Taylor (2003) | N=26 | FE | Respond to target stimuli (cars, planes, butterflies) | 500 ms | No sgn | N and Su smallest amplitudes. | F, D, S > N, H, Su | F>N, H, D, Su, Sa, A. |
| Pizzagalli et al. (2002) | N=18 | aFE | Passive viewing. | 450 ms | x | dislike faces > liked faces | x | liked faces > disliked faces |
| Watanabe et al. (2002) | N=14 | G | Focus at midway point between eyes of the face | 800 ms | x | x | ns | right hemisphere: right avered > straight eyes |
| Puce et al (2000) | N=20 | FM | Maintain fixation at point in center of the eyes | 500 ms | x | x | Mouth Open < Mouth closed; Eyes Averted < straight | Mouth Open > Closed; Eyes Averted > straight |
| Schupp (2004) | N=20 | FE | Attention to the emotion or orientation of the faces. | 1000 ms | x | x | $x^{s}$ | ns: A, H, N |
| Ashley et al. (2004) | N=16 | FE | Response to immediate stimulus repetition | 750 ms | x | x | ns | orientation x emotion sgn. H,D upr > inv; N inv > upr. |
| Eimer et al. (2003) | N=15 | FE | 1.Decide emotional or neutral expression | 300 ms | x | x | $x^{ma}$ | ns: A, D, F, H, Sa, Su, N. |
| Eimer et al. (2003) | N=15 | FE | 2.Line orientation task | 300 ms | x | x | $x^{ma}$ | ns: A, D, F, H, Sa, Su, N. |
| Holmes et al. (2003) | N=20 | FE | Detect identical stimuli (faces or houses) | 300 ms | x | x | $x^{ma}$ | ns: F, N |
| Eimer & Holmes (2002) | N=18 | FE | Respond to immediate stimulus repetitions | 300 ms | x | x | $x^{ma}$ | ns: F, N |
| Krolak-Salmon et al. (2001) | N=10 | FE | Counting targets of males/females; Counting surprised faces | 400 ms | ns | ns | ns | ns |

### *1.4.2. Recognition memory for faces*

Memory-related ERP effects have been studied for the N170 and for the long-latency components. In a few studies, decreased N170 amplitudes have been found for faces that were repeated as compared to new faces (Heisz et al., 2006; Itier & Taylor, 2004a). Others have not found familiarity effects on the N170 by using faces of celebrities (Bentin & Deouell, 2000; Henson et al., 2003; Schweinberger et al., 2002) and unfamiliar faces (Henson et al., 2003).

More frequently, ERP correlates of recognition memory have been investigated for long-latency ERPs, which will be referred to as the late positive complex (LPC). Two methods of data analyses have been used frequently to analyze these ERP components. First, the statistical significance of a difference between two or more conditions has been determined by using a sample-by-sample t-test analysis (Guthrie & Buchwald, 1991), which shows at what time-points grand average waves differ significantly. It is a useful method for data exploration, especially if no specific hypotheses are formulated.

Second, a related method of time-course analyses is the selection of predefined time-segments, and analyzing the difference for each segment between conditions. For instance, a face recognition memory study is reported in chapter 6, using a window starting from 200 ms after stimulus onset, till 800 ms after stimulus onset, divided into six segments of 100 ms. This time-course analyses is based on previous recognition memory studies that have shown differences between correctly recognized faces as compared to correctly rejected faces on the LPC, as the LPC is more positive-going for correctly recognized faces (Johansson et al., 2004; Joyce & Kutas, 2005; Paller, Gonsalves, Grabowecky, Bozic, & Yamada, 2000; Schweinberger, 1995; Schweinberger et al., 2002; Yovel & Paller, 2004). This time-course analysis by using predefined time-segments has been used frequently in recognition memory studies. Experimental hypotheses can be formulated more precisely by using this method, and therefore it is used for the analyses of the LPC in chapter 6.

A previous study has shown that negative facial expressions increase the positive amplitude of the LPC relative to both positive and neutral expressions (Johansson et al., 2004). In addition, the LPC response to objects was sensitive to the emotional context in which the object had been encoded, as increased LPC amplitudes were found for objects that were encoded in emotional scenes (Smith, Dolan, & Rugg, 2004). In chapter 6, it was investigated how the emotional scene that accompanies faces during encoding influences recognition memory for faces in a test-phase in which no contexts were presented anymore.

## 1.5 Experiments and hypotheses

Behavioral and ERP experiments were conducted to investigate how contexts influence face recognition. A distinction was made between general effects and specific emotion effects.

General context effects were analysed in order to investigate how face processing is affected by meaningful scenes as compared to faces that are presented without a scene context (chapter 2) or in scrambled versions of the scenes (chapter 4, 5, 6). It was expected that meaningful context information distracts attention away from face processing. This will be expressed in smaller N170 amplitudes and worse behavioral performance for faces in meaningful contexts.

Specific emotion effects of contexts in face processing were investigated by using behavioral and ERP methods in healthy participants and schizophrenia patients. We studied the neural correlates of perceiving facial expressions (fear, neutral) in emotional scenes (fearful, neutral) by using ERPs in healthy participants. Scene stimuli were edited in which a a facial expression was positioned centrally (chapter 2). Participants were instructed to indicate the orientation of the presented face (i.e., orientation-decision task). This task was utilized in order to prevent that task performance was related to the emotional content of the stimuli. It was predicted that if fearful faces were accompanied by fearful scenes that the P1 and N170 amplitude would increase (chapter 2). This hypothesis is based on the finding that emotional properties of facial expressions and emotional scenes both influence early perception.

The ERP results of the former study were the starting point to investigate whether the task instructions to categorize facial expressions would affect behavioral and ERP correlates of face processing. It was hypothesized that if emotional salient scenes influence the perception of facial expressions, this may also be reflected on a behavioral level. By using a categorization task for facial expressions (fear, disgusted, happiness) that were accompanied by emotional congruent scenes or incongruent scenes, it was hypothesized that categorization would be better and faster for facial expressions in congruent emotional scenes, for example a facial expression of happiness positioned in front of flowers in contrast to a garbage dump (see Figure 1.1 and chapter 3). Following this experiment, ERPs were acquired to test whether a categorization task for facial expressions would influence the context effects that had been observed previously. Participants were required to categorize facial expressions (fear, happiness) in emotional scene contexts (fear, happy, neutral). Given the emotion effects on the N170 that were found across different task conditions (Table 1.2), it was expected that the N170 amplitude would be increased for fearful faces in fearful scenes as compared to fearful faces in happy and neutral scenes.

A similar experiment was performed with schizophrenia patients. Based on the recognition deficits for facial expressions in schizophrenia, it was hypothesized that categorization of facial expressions would not be influenced by emotional scenes in schizophrenia patients. Because previous studies have shown that emotion recognition deficits in schizophrenia may relate to early stages of encoding, it was expected that schizophrenia patients do not show significant differences on the N170 amplitude as a function of context emotion (chapter 5).

Last, it was examined whether behavioral and ERP measures of face recognition memory are influenced by the emotional scene (fearful, neutral) in which the faces had been studied (chapter 6). A study-test paradigm was used. In the study-phase, neutral faces were presented within fearful or neutral scene contexts. In the test-phase, participants were required to perform an old/new discrimination task. Based on results in the previous experiments, we hypothesized that the N170 would increase for faces that were encoded in emotional scenes compared to faces encoded in neutral scenes. It was further hypothesized that if faces are distracted by emotional relative to neutral contexts, a more positive-going LPC would be found for correctly recognized faces that had been encoded in neutral contexts relative to emotional contexts.

## Chapter 2

## Context influences early perceptual analysis of faces

## An electrophysiological study

*Abstract*

Electrophysiological and haemodynamic correlates of processing isolated faces have been investigated extensively over the last decade. A question not addressed thus far is whether the visual scene, which normally surrounds a face or a facial expression, has an influence on how the face is processed. Here we investigated this issue by presenting faces in natural contexts and measuring whether the emotional content of the scene influences processing of a facial expression. Event related potentials (ERPs) were recorded to faces (fearful / neutral) embedded in scene contexts (fearful / neutral) while participants performed an orientation-decision task (face upright or inverted). Two additional experiments were run, one to examine the effects of context without a face, the other, to evaluate the effects of faces-only. Faces without context showed the largest N170 amplitudes. The presence of a face in a fearful context enhances the N170 amplitude over a face in a neutral context, an effect that is strongest for fearful faces on left occipito-temporal sites. This N170 effect, and the corresponding topographic distribution, was not found for contexts-only, indicating that the increased N170 amplitude results from the combination of face and fearful context. These findings suggest that the context in which a face appears may influence how it is encoded.

A slightly modified version of this chapter has been published as: Righart, R. & de Gelder, B. (2006). Context influences early perceptual analysis of faces. An electrophysiological study. *Cerebral Cortex, 16,* 1249-1257.

## 2.1 Introduction

In everyday life, faces, just like objects, appear within a natural environment. It has been shown previously that natural contexts play an important role in the recognition of objects. A football player, for instance, is usually seen in a football stadium, and his presence is rather unusual in other contexts (Davenport & Potter, 2004).

If the probability is high that a certain context surrounds a visual object, the processing of that object is facilitated, whereas unexpected contexts tend to inhibit it (see for a review Bar, 2004; Davenport & Potter, 2004; Ganis & Kutas, 2003; but see also Hollingworth & Henderson, 1998; Palmer, 1975).

Context influences have been investigated by varying the semantic relationship between a target object and its background context. In such experiments, an object is embedded either in a context in which it is highly expected (congruent object-context compounds) or in a context in which it is highly unexpected (incongruent compounds). Objects that were placed in semantically congruent contexts were recognized more accurately (Davenport & Potter, 2004) and faster (Ganis & Kutas, 2003) than objects in semantically incongruent contexts. Further, the time course of these context effects was related to a decrease on the N400 when target object and context were congruent (Ganis & Kutas, 2003). These findings indicate that knowledge about the visual world may influence our expectations of what objects should appear in a visual scene (Bar, 2004).

A different factor that may influence stimulus processing besides semantic congruency between target object and context is the presence of emotional information (Smith, Dolan et al., 2004; Smith, Henson, Dolan, & Rugg, 2004). Emotional scenes are rapidly evaluated by the visual system for both salient biological (e.g., snake) and artefactual (e.g., gun) elements. Fast evaluation of salient stimuli is reflected in a prioritizing of threatening scenes, as is seen in rapid gaze shifts (Calvo & Lang, 2004) and faster response times for the detection of fear-relevant pictures (Öhman, Flykt et al., 2001) as well as enhanced P1 amplitudes for threatening scenes (Carretie, Hinojosa, Martin-Loeches, Mercado, & Tapia, 2004; Smith, Cacioppo, Larsen, & Chartrand, 2003). Further, single-neuron responses in the right ventral prefrontal cortex for aversive scenes diverge from neutral scenes at around 120 ms (Kawasaki et al., 2001).

Similarly, studies have shown that facial expressions can be discriminated from each other rapidly (White, 1995). In electrophysiological studies, it was observed that the N170 amplitude is larger for faces than for other objects (Bentin et al., 1996; Itier & Taylor, 2004b). The N170 is considered to be associated with structural encoding of the face (Bentin et al., 1996; Eimer, 2000) and may be insensitive for facial expressions (Eimer & Holmes, 2002; Holmes et al., 2003). However, recent studies indicate that the N170 amplitude increases to fearful expressions (Batty & Taylor, 2003; Stekelenburg & de Gelder, 2004). The N170 is also sensitive to social relevant properties of the face (Pizzagalli et al., 2002). Furthermore, effects of facial expression have also been observed at earlier components (~100 ms) at occipital (Batty & Taylor, 2003; Eger, Jedynak, Iwaki, & Skrandies, 2003; Pizzagalli et al., 2002) and frontal electrode sites (Eimer & Holmes, 2002; Holmes et al., 2003) and inferior occipital sources have been found by magneto-encephalography at 110 ms after stimulus onset (Halgren, Raij, Marinkovic, Jousmaki, & Hari, 2000). Further, the extrastriate regions of the brain respond to the emotional intensity of facial expressions, which implicates that visual areas are involved in emotional analysis of stimuli (Surguladze et al., 2003). Thus, the available evidence indicates that the emotional content of faces as well as of scenes is discriminated from neutral content at an early stage of processing. A few studies have already explored the electrophysiological (Guillaume & Tiberghien, 2001; G. A. Rousselet, M. J. M. Mace, & M. Fabre-Thorpe, 2004b; Tsivilis, Otten, & Rugg, 2001) and haemodynamic (Yi et al., 2004) correlates of processing faces in neutral contexts. The influence of emotional contexts on processing faces has not been investigated thus far.

In the present study, we focused on the early electrophysiological correlates (P1 and N170) of facial expressions embedded in emotional contexts. Our hypothesis was that since both emotional attributes of faces and of scenes influence perception, the combination of emotional faces and emotional contexts might increase processing of faces in contexts. If fearful faces are accompanied by a fearful context, the amplitudes of the P1 and N170 component may increase.

## 2.2 Method

### *2.2.1 Participants*

Twelve neurological healthy participants (10 female) with normal or corrected-to-normal vision volunteered ($M = 21.1$ years). All gave informed consent. Ten were right handed. The study was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki.

### *2.2.2 Materials and Procedure*

Stimuli consisted of 24 face photographs (6 females and 6 males posing a fearful or neutral expression) taken from a validated image database (Ekman & Friesen, 1976) and 12 scenes used as contexts (6 fearful related contexts, e.g., knife, crashed car, injection needle; 6 neutral contexts, e.g., house, sofa, guitar). Scenes were selected from the International Affective

Picture System (IAPS) (Lang, Bradley, & Cuthbert, 1999) and complemented with scenes found on the web. Participants were not familiar with the selected stimuli.

The pictures were evaluated by another group of participants (N=15) on arousal (nine-point scale from 1=calm to 9=extremely arousing) and valence (1=very unpleasant to 9=very pleasant). Emotional arousal rates of the fearful pictures were reliably different from neutral pictures (Fearful $M = 7.31$, $SD = 0.64$; Neutral $M = 3.77$, $SD = 0.64$; $t(14) = 14.33$, $P < 0.001$). Fearful pictures were evaluated significantly different on emotional valence than neutral pictures (Fearful $M = 3.43$, $SD = 1.23$; Neutral $M = 6.28$, $SD = 0.67$; $t(14) = 7.08$, $P < 0.001$ ). Every fearful picture was evaluated as being more unpleasant and more arousing than any of the neutral pictures.

All stimuli were greyscale pictures. Faces were overlaid on the centre of a context stimulus in such a way that faces did not occlude the critical parts of the context (Figure 2.1). The height and width of the facial images were 6½ cm x 4½ cm (4.7° x 3.2°) and for context images 20 cm x 30 cm (14.3° x 21.2°) respectively. Participants sat in an electrically shielded cabin at 80 cm distance from the monitor.

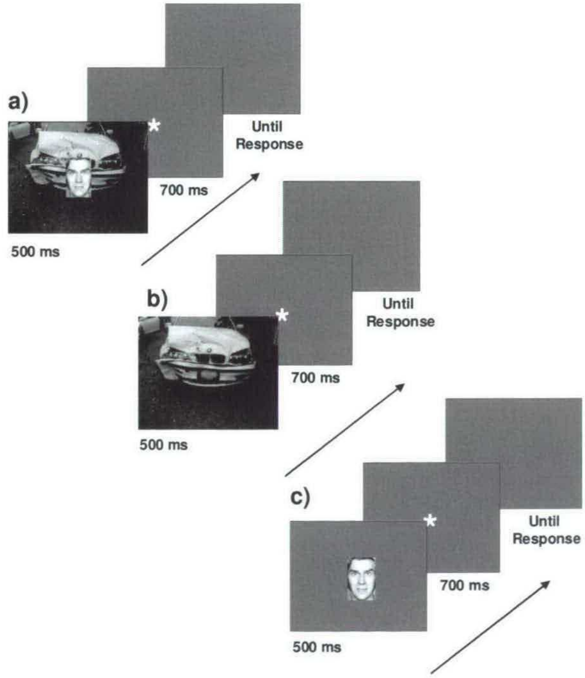



**Figure 2.1 Stimuli.** Design of (a) faces in context, (b) contexts-only and (c) faces-only experiment. Participants were instructed to indicate the orientation of the face in the faces in context and faces-only experiment. Participants had to detect a target trial (i.e., telephone) among context trials in the contexts-only experiment.

This study comprised three experiments: a main experiment in which face-context compounds were presented, and two additional experiments containing faces-only and contexts-only. The order of the three experiments was counterbalanced across participants. All individual experiments were preceded by a short training session to get familiarized with the procedures.

The face-context compounds were shown for 500 ms (Figure 2.1). Stimuli were followed by a central fixation cross which was presented for 700 ms. Participants performed an orientation-decision task, in which they had to indicate the orientation of the face (upright or inverted) by pressing one of two buttons after disappearance of the fixation cross. This delay in response was used in order to avoid contamination by motor response related artefacts. The orientation task was chosen to prevent any task-related effects for emotions in face or context. Compounds were created of fearful faces with fearful contexts and neutral contexts, and neutral faces with the same fearful contexts and neutral contexts. These compounds were created for upright and inverted faces resulting in eight conditions. Each condition contained 72 unique face-context combinations. Stimuli were delivered in three separate blocks of 192 trials. Each of the blocks contained each condition equally (i.e., 8 x 24 face-context combinations).

In the contexts-only experiment, the same context stimuli were presented but without faces. Target trials (i.e., a telephone) were presented in between. Participants were instructed to press a button when the target trial was presented and the remaining button for a non-target context. The non-target context stimuli were the same stimuli as the backgrounds in the face-context experiment. We hypothesized that if contexts influence ERP components of face processing, then this ERP effect should be different from potentials generated by contexts-only.

Another experiment was run to control for the effects of face stimuli separate from context. The same face stimuli but without contexts (i.e., faces-only) were presented to provide a baseline against which faces with contexts could be compared. The face-stimuli were overlaid on a grey background, similarly sized to the face and background in the face-context experiment. As in the face-context experiment, participants had to indicate the orientation of the face. All other specifications are similar to the face-context experiment described above. We hypothesized that if contexts affect face processing, the ERP components for faces-only could be differentiated from components generated for faces in contexts.

*2.2.3 EEG Recording*

EEG was recorded from 49 locations using active Ag-AgCl electrodes (BioSemi Active2) mounted in an elastic cap, referenced to an additional active electrode (Common Mode Sense) during recording. EEG signals were band-pass filtered (0,1 Hz-30 Hz, 24 dB/Octave). The sampling rate was 256 Hz. All electrodes were off-line referenced to an average reference. Horizontal EOGs were recorded from two electrodes placed at the outer canthi of both eyes. Vertical EOGs were recorded from electrodes on the infraorbital and supraorbital regions of the right eye in line with the pupil. Raw EEG data were segmented into epochs starting 100 ms before stimulus onset to 900 ms after the stimulus onset. Data were baseline corrected to the first 100 ms of the epoch.

After EOG correction using the algorithm of (Gratton, Coles, & Donchin, 1983), epochs with amplitude exceeding 100 μV at any channel were rejected from analyses. ERPs were averaged for each condition. Only trials on which participants responded correctly were averaged. For the contexts-only experiment, non-target trials were averaged in order to compare ERPs for contexts-only with faces in contexts. The target trials were shown infrequently to maintain fixation and attention and were therefore not analyzed. Electrode selection was based on previous studies showing maximal amplitudes for P1 and N170 on these sites (Batty & Taylor, 2003; Stekelenburg & de Gelder, 2004). Peak latency and amplitude of P1 were scored at occipital sites (O1/2) and parieto-occipital sites (PO3/4) as the maximal positive peak in the time window 100 – 150 ms with respect to baseline. For N170, peak latency and amplitude were scored at occipito-temporal sites (P5/6, P7/8 and PO7/8) as the maximal negative peak in the time window 140 – 220 ms (for electrode positions, see Figure 2.2).

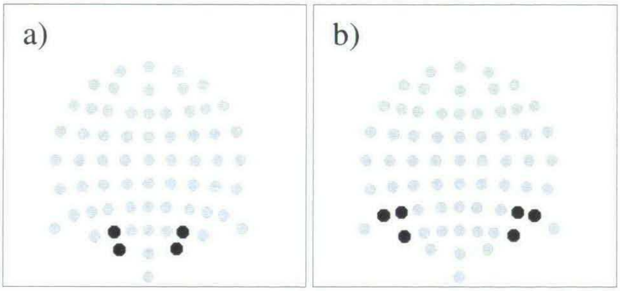



**Figure 2.2 Electrode set up.** Scalp locations on which P1 *(a)* and N170 *(b)* were measured.

*2.2.4 Analyses*

P1 and N170 latencies and amplitudes for faces in contexts were analyzed with multivariate analyses for repeated measures containing the within-subject factors facial expression (fearful, neutral), context emotion (fearful, neutral), hemisphere (left, right), and electrode position.

Additional comparisons were made for the faces-only and contexts-only experiments. First, for faces-only, fearful / neutral faces in contexts were tested against fearful / neutral faces without a context by adding the factor context presence (present, absent). Second, for contexts-only, the effects of faces in fearful / neutral contexts were tested against fearful / neutral context without a face by adding the factor face presence (present, absent). Scalp topographic distributions were then analyzed for the N170 component for the comparison of faces in fearful / neutral contexts with fearful / neutral contexts-only. Difference waves (i.e. fearful contexts – neutral contexts) were calculated separately for faces in context and contexts-only. Amplitudes were vector normalized according to the method described by McCarthy and Wood (1985) and subjected to repeated measures ANOVA. Additionally, difference waves (i.e. faces in context – context-only) were calculated to test scalp distributions of fearful contexts against neutral contexts. *P*-values were corrected by Greenhouse-Geisser epsilon correction, if appropriate. Statistics are indicated with original degrees of freedom (Picton et al., 2000).

Finally, face inversion effects for P1 and N170 latencies and amplitudes were analyzed with multivariate analysis for repeated measures containing the within-subject factors facial expression (fearful, neutral) and orientation (upright, inversion).

## 2.3 Results

*2.3.1 Behavioral results*

Performance on the orientation-decision task (upright vs inverted) was nearly flawless. The accuracy for upright faces was 99.2% and for inverted faces 99.1%. No significant differences were found among any of the conditions ($p > 0.05$). Response times were not analyzed because of the delayed-response paradigm.

*2.3.2 ERP results*

*P1 Latency*

There was a main effect of facial expression, $F(1,11) = 9.09$, $P < 0.05$, in that latencies were prolonged for fearful ($M$ = 135 ms) as compared to neutral ($M$ = 133 ms) faces. No main effect of context emotion was observed, $p > 0.05$. The factor face presence was added to the analysis to test whether contexts with faces could be differentiated from contexts without faces. The interaction between face presence and context emotion was near significance, $F(1,11) = 4.15$, $p = 0.07$. Post-hoc comparisons showed that latencies were prolonged for neutral context containing a face ($M$ = 135 ms) as compared to neutral contexts without a face ($M$ = 129 ms) ($p < 0.05$). Further, faces in contexts were compared to faces without a context. Latencies did not vary significantly as a consequence of context presence (all $ps > 0.05$). The latency difference between fearful and neutral faces was not significant for faces without any context ($p > 0.05$).

*P1 Amplitude*

No effects were found for facial expression ($p > 0.05$). The main effect for context emotion on P1 amplitude was marginally significant, $F(1,11) = 4.04$, $p = 0.07$, reflecting increased amplitudes for faces accompanied by fearful contexts ($M$ = 10.91 μV) as compared to neutral contexts ($M$ = 10.28 μV). However, this effect was not specific for faces in context, as there was no significant interaction between context emotion and face presence ($p > 0.05$).

A main effect was found for face presence, $F(1,11) = 26.28$, $p < 0.001$, as amplitudes were larger for contexts in which a face was present ($M$ = 10.59 μV) than absent ($M$ = 9.52 μV). The amplitudes were marginally increased for faces without a context ($M$ = 11.76 μV) compared to faces in a context ($M$ = 10.59 μV), $F(1,11) = 3.84$, $p = 0.07$), reaching significance on the posterior electrodes O1/2 ($p < 0.05$). Like faces in context, no main effects were found for facial expression without a context ($p > 0.05$).

*N170 Latency*

Latencies did not vary significantly as a consequence of context emotion ($p > 0.05$). A main effect was observed for facial expression, $F(1,11) = 12.13$, $p < 0.01$, as reflected by longer latencies for fearful faces ($M$ = 186 ms) than neutral faces ($M$ = 182 ms). The factor facial expression had a significant interaction with hemisphere and electrode position, $F(2,10) = 6.42$, $p < 0.05$. Post-hoc comparisons showed that latencies for fearful faces were longer than neutral faces on each electrode, but reached significance only for the right electrode P6 ($p < 0.05$). Latencies did not vary as a consequence of face presence ($p > 0.05$).

For faces-only, latencies were longer for fearful faces ($M$ = 189 ms) over neutral faces ($M$ = 182 ms) as well, $F(1,11) = 29.28$, $p < 0.05$. This difference was significant at each electrode ($p < 0.05$).

*N170 Amplitude*

The factor electrode position showed a main effect, $F(2,10) = 37.39$, $p < 0.001$, as N170 amplitudes were more negative for electrodes P7/8 ($M$ = -2.22 µV) compared to electrodes P5/6 ($M$ = +1.74 µV) and PO7/8 ($M$ = +1.84 µV). No main effects were found for hemisphere or facial expression ($p > 0.05$). A main effect was observed for context emotion, $F(1,11) = 28.63$, $p < 0.001$), but this effect was qualified by an interaction with hemisphere, $F(1,11) = 12.16$, $p < 0.01$. Post-hoc comparisons showed that N170 amplitudes were more negative for faces in fearful contexts as compared to faces in neutral contexts, but only significantly for electrodes on the left hemisphere, $t(11) = 7.90$, $p < 0.001$. As indicated in Figure 2.3, N170 amplitudes were more negative for faces accompanied by fearful contexts ($M$ = -3.34 µV) than those accompanied by neutral contexts ($M$ = -1.05 µV) on electrode P7, $t(11) = 6.07$, $p < 0.001$) but not on electrode P8 ($M$ = -2.21 µV and $M$ = -2.29 µV resp.). Indeed, all 12 participants showed this pattern on electrode P7 (binomial $P(12/12) < 0.001$).

To test whether this effect is truly explained by the face-context combination, comparisons were made to contexts-only. Therefore, the factor face presence (collapsed across facial expressions) was added to the repeated measures model. First, a main effect was found for face presence, $F(1,11) = 8.60$, $p < 0.05$) in that amplitudes were more negative for contexts in which a face was present as compared to contexts in which a face was absent (Fig 2.3).

Second, an interaction was found between context emotion and face presence, $F(1,11) = 12.59$, $p < 0.01$, for which post-hoc comparisons showed profound differences between fearful and neutral contexts when a face was present, $t(11) = 5.35$, $P < 0.001$, whereas no significant differences were observed between fearful and neutral contexts for which the face was absent ($p > 0.05$).

In addition, scalp topographies for this interaction were significantly different, $F(48,528) = 2.43$, $p < 0.05$ (Figure 2.4). This interaction critically suggests that the observed effects were caused by the combination of face and context and not by contexts only and may suggest that the underlying sources or relative source strengths are (at least partly) different (Picton et al., 2000). The topography for faces in fearful contexts against faces in neutral contexts were not significantly different, $F(48,528) = 2.15$, $p > 0.05$ (see Figure 2.4).

In addition to the two-way interaction between context emotion and hemisphere, a three-way interaction was observed for facial expression, context emotion and hemisphere, $F(1,11) = 5.11$, $p < 0.05$. In post-hoc comparisons it was found that the N170 amplitudes were more increased for fearful faces than for neutral faces, when they were in the presence of a fearful context, but only on left hemispheric electrodes, $t(11) = 3.82$, $p < 0.01$. As illustrated for P7/8 in Figure 2.3, sharply enhanced negativities were found for fearful faces in the presence of a fearful context on electrode P7 ($M$ = -4.11 μV) as compared to neutral faces in a fearful context ($M$ = -2.57 μV), $t(11) = 3.10$, $p < 0.05$. 10 out of 12 participants confirmed to this pattern (binomial $P(10/12) < 0.05$) on electrode P7, whereas 9 out of 12 participants (binomial $P(9/12) = 0.07$) showed such an effect on site P8 but that difference did not reach significance on post-hoc testing ($p > 0.05$).

To interpret the effects of faces in context against faces without context information, the effect of context presence was analyzed. As indicated in Figure 2.3, a main effect was found for context presence, $F(1,11) = 31.48$, $p < 0.001$. N170 amplitudes were more negative for faces in which the context (P7 = -5.73 μV; P8 = -5.92 μV) was absent than faces in which the context was present (P7 = -2.20 μV; P8 = -2.25 μV).

For faces without contexts, a main effect was found for facial expression, $F(1,11) = 5.38$, $p < 0.05$. Amplitudes of fearful expressions ($M$ = -3.25 μV) were more negative than neutral expressions ($M$ = -2.53 μV). However, the interaction with hemisphere was not significant ($p > 0.05$).

*Inversion effects*

No significant differences were found between upright and inverted faces on P1 latencies and amplitudes neither for faces in context nor for faces-only ($p > 0.05$). Although for P1 amplitudes in the faces-only condition a significant interaction was observed between orientation and hemisphere ($F(1,11) = 5.26$, $p < 0.05$), this effect was not explained by inversion ($p > 0.05$) but by P1 amplitudes being larger on right electrodes ($M$ = 12.55 μV) than left electrodes for upright faces ($M$ = 10.97 μV) ($p = 0.08$).

N170 latencies were prolonged for inverted ($M$ = 189 ms) as compared to upright faces in context ($M$ = 184 ms), $F(1,11) = 15.26$, $p < 0.01$. An interaction between orientation and facial expression was found for the N170 amplitude, $F(1,11) = 20.25$, $p < 0.001$. Inversion effects on N170 amplitude were found for neutral faces but not for fearful faces in context. For neutral faces, amplitudes were more negative for inverted ($M$ = -0.26 μV) than upright faces ($M$ = 0.51 μV), $t(11) = 2.67$, $p < 0.05$ whereas differences were not significant for fearful faces ($p > 0.05$).

Similar results were obtained for faces-only as for faces in context. N170 latencies were longer for inverted ($M$ = 196 ms) than upright faces ($M$ = 186 ms), $F(1,11) = 30.30$, $p < 0.001$. Again, an interaction between orientation and facial expression was found for the N170 amplitude, $F(1,11) = 7.65$, $p < 0.05$). For neutral faces, amplitudes were more negative for inverted ($M$ = -3.69 μV) than upright faces ($M$ = -2.53 μV), $t(11) = 2.26$, $p < 0.05$, whereas differences were not significant for fearful faces ($p > 0.05$).

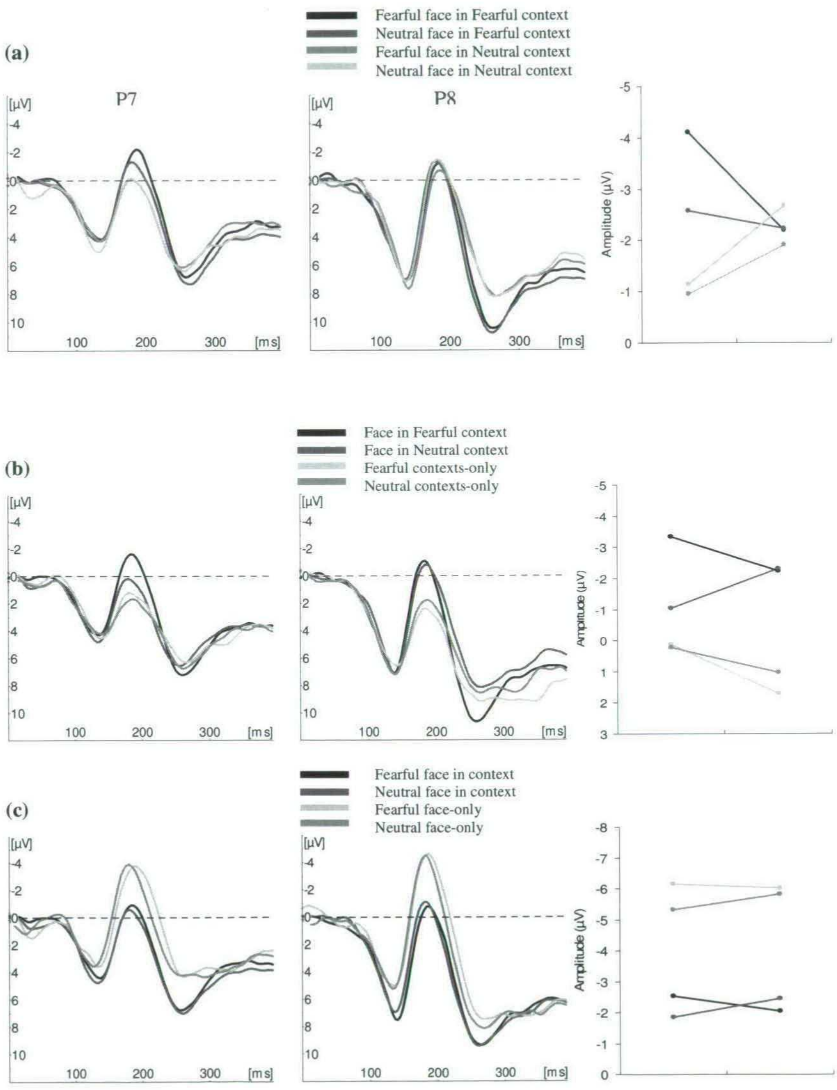



**Figure 2.3. Grand Average ERP waveforms and plots of the N170 amplitudes.** Electrode site P7 (left) and P8 (right) are displayed. Negative amplitudes are plotted upward. *(a)* Faces in Context. N170 amplitudes were enhanced for faces in fearful contexts, particularly for fearful faces on left occipito-temporal sites. *(b)* Contexts-only. There were larger amplitudes for faces in fearful contexts than neutral contexts. These differences were not observed for contexts-only. *(c)* Faces-only. N170 amplitudes were larger for faces-only than faces in context.

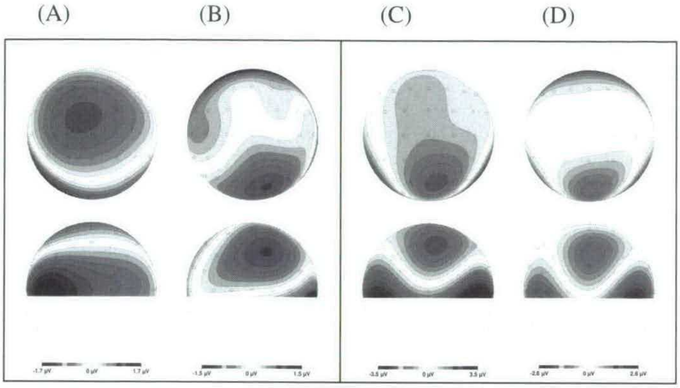



**Figure 2.4 Scalp topographies.** Scalp topography at 180 ms for (*a*) faces in fearful contexts minus faces in neutral contexts tested against (*b*) fearful contexts-only minus neutral contexts-only.
Although a topographic difference is present between fearful and neutral contexts-only (in b), reflected as a parietal negativity, a similar topography is anticipated if face presence had no effect (in a). However, a different topography is seen, with a clear left occipito-temporal negativity and central positivity.
(*c*) Scalp topography at 180 ms for faces in fearful context minus fearful contexts-only tested against (*d*) faces in neutral context minus neutral contexts-only. A similar topography with clear occipito-temporal negativities is seen.

## 2.4 Discussion

We investigated the electrophysiological correlates of perceiving faces in emotional contexts. Several important findings emerge from this study. Firstly, the N170 for face processing is sensitive to the presence of context information, which reduces the amplitude, compared to isolated faces. Secondly and more importantly, the N170 is influenced by the emotional information the context provides. The N170 amplitude was more negative when a face, particularly a fearful face, was presented in a fearful context as compared to a neutral context.

### *2.4.1 General context effects*

A comparison of the ERP components for faces without a context *vs.* faces within a context indicates that early stages of face processing are sensitive to the presence of a context picture. Given the functional meaning of the stage of processing as reflected by the N170, the question arises whether the presence of a context leads to a reduced structural encoding. Previous findings indicate that frontal face views compared to profiles and to faces not clearly standing off from the background yield higher N170 amplitudes (Eimer, 2000). This may indicate better structural encoding for the isolated faces. Likewise, it has been found that N170 amplitudes increase linearly as the level of noise around a face is decreased stepwise (Jemel et al., 2003). The data therefore suggest that structural encoding will be better when faces are shown without a context picture. This functional interpretation of the higher N170 amplitudes should be tested in a subsequent memory experiment. But at present this interpretation is consistent with available behavioral data. Higher identification rates have been found for target objects that were isolated from any context information as opposed to targets in either congruent or incongruent contexts (Davenport & Potter, 2004).

An alternative view is that the N170 amplitudes for faces in contexts are smaller than for isolated faces because of the perceptual load of the context information. In terms of Lavie (1995), it may be hypothesized that load in processing the target relevant stimulus (i.e., the face) determines the degree of processing target irrelevant information (i.e., the context). It was shown that irrelevant distracter letters could interfere with the processing of relevant target letters only if perceptual load was low (Lavie, 1995). Similarly, Yi *et al.* (2004) observed that task irrelevant scenes activated scene sensitive areas of the brain (i.e., parahippocampal place area, a region of the medial temporal cortex), only if face discrimination was relatively easy. In the present study perceptual load was low as well, which is reflected in the high accuracies on the face orientation task. Therefore, it may be hypothesized that the task related stimulus did not demand all available attentional capacity. The irrelevant context may have captured the remaining capacity.

The context information may capture attention and therefore decrease the N170 amplitude (Cauquil et al., 2000; Holmes et al., 2003), which may influence face encoding. A related finding has been reported in a study by Rolls *et al* (2003), in which macaque monkeys were shown objects against either a grey context or against a complex natural scene context. Single-neuron recordings showed that the receptive field (i.e., region of the visual field that causes a visual neuron to respond) of the inferior temporal cortex neurons was reduced if complexity of the surrounding context increased. It was suggested that this is due to attentional competition between object and context processing (Rolls et al., 2003). A similar mechanism may underlie the effects for contexts on face processing in the present study[3].

However, it should be noted that it is unlikely that all attention was dedicated to the processing of contexts-only. The P1 and N170 amplitude were both larger for faces in contexts than contexts-only. Previous data have shown that the P1 and N170 are both increased for faces compared to other objects (Itier & Taylor, 2004a) or scenes (Puce, Allison, & McCarthy, 1999). The time-course of these stages of face processing is consistent with single cell recordings (Sugase, Yamane, Ueno, & Kawano, 1999). The data implicate that faces were effectively differentiated from their contexts. The data replicate previous findings

---

[3] It should be noted here that perceptual load could be better distinguished from structural encoding by using phase-scrambled scenes. According to the structural encoding hypothesis, face encoding should not be different in scrambled or intact scenes. The perceptual load hypothesis would still predict stronger competition from the intact scene. However, we decided to select grey backgrounds in order to reliably replicate the design of earlier studies. We thank an anonymous reviewer for this suggestion.

for the N170 obtained for faces in neutral contexts (Guillaume & Tiberghien, 2001; Rousselet et al., 2004b).

### *2.4.2 Specific emotion effects of contexts*

The central issue of the present study concerns not the extent of structural encoding in isolated vs. contextualized faces but the differential effect of emotional contexts observed on P1 and N170. Comparing the ERP components for faces in fearful contexts to faces in neutral contexts showed that the N170 amplitude for faces is sensitive to the emotion in context. N170 amplitudes were more negative for faces in a fearful context as compared to neutral contexts. An additional experiment containing contexts without a face served as a control to show that the differential effects for emotion in face-context compounds were not generated by the contexts-only.

A number of arguments plead against an explanation of the observed N170 effects by contexts-only. First, whereas the critical difference was found on N170 amplitude for faces in fearful compared to neutral contexts, no significant difference was found between fearful and neutral contexts without a face. This difference between faces in contexts and contexts-only was also confirmed by scalp topography analyses. It is therefore unlikely that low-level features of the context have generated these differential ERPs. Secondly, participants made correct face orientation decisions and presentation of inverted faces resulted in the signature inversion effect of prolonged N170 latencies and increased N170 amplitudes (Bentin et al., 1996; Itier & Taylor, 2004b; Rousselet et al., 2004b). Therefore, these arguments do not support an explanation of differential emotion effects on N170 by contexts-only. We conclude that the findings are critically dependent on the combination of face and context.

An important three way interaction was observed between facial expression, context emotion and hemisphere. N170 amplitudes for fearful faces differed significantly from neutral faces on left occipito-temporal sites, but only if faces were presented in a fearful context. Fearful contexts may add important information about the facial expression (see discussion below), which may influence encoding. However, it cannot be determined from these data whether fearful contexts enhance processing of a fearful face or whether neutral contexts suppress processing of a fearful face compared to a neutral face. Future work should address whether congruency effects are present for specific categories of facial expressions (e.g., fear, disgust, happiness) that are either matched or mismatched with their accompanying contexts.

The finding that face processing is sensitive to the emotional context as reflected in increased N170 amplitudes is consistent with fMRI data. Facial expressions of fear and fearful scenes both increase amygdala responses (Hariri, Tessitore, Mattay, Fera, & Weinberger, 2002) and increase the response in the fusiform gyrus (Lang et al., 1998; Morris et al., 1998; Surguladze et al., 2003). Activation of the fusiform gyrus is associated with the intensity of fearful expressions (Morris et al., 1998; Surguladze et al., 2003). Enhanced responses in fusiform gyrus may indicate feedback modulation from the amygdala to the fusiform gyrus (Morris et al., 1998; Vuilleumier, Richardson, Armony, Driver, & Dolan, 2004). As the N170 may have its source in the fusiform gyrus (Pizzagalli et al., 2002) its enhancement may reflect feedback modulations from the amygdala to enhance structural encoding. The result that P1 and N170 latencies were prolonged for fearful facial expressions may at first hand seem counter-intuitive. However, slower N170 peak latencies for fear over neutral faces have been reported previously (Batty & Taylor, 2003) and may on the other hand reflect prolonged activation to encode available information. Future studies should establish the meaning of peak latencies in processing facial expressions.

The observed effects for the N170 on the left occipito-temporal electrode sites may be related to a predominant response of the left hemisphere to emotion. Although most neuroimaging studies have found a bilateral response of amygdala and fusiform gyrus to emotion in scenes (Lang et al., 1998) and facial expressions (Morris et al., 1998), or a predominant right fusiform gyrus response (Surguladze et al., 2003), a few recent studies have shown left fusiform gyrus responses (Kim et al., 2004; Taylor, Phan, Decker, & Liberzon, 2003). Kim *et al.* (2004) have observed left fusiform gyrus responses to sad faces that were cued by negative sentences (e.g., about losing money) compared to positive sentences (e.g., about winning money). The authors propose that right hemisphere responses may be related to the ambiguous valence of some facial expressions, whereas left hemisphere responses may be evoked when valences are clearly determined (Kim et al., 2004; Phelps et al., 2001). In the present ERP study, fearful contexts possibly fill in the ambiguous nature of the facial expressions, which may have introduced the left hemispheric effects. However, further source analyses and fMRI studies should determine the role of the left hemisphere in processing emotions.

Alternatively, the N170 may have its source in the superior temporal sulcus (Henson et al., 2003; Itier & Taylor, 2004c), which has been related to social perception (Allison, Puce, & McCarthy, 2000). In threatening circumstances, finer encoding and sustained analysis of faces may be necessary to prepare an appropriate reaction and act adaptively. In accordance with this, in earlier studies it has been shown that the N170 is sensitive to socially relevant cues, such as eyes (Bentin et al., 1996), direction of eye-gaze (Watanabe, Miki, & Kakigi, 2002), facial motion (Puce et al., 2003), biological motion (Jokisch et al., 2005), affective facial features (Pizzagalli et al., 2002), facial expressions (Batty & Taylor, 2003; Stekelenburg & de Gelder, 2004) and expressional change (Miyoshi et al., 2004). Consistent with this, several neuroimaging and single cell-recording studies have shown that the superior temporal sulcus is activated by socially relevant information (Allison et al., 2000).

The N170 component has been related to structural encoding that is utilized for facial identification (Eimer, 2000). The present results do not allow a definite answer to the question whether the larger N170 amplitudes reflect enhanced encoding of either identity or expression of the face, or both, and whether this predicts better recall of facial identity (Haxby et al., 2000). Similar to studies on the effect of emotional scene context on object recognition and memory, future studies of face recognition need to determine how the presence of an emotional context influences encoding and subsequent retrieval (Erk et al., 2003; Smith, Dolan et al., 2004; Smith, Henson et al., 2004).

*Acknowledgements*

We are grateful to J. Stekelenburg, R. Mark and W.A.C. van de Riet and two anonymous reviewers for helpful comments and suggestions on the manuscript.

# Chapter 3

## Recognition of facial expressions is influenced by emotional scene gist

## A behavioral study

*Abstract*

Recognition of facial expressions has traditionally been investigated by presenting facial expressions without any context information. The present study addressed the question whether emotional scenes influence the explicit recognition of facial expressions. In three experiments, participants were required to categorize facial expressions (disgust, fear and happiness) that were shown against a background consisting of natural scenes with either a congruent or incongruent emotional significance. A significant interaction was found between facial expressions and the emotional content of the scenes, showing a response advantage for the recognition of facial expressions accompanied by congruent scenes. This advantage was robust across different task conditions and presentation durations. Taken together, the results show that the surrounding scene is an important factor in recognizing facial expressions. We conclude that the influence of emotional gist on recognition of facial expressions is robust enough to withstand short presentation times and concurrent attentional load.

A slightly modified version of this chapter has been submitted as: Righart, R. & de Gelder, B. (2007) Recognition of facial expressions is influenced by emotional scene gist.

## 3.1 Introduction

Facial expressions provide the observer with very salient emotional and communicative cues. Many behavioral studies have shown how humans and non-human animals express and recognize emotions (Darwin, 1998; Ekman, 1992). Darwin argued that many facial expressions can be traced back to the concrete circumstances in which they originated (Darwin, 1998). Contemporary theorists have occasionally revisited this theme, and have shown that when we perceive a facial expression, we in fact tacitly fill-in information about the person's reaction to and interaction with the environment for which this facial expression is appropriate. When recognition of facial expressions is tested, and the facial expression stimuli only show the person's head and shoulders without any context information, individuals nevertheless perceive the facial expression as part of a person's response to an event in his or her environment (e.g., they report that the person is shielding from something, withdrawing from something, accepting or not accepting something; (Frijda & Tcherkassof, 1997). Thus, it appears that faces are perceived, not as displaying a facial expression *perse*, nor as signaling some emotional state, but as a person actively responding to an event in the environment (Frijda, 1986; Frijda & Tcherkassof, 1997). In the present study, we investigated whether emotional scene information influences how we recognize facial expressions.

The issue of context influences on the recognition of facial expressions addressed by our study has not been investigated in the literature so far, and therefore specific predictions are scarce. Nevertheless, based on existing studies of face perception on the one hand and scene perception on the other, different outcomes may be envisaged when a facial expression needs to be recognized in the presence of emotional scene information. One possibility is that emotional scenes hamper face processing by diverting attention away from faces. This perspective is based on the finding that individuals prefer to direct their attention towards fearful scenes (Blanchette, 2006; Calvo & Lang, 2005; Öhman, Flykt et al., 2001; see also Tipples, Young, Quinlan, Broks, & Ellis, 2002). This attentional bias for fearful scenes may divert attention from the primary task. For tasks other than face recognition, for example a bar-orientation task, target processing is slowed down when the target is placed in a fearful scene (Erthal et al., 2005; see also Algom, Chajut, & Lev, 2004). If this effect generalizes across different kind of targets, a similar effect may be observed for faces in contexts.

An alternative view holds that facial expressions have a processing advantage irrespective of what context information is present. Eastwood, Smilek & Merikle (2003) have shown that negative facial expressions slow down target processing, even when the target bears no direct relationship to the facial expression (i.e. counting component parts of schematic faces). Presumably, negative faces capture attention in the early stages of processing. This is suggested by findings from experiments using detection tasks that have shown faster response times for negative facial expressions (Hansen & Hansen, 1988; Öhman, Lundqvist et al., 2001). However, response times may depend on the experimental setting and the task that is used. If participants have to discriminate among facial expressions, happy faces tend to be discriminated faster than other expressions, for instance expressions of anger, fear and disgust (Calder et al., 2000; Kirita & Endo, 1995; Kirouac & Dore, 1983; Leppänen et al., 2003; Palermo & Coltheart, 2004). Differences between detection and discrimination may be explained by task requirements, as discrimination requires a more detailed analysis (Leppänen et al., 2003). Therefore, negative facial expressions may not only capture attention, but also hold attention, resulting in longer response times (see also Fox et al., 2001). However, these studies did not manipulate context, and thus far it is unclear how emotional information from face and context affect processing.

A third possibility is that recognition of facial expressions is influenced by concurrently presented emotional scenes, in such a way that information from facial expressions and emotional scenes will interact in the course of processing. This is predicted by a few recent studies presenting other emotional stimuli in the context of faces. Event-related potential studies suggest that facial expressions and information from contexts such as the concurrently presented bodily expressions (Meeren et al., 2005) or emotional scenes (chapter 2) interact on an early stage of processing. In addition, behavioral data have shown that individuals are better and faster in identifying facial expressions of fear and anger when they were accompanied by congruent bodily expressions compared to bodily expressions that were not congruent to the facial expression (Meeren et al., 2005). Comparable congruency effects of emotion have been observed for negative and positive word stimuli that were overlaid on emotional congruent or incongruent line-drawing of animals (de Houwer & Hermans, 1994).

The present study addressed the question whether emotional scenes affect how facial expressions are recognized. We used an explicit recognition task in which participants were required to categorize a facial expression (disgust, fear, happiness) presented together with a scene. The emotional scenes were selected for evoking disgust, fear and happiness. We hypothesized that the categorization of facial expressions of emotions is improved if the context has a similar emotional value (i.e., congruency effect). In contrast, recognition is hampered when the emotion of the facial expression and that of scene context are incongruent.

## 3.2 Experiment 1

In experiment 1, facial expressions of fear, disgust and happiness were presented in emotional congruent and incongruent context scenes. Participants were required to categorize facial expressions.

*3.2.1 Method*

*3.2.1.1 Participants*

Fifteen neurologically healthy undergraduate students (12 female) from Tilburg University participated for course credits. All participants ($M = 21.8$ years) were right handed and had normal or corrected to normal vision. Informed consent was given by all participants.

*3.2.1.2 Stimuli and procedure.*

Materials consisted of face- and context stimuli. Face stimuli consisted of 36 color face photographs balanced for gender, taken from the Karolinska Directed Emotional Faces set (Lundqvist, Flykt, & Ohman, 1998). Facial expressions were happiness (12), disgust (12) or fear (12). Context stimuli were 36 pictures of scenes taken from the International Affective Picture System (Lang et al., 1999) and selected from the web. Context scenes depicted emotions of happiness (12), disgust (12) and fear (12). The scene stimuli were selected after a validation procedure. For this validation, another group of participants (N=15) was recruited in which pictures were judged for arousal (nine-point scale from 1=calm to 9=extremely arousing) and valence (1=very unpleasant to 9=very pleasant). The pictures were presented randomly along with non-emotional pictures. A main effect was found for emotional arousal, $F(3,42) = 95.63$, $p < 0.001$, which was highest for fear ($M = 7.59$, $SD = 0.87$), followed by disgust ($M = 7.11$, $SD = 0.76$) and happiness ($M = 3.38$, $SD = 1.07$), $t(14) = 3.61$, $p < .01$ for comparison fear-disgust, $t(14) = 12.02$, $p < .001$ for fear-happiness, and $t(14) = 12.12$, $p < .001$ for disgust-happiness. An effect was also found for emotional valence, $F(3,42) = 76.54$, $p < 0.001$, with rates that were highest for happy pictures ($M = 7.71$, $SD = 0.68$), followed by fearful pictures ($M = 3.33$, $SD = 1.42$) and disgust pictures ($M = 2.75$, $SD = 1.02$), $t(14) = 2.54$, $p < .05$ for fear-disgust, $t(14) = 8.70$, $p < .001$ for fear-happiness, and $t(14) = 12.04$, $p < .001$ for disgust-happiness.

Faces in context were constructed by positioning faces centrally upon emotional scenes (For an example, see Figure 3.1). Faces and emotional scenes were carefully edited and sized in such a way that face pictures did not occlude parts of the context pictures that could possibly be critical for recognition. The height and width of the facial images were 9.0 x 7.0 cm (6.4° x 5.0°) and for the emotional scenes 26.0 x 35.0 cm (18.5° x 24.7°). Facial expressions of fear (F), disgust (D) and happiness (H) were paired with emotional scenes of fear (F), disgust (D) and happiness (H), and were presented in three separate blocks.

Face-context combinations were presented for as long as participants needed to respond with a maximum presentation-time of 6000 ms, and were presented on a monitor 80 cm in front of the participant. Participants performed a two-alternative forced choice task on facial expressions. They were instructed to respond as accurately and as fast as possible. Responses were recorded from stimulus-onset. No feedback was given.

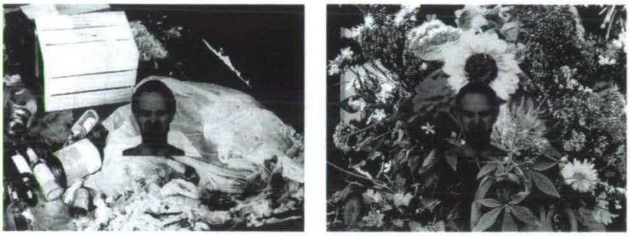


**Figure 3.1 Stimuli.** Example of a face-context stimulus. Facial expressions of fear, happiness and disgust were paired with context emotions of fear, happiness and disgust. These pairs could constitute congruent emotions (left), for instance a facial expression of disgust in a context of garbage, or incongruent emotions (right), the same expression shown among flowers.

The experiment was divided into three blocks. Each block contained two emotional categories for faces and contexts, resulting in three blocks constituting all pairwise combinations of fear, happy and disgust. Thus, a block could contain the emotions fear and disgust (with four conditions of face-context pairs: FF, DD, FD, DF), fear and happiness (FF, HH, FH, HF), and happiness and disgust (HH, DD, HD, DH).

The block order was counterbalanced across participants. Each condition contained 48 trials, two conditions being congruent, two being incongruent pairs of facial expression and context emotion. Each face stimulus served in both congruent and incongruent contexts. A face was never shown more than once in combination with one particular emotional scene.

*3.2.1.3 Analyses*

Differences across conditions were analyzed for error rates (percentage of incorrect responses) and response times (average for correct responded trials). Response time data were inspected on outliers for each participant separately. Response times that were more than 2.5 standard deviations (*SD*) from the mean of each condition were removed from analysis. Using these criteria, less than 1% of the trials were removed. For the percentage of errors and response times, the mean (*M*) and standard error of mean (*SEM*) will be reported.

The main and interaction effects for the factors facial expression and context emotion were analyzed for each experimental block separately (levels: fear and disgust; fear and happiness; happiness and disgust) in a repeated measures analysis of variance ($\alpha = 0.05$, one-tailed). *p*-values were corrected by Greenhouse-Geisser epsilon, if appropriate. For significant ANOVA effects, we report partial eta squared ($\eta_p^2$) as an estimate of effect size. For significant t-test effects we report Cohen's d. Cohen's d is defined as the difference between two means divided by the pooled standard deviations for those means (Cohen, 1988).

*3.2.2 Results*

The error rates were on average less than 5%. In the block FD, a significant main effect was found for facial expression, $F(1, 14) = 22.12$, $p < .001$, $\eta_p^2 = .61$, as more errors were made for fearful faces ($M = 5.2\%$, $SEM = 0.7\%$) as compared to disgusted faces ($M = 2.2\%$, $SEM = 0.5\%$). No significant effects were found for the blocks FH and HD.

The analyses for response times showed a main effect for facial expression and interaction effects between facial expression and context emotion in all blocks. None of the blocks showed a main effect for context emotion. A main effect was observed for facial expression in the block FD, $F(1,14) = 15.17$, $p < .001$, $\eta_p^2 = .52$, as it was found that response times were faster to disgusted facial expressions ($M = 762$ ms, $SEM = 26$ ms) than fearful facial expressions ($M = 860$ ms, $SEM = 36$ ms). The interaction-effect between facial expression and context emotion was significant, $F(1,14) = 7.39$, $p < .05$, $\eta_p^2 = .35$, and showed that response times were faster for fearful facial expressions in fearful contexts ($M = 838$ ms, $SEM = 36$ ms) compared to disgusted contexts ($M = 882$ ms, $SEM = 39$ ms), $t(14) = 2.01$, $p < 0.05$, $d = .30$. The difference between disgusted facial expressions in disgusted contexts ($M = 759$ ms, $SEM = 29$ ms) and disgusted facial expressions in fearful contexts did not attain significance ($M = 766$ ms, $SEM = 26$ ms), $t(14) = .35$, $p = .73$ (Figure 3.2).

The experimental block FH showed a main effect for facial expression, $F(1,14) = 11.76$, $p < .01$, $\eta_p^2 = .46$, reflecting faster response times to happy facial expressions ($M = 604$ ms, $SEM = 22$ ms) over fearful facial expressions ($M = 649$ ms, $SEM = 29$ ms). This effect was modified by an interaction with context emotion, $F(1,14) = 11.36$, $p < .01$, $\eta_p^2 = .45$, in that responses to happy facial expressions were faster if accompanied by happy contexts ($M = 590$ ms, $SEM = 22$ ms) compared to fearful contexts ($M = 617$ ms, $SEM = 23$ ms), $t(14) = 2.89$, $p < 0.01$, $d = .32$, while responses were not significantly faster for fearful facial expressions in fearful contexts ($M = 640$ ms, $SEM = 32$ ms) compared to happy contexts ($M = 658$ ms, $SEM = 28$ ms), $t(14) = 1.61$, $p = .13$.

The experimental block DH showed a main effect for facial expression, $F(1,14) = 5.66$, $p < .05$, $\eta_p^2 = .29$, as response times to disgust facial expressions ($M = 594$ ms, $SEM = 15$ ms) were faster than happy facial expressions ($M = 617$ ms, $SEM = 20$ ms). The interaction between facial expression and context emotion was significant, $F(1,14) = 11.01$, $p < .01$, $\eta_p^2 = .44$, showing that responses were faster to happy facial expressions if accompanied by happy contexts ($M = 600$ ms, $SEM = 17$ ms) compared to disgusted contexts ($M = 634$ ms, $SEM = 24$ ms), $t(14) = 2.66$, $p < .01$, $d = .42$, whereas the difference between facial expressions of disgust in a disgusted contexts ($M = 588$ ms, $SEM = 17$ ms) compared to a happy contexts ($M = 601$ ms, $SEM = 13$ ms) was not significant, $t(14) = 1.37$, $p = .09$.

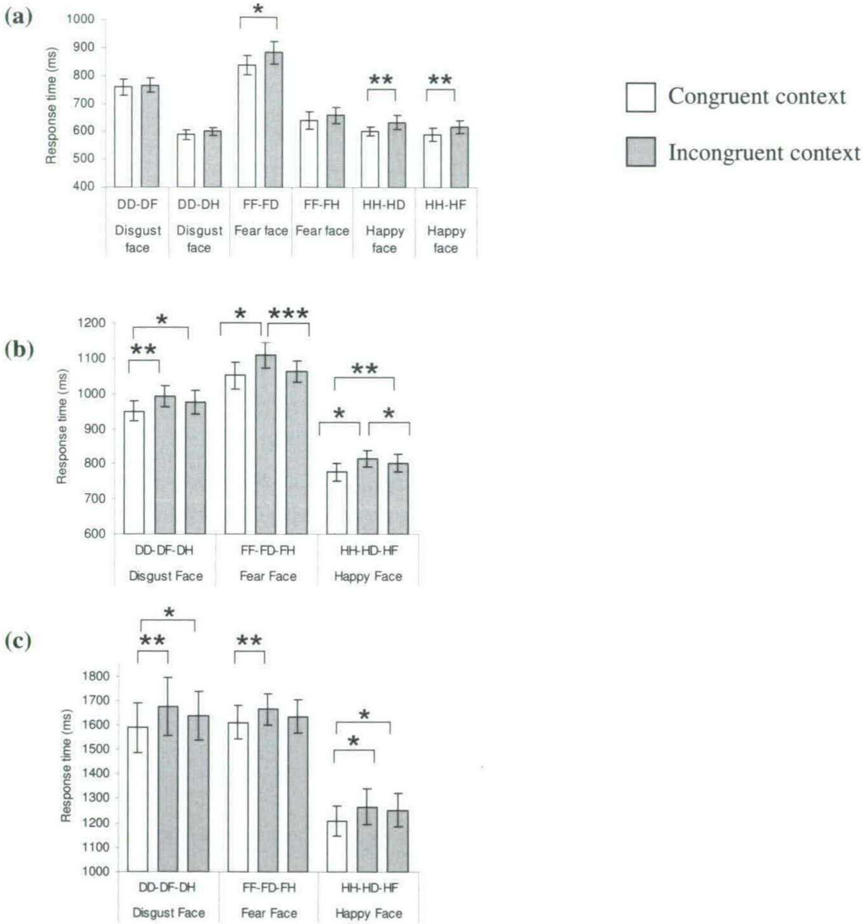



**Figure 3.2 Response times for facial expressions as a function of emotional scenes.** Mean response times for facial expressions in congruent (white bar) and incongruent contexts (gray bar). In experiment 1, a two-alternative choice task was used in which participants categorized facial expressions of fear, disgust and happiness (a). In experiment 2, a three-alternative choice task was used and presentation times were shortened to 200 ms (b), and in experiment 3 task load was increased due to a distracter task. Abbreviations DD, DF, DH; FF, FD, FH; HH, HD, HF: first letter is facial expression, second letter is context emotions. D=disgust; F=fear; H=happy. Error bars represent one standard error around the mean. Asterisks denote the significance level of a pairwise comparison: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

### *3.2.3 Discussion*

The present data show a clear difference for facial expressions as a function of context emotion, which is reflected in better and faster recognition of facial expressions of fear and happiness accompanied by contexts with a similar emotional content. However, this congruency effect is sensitive to the emotional content of the face-context pairs, as the effects were not significant for all pairwise combinations of face and context.

A number of reasons may explain why different effects were found across fear, disgusted and happy facial expressions. Firstly, the difference between congruent and incongruent face-context pairs may be specific for certain facial expressions. For instance, some facial expressions may be recognized clearly regardless of the accompanying context information, whereas others may be perceived as more ambiguous and depend on the context (Russell, 1997). Although the facial stimuli and context scenes were carefully selected for clarity of the emotion, in the next experiments we have added a classification task after the main experiment to verify whether participants could recognize the selected facial expressions. In this additional validation, facial expressions were shown without contexts, and contexts without faces (i.e., see for a discussion about source clarity (Carroll & Russell, 1996; Russell, 1997).

Secondly, the results may partly be explained by the response alternatives that were used. A drawback of the two-alternative choice design is that categorization may depend on the combinations that were tested in a certain block. In fact, fearful facial expressions paired with fearful contexts showed a higher error rate and a longer mean response time when shown in FD blocks than in FH blocks, a pattern which was also obtained for response times of facial expressions of disgust paired with disgusted context shown in FD blocks than in HD block (see Figure 3.2). In fact, similar effects for recognizing facial expressions have been observed in previous research (Tanaka-Matsumi et al., 1995). Therefore, in the next experiment we have used a three-alternative design, requiring participants to select between emotions of fear, disgust and happiness in the same block, which equalizes the combinations across blocks.

Thirdly, the relatively long presentation times that were used here may have increased the number of saccades to details in face and context. Global perceptual analyses of the scene may rely on relatively short presentation times. By using short presentation times, participants have only time for a single fixation (Loftus, Nelson, & Kallman, 1983). In the next experiment, stimuli were presented for 200 ms. A study by Calvo and Lang (2005) has shown that emotional scenes with short presentation times still receive preferential processing, even when positioned in the periphery. Therefore, our hypotheses were similar to experiment 1.

## 3.3 Experiment 2

In experiment 2, facial expressions of fear, disgust and happiness were presented in emotional congruent and incongruent context scenes. Participants were required to categorize facial expressions. Emotion categorization was investigated by using a three-alternative forced choice task. Presentation-times were shortened to 200 ms to investigate whether the effects found in experiment 1 persist.

### *3.3.1 Method*

#### *3.3.1.1 Participants*

Participants were 22 neurologically healthy undergraduate students (13 females; overall age $M = 20.0$ years) from Tilburg University. They participated for course credits. Three were left-handed. All had normal or corrected to normal vision and provided informed consent.

#### *3.3.1.2 Stimuli and Procedure*

The materials were similar to experiment 1. Facial expressions (F, D, H) were paired with context emotions (F, D, H). Stimuli were presented for 200 ms. Participants were required to categorize facial expressions and used three buttons indicating fearful, disgusted, and happy expressions. Response times were recorded from the onset of stimulus presentation. Response mapping was counterbalanced across participants. Participants were instructed to respond by their index finger. To control the distance to each of the response buttons, participants were required to rest their index finger during the intertrial interval on a keypad, which was monitored by a video-camera.

The experiment comprised eight separate blocks containing each 72 trials. The order of blocks was counterbalanced across subjects. Each block contained all possible emotion combinations of face and context resulting in congruent pairs (FF, DD, HH, each 96 trials) and incongruent pairs (FD, FH; DF, DH; HD, HF, each 48 trials).

In order to check whether the participants were capable to categorize each emotion correctly, after the experiment, participants had to classify 36 facial expressions and 36 emotion-inducing scenes (12 for each emotion). They were removed from analyses if they failed to classify more than eight out of twelve facial expressions or context emotions correctly for at least one of the emotions. Similar to the main experiment, participants made a decision among the three alternatives (fear, disgust, happiness). The stimuli were similar to the ones shown in the main experiment, but faces and contexts were now shown separately. Faces were shown on a gray background. The size of the stimuli was similar to the stimuli in the main experiment. The order of the two blocks (i.e., faces and scenes) was counterbalanced across subjects.

*3.3.1.3 Analyses*

Error rates and response times were analyzed for each condition. Response time data were inspected for outliers for each participant separately. Using similar criteria for outlier removal as experiment 1, on average 2.6% of all trials were removed.

Main and interaction effects for facial expression (disgusted, fearful, happy) and context emotion (disgusted, fearful, happy) were tested by a repeated measures analysis of variance ($\alpha = 0.05$, one-tailed).

*3.3.2 Results*

Two participants were removed from analyses because they failed to classify more than eight out of twelve facial expressions correctly. The remaining participants were all accurate in the classification task.

The error rates for judging facial expressions in contexts were on average 8.3%. A main effect was found for facial expression, $F(2, 38) = 29.45, p < .001, \eta_p^2 = .61$. Error rates differed significantly across facial expressions of fear ($M$ = 17.2%, $SEM$ = 2.4%), disgust ($M$ = 6.6%, $SEM$ = 0.9%) and happiness ($M$ = 1.0%, $SEM$ = 0.5%), $t(19) = 3.82, p < .01, d$ = 1.29, for fear-disgust, $t(19) = 5.61, p < .001; d = 1.65$, for disgust-happiness, and $t(19) = 7.20$, $p < .001, d = 2.06$ for fear-happiness. A main effect was also found for context emotion, $F(1.67; 31.68) = 7.38, p < .01, \eta_p^2 = .28$, as more errors were made for faces in disgusted contexts ($M$ = 8.9%, $SEM$ = 1.0%) and for faces in fearful contexts ($M$ = 8.7%, $SEM$ = 0.9%) than faces in happy contexts ($M$ = 7.2%, $SEM$ = 0.9%), $t(19) = 4.30, p < .001, d = .37$ for disgust-happiness, $t(19) = 2.54, p < .05, d = .42$ for fear-happiness.

The main effects were modified by an interaction effect, $F(4,76) = 10.78, p < .001, \eta_p^2 = .36$. Less errors were committed for disgusted facial expressions accompanied by disgusted contexts ($M$ = 4.6%, $SEM$ = 0.8%) and happy contexts ($M$ = 5.2%, $SEM$ = 1.0%) than fearful context ($M$ = 9.9 %, $SEM$ = 1.7%), $t(19) = 3.06, p < .01, d = .89$ and $t(19) = 3.29, p < .01, d = .74$ respectively. Further, less errors were made for fearful facial expressions in fearful contexts ($M$ = 14.7%, $SEM$ = 2.4%) and happy contexts ($M$ = 15.8%, $SEM$ = 2.6%) compared to disgusted contexts ($M$ = 20.9%, $SEM$ = 2.6%), $t(19) = 4.03, p < .001, d = .77$ and $t(19) = 4.38, p < .001, d = .44$ respectively. Less errors were committed if happy facial expressions were presented in happy contexts ($M$ = 0.4 %, $SEM$ = 0.9%) compared to fearful contexts ($M$ = 1.5%, $SEM$ = 0.6%), $t(19) = 1.99, p < .05$, d = .49.

The analysis of response times showed a main effect for facial expression, $F(2,38) = 61.56, p < .001, \eta_p^2 = .76$, and context emotion, $F(2,38) = 3.78, p < .05, \eta_p^2 = .17$. Responses differed significantly across all facial expressions, being fastest for happy facial expressions

($M = 803$ ms, $SEM = 22$ ms), intermediate for disgusted facial expressions ($M = 953$ ms, $SEM = 26$ ms) and slowest for fearful facial expressions ($M = 1{,}061$ ms, $SEM = 35$ ms), $t(19) = 7.14$, $p < 0.001$, $d = 1.40$ for happy-disgust, $t(19) = 9.95$, $p < 0.001$, $d = .97$ for happy-fear, $t(19) = 4.72$, $p < 0.001$, $d = .79$ for disgust-fear. Responses were faster for faces in happy contexts ($M = 932$ ms, $SEM = 24$ ms) than faces in disgusted contexts ($M = 947$ ms, $SEM = 25$ ms), $t(19) = 3.32$, $p < 0.01$, $d = .14$, but there were no significant differences with faces in fearful contexts ($M = 939$ ms, $SEM = 25$ ms), $t(19) = 1.58$, $p = .13$ for disgust-fear, $t(19) = 1.03$, $p = .32$ for happiness-fear.

The main effects were qualified by a significant interaction, $F(2.28, 43.39) = 7.47$, $p < .01$, $\eta_p^2 = .28$. Comparisons between contexts showed that responses to facial expressions of disgust were faster if accompanied by a disgusted context ($M = 929$ ms, $SEM = 26$ ms) compared to happy context ($M = 958$ ms, $SEM = 28$ ms) and fearful context ($M = 972$ ms, $SEM = 26$ ms), $t(19) = 2.13$, $p < .05$, $d = .23$ for DD-DH, and $t(19) = 3.58$, $p < .01$, $d = .37$ for DD-DF. For fearful facial expressions, responses were faster if accompanied by fearful contexts ($M = 1{,}037$ ms, $SEM = 38$ ms) and happy contexts ($M = 1{,}053$ ms, $SEM = 35$ ms) as compared to disgusted contexts ($M = 1{,}094$ ms, $SEM = 35$ ms), $t(19) = 5.30$, $p < .001$, $d = .35$ for FF-FD and, $t(19) = 2.12$, $p < .05$, $d = .26$ for FH-FD. Responses for happy facial expressions were fastest if accompanied by happy contexts ($M = 784$ ms, $SEM = 23$ ms) as compared to disgusted contexts ($M = 818$ ms, $SEM = 23$ ms) and fearful contexts ($M = 807$ ms, $SEM = 21$ ms), $t(19) = 5.58$, $p < .01$, d = .33 for HH-HD, $t(19) = 3.30$, $p < .001$, d = .22 for HH-HF, and $t(19) = 1.96$, $p < .05$, d = .11 for HF-HD.

### *3.3.3 Discussion*

The results of experiment 2 are similar to those of experiment 1 in showing better and faster recognition of facial expressions that were accompanied by congruent emotional contexts, even when exposure duration only allowed single fixations (Loftus et al., 1983). Task load was also increased as participants chose among facial expressions of fear, disgust and happiness in a three-alternative forced choice design. Although these task requirements may have resulted in slightly increased error rates and response times, response advantages were, in contrast to experiment 1, also found for facial expressions of disgust in congruent contexts.

In previous studies it has been shown that the task load determines to what degree context information is processed. (Lavie, 1995) has shown that irrelevant peripheral distracters interfere with target processing only if the task load is relatively low. If the task load was high enough, distracters did not interfere anymore, which may relate to the capacity limits consumed by the target task.

Similar effects of interference from distracters in context have been observed for emotional scenes. Task irrelevant emotional information slows down response times for target processing (Algom et al., 2004; Erthal et al., 2005). But if the task load was increased by making the central task harder (e.g., by decreasing angular difference of bars in a bar-orientation task), unpleasant scenes did not slow down responses anymore (Erthal et al., 2005).

In the third experiment, a concurrent task was used to investigate how the imposed task load relates to the capacity for context processing. Besides the main task of categorizing the facial expression, participants were required to categorize a character that was overlaid on the face stimulus centrally. If the concurrent task reduces the attentional resources available for context processing, it is expected that the effects of contexts will disappear. If facial expression and context emotion show an interaction irrespective of the amount of attention dedicated, response times may nevertheless be delayed as a consequence of the additional task, but response times may still be faster for congruent pairs of facial expression and context emotion than incongruent pairs.

## 3.4 Experiment 3

In experiment 3, facial expressions of fear, disgust and happiness were presented in emotional congruent and incongruent context scenes. Participants were required to categorize facial expressions. An additional task was introduced to investigate whether attentional competition influences the observed effects in experiment 2.

### *3.4.1 Method*

The experiment was similar to experiment 2, except that the additional task now required participants to categorize a character symbol displayed in the center of the face.

#### *3.4.1.1 Participants*

Sixteen neurologically healthy undergraduate students (8 females, overall age $M = 23$ years) were recruited from Tilburg University. Two participants were left-handed. All had normal or corrected to normal vision and provided informed consent.

#### *3.4.1.2 Stimuli and procedure*

Similar face-context pairs were used, except that a “x” or “o” character was centered upon each face. Characters were presented in a small font size (10 points) or a large font size (20 points) to avoid that features related to the font size would cue the correct response. The overlap of the character upon the face-context stimulus was randomized to avoid a predictive relationship between the character and face-context stimuli. In addition, the combinations of face- and character stimuli were counterbalanced across participants. Stimuli had a similar onset and were presented for 200 ms. Participants were instructed to categorize the facial expression and use three buttons for disgust, fear and happiness. After this decision, participants made a two-key response as to indicate what character was presented. Response-mapping was counterbalanced across participants.

*3.4.1.3 Analyses*

The data were analyzed in the same way as in experiment 2. Using similar criteria for outlier removal as in experiment 1, 2.4% of all trials were removed on average.

*3.4.2 Results*

Two participants were removed from analyses, one because of an excessive number of misclassifications for fearful faces, the other because of poor performance on the categorization task for characters (*M* = 52% error). The remaining participants showed a relatively small percentage of errors (*M* = 8.9%, *SEM* = 1.3%).

The error rates for categorizing facial expressions in contexts were on average 6.6%. A significant main effect was found for facial expression, $F(1{,}54; 20{,}07) = 6.14$, $p < .01$, $\eta_p^2 = .32$, in that more errors were made for fearful facial expressions (*M* = 10.4%, *SEM* = 2.7%) and disgusted facial expressions (*M* = 8.3%, *SEM* = 2.3%) than happy facial expressions (*M* = 1.0%, *SEM* = 0.4%) $t(13) = 3.61$, $p < .01$, $d = 1.31$ for fear-happiness, and $t(13) = 3.28$, $p < .01$, $d = 1.18$ for disgust-happiness. The main effect for context emotion was not significant. However, a significant interaction was found between facial expression and context emotion, $F(1.70; 22.10) = 3.86$, $p < .05$, $\eta_p^2 = .23$. Less errors were committed for disgusted facial expressions in disgusted contexts (*M* = 6.5%, *SEM* = 1.9%) compared to fearful contexts (*M* = 10.3%, *SEM* = 2.8%) and happy contexts (*M* = 8.0%, *SEM* = 2.2%), $t(13) = 3.06$, $p < .01$, $d = .42$ for DD-DF, $t(13) = 2.62$, $p < .05$, $d = .20$ for DD-DH, and $t(13) = 2.69$, $p < .01$, $d = .24$ for DF-DH. No significant different context effects were found for facial expressions of fear and happiness ($p > .05$).

Analyses for response times showed a main effect for facial expression, $F(2{,}26) = 20.63$, $p < .001$, $\eta_p^2 = .61$. Responses were fastest for happy facial expressions (*M* = 1,242 ms, *SEM* = 67 ms) as compared to disgusted facial expressions (*M* = 1,634 ms, *SEM* = 106 ms) and fearful facial expression (*M* = 1,637 ms, *SEM* = 65 ms), $t(13) = 4.51$, $p < .001$, $d = 1.18$, and $t(13) = 6.63$, $p < .001$, $d = 1.60$ respectively. The main effect for context emotion was not significant ($p > .05$).

The interaction between facial expression and context emotion was significant, $F(2.46; 32.01) = 4.74$, $p < .01$, $\eta_p^2 = .27$ (Figure 3.2). Responses were faster to disgusted facial expressions in disgusted contexts (*M* = 1,589 ms, *SEM* = 101 ms) compared to fearful contexts (*M* = 1,676 ms, *SEM* = 119 ms) and happy contexts (*M* = 1,637 ms, *SEM* = 100 ms),

$t(13) = 2.73$, $p < .01$, $d = .21$ for DD-DF, and $t(13) = 2.49$, $p < .05$, $d = .13$ for DD-DH. Responses were also faster to fearful facial expressions in fearful contexts ($M$ = 1,611 ms, $SEM$ = 68 ms) than disgusted contexts ($M$ = 1,664 ms, $SEM$ = 63 ms), $t(13) = 2.82$, $p < .01$, $d = .22$). For happy faces, responses were faster if presented in a happy context ($M$ = 1,208 ms, $SEM$ = 64 ms) compared to a disgusted context ($M$ = 1,266 ms, $SEM$ = 72 ms) and fearful context ($M$ = 1,253 ms, $SEM$ = 68 ms), $t(13) = 1.95$, $p < .05$, $d = .18$ for HH-HD, and $t(13) = 1.89$, $p < .05$, $d = .23$ for HH-HF.

### *3.4.3 Discussion*

The general increase in response times compared to the previous experiments indicates that the concurrent task competed for attentional resources. However, the results of experiment 2 were largely replicated and they clearly indicate that this increased load did not override the effects of the context. The pattern of accuracies was similar to experiment 2 in that disgusted facial expressions were better recognized in a disgust-inducing context. Response times were fastest for facial expressions that were accompanied by congruent contexts. Note that response times were not significantly faster for fearful facial expressions in fearful contexts compared to happy contexts. This pattern was also observed in the other experiments.

## 3.5 General discussion

The present study shows that recognition of facial expression is significantly affected by the accompanying but task-irrelevant emotional context shown in the background. In experiment 1, facial expressions of fear were recognized faster in fearful contexts compared to disgusted contexts, and happy facial expressions were faster recognized in happy contexts than fearful and disgusted contexts. Experiment 2 used shorter presentation times and experiment 3 used a concurrent character categorization task. In addition to the effects observed in experiment 1, faster responses were found for disgusted facial expressions in disgusted contexts compared to these expressions in fearful and happy contexts in experiment 2 and 3. Taken together, our results demonstrate that the effects of context emotion on facial expression processing are robust across different task conditions and presentation times.

The effects of context on face perception may depend on global analysis of scene information, whereby the emotional gist of the scene is extracted. Given the short presentation times, it is rather unlikely that detailed analysis of scenes has finished (see also Calvo & Lang, 2005). This proposal is consistent with results obtained by an event-related potential (ERP) study, in which participants were presented with facial expressions of fearful and neutral faces in fearful or neutral contexts. It was shown that the N170 amplitude was increased for faces in fearful contexts, especially for fearful facial expressions (See chapter 2). The N170 occurs in an early stage of face processing (~160–200 ms across participants), and the observed differences suggest that the emotional information from face and scene is extracted rapidly. In a model by Bar (2004), it has been proposed that scene analyses may rely on the gist of the scene, in which the semantic meaning of the scene is extracted by fast analyses of the global properties of the scene. The results of experiment 2 and 3 using short presentation times suggest a key role for perceiving the emotional gist of the scene, which may result in faster recognition of facial expressions.

A central question in the literature about emotion perception is whether negative emotions have privileged access to processing resources (Calvo & Lang, 2005; Öhman, Flykt et al., 2001; but see also Tipples et al., 2002), possibly by a subcortical fast route (LeDoux, 1996). This privileged access may divert attention from target processing. For example, in a lexical-decision task (Algom et al., 2004) and a bar-orientation task (Erthal et al., 2005) it was shown that response times were slowed down by the emotional value of the word and scene respectively. Importantly, we did not observe a plain slow-down for faces in negative scene contexts, but an interaction between facial expression and context emotion. The face-context pairs differ from the word reading and the bar-orientation task in that the dimensions of facial expression and context overlap on emotion. The dimensions are associated with the same task appropriate response, for which response times may differ between congruent and incongruent pairs of faces and contexts (see discussion about the emotional Stroop task by Algom et al., 2004). The present data show an interaction between facial expression and context emotion, and suggest that the underlying mechanism is different from the aforementioned studies that found a general slow-down for non-face targets in contexts. This is further supported by the finding that the interaction between facial expression and context emotion was maintained under increased task load, which contrasts to the results obtained by Erthal et al. (2005).

The observed interaction effect transcends the positive-negative valence dimension (e.g., positive facial expressions processed faster in positive contexts than in negative contexts and vice versa). The differentiation within the negative emotion category showed that categorization of facial expressions of disgust was faster in disgusted contexts than fearful contexts, as well as the opposite, that recognition of fearful facial expressions was faster in fearful contexts than disgusted contexts. One purpose of fast and efficient processing of these specific facial expressions is to react adaptively and rapidly to the environment (Ekman, 1992). Happy emotions are social rewarding and may evoke affiliation (Keltner & Kring, 1998). Perception of fear is crucial to monitor for potentially dangerous situations (LeDoux, 1996). The recognition of a fearful expression in a certain context may direct attention to objects in the environment that were the triggers of these expressions (Frijda & Tcherkassof, 1997). The evocation of disgust keeps the individual away from a harmful object or contamination (Rozin & Fallon, 1987). There may even though be individual (Fox et al., 2001; Keltner & Kring, 1998) or cultural differences (Ekman, 1992; Mesquita & Frijda, 1992) in what events call forth an emotion.

The dissociation between fear and disgust processing found in the present study adds to studies that have suggested relatively separate neural systems for these emotions (see review Calder, Lawrence, & Young, 2001). Brain-imaging studies have shown that the amygdala responds mainly to facial expressions of fear, and that the insula-basal ganglia system responds mainly to expressions of disgust (but see also Fitzgerald, Angstadt, Jelsone, Nathan, & Phan, 2006; Phillips et al., 1997). Neuropsychological case studies have confirmed that these brain regions are essential for recognizing fear and disgust respectively (Adolphs, Tranel, Damasio, & Damasio, 1994; Sprengelmeyer et al., 1996). It would be interesting to study whether the same brain regions are involved in processing faces in contexts, and whether the response in the respective brain regions of interest are modified by congruency between face and context.

The present results are potentially relevant for other areas of context research as well. Previous studies have found that recognition of facial expressions can be influenced by emotional contexts, as has been shown in the context of emotional voices (de Gelder & Vroomen, 2000) and bodily expressions (Meeren et al., 2005; Van den Stock, Righart, & de Gelder, 2006). These studies have repeatedly shown that individuals are better and faster in identifying facial expressions that were accompanied by congruent information from body and voice as compared to incongruent information (de Gelder et al., 2006). In addition, previous work has used contexts that preceded the presentation of facial expressions. In some studies, context consisted of a storyline told by the experimenter. It was found that the emotional content of the storyline could affect how individuals interpreted the subsequently shown facial expressions (Carroll & Russell, 1996; Fernández-Dols & Carroll, 1997; Russell, 1997). However, the aforementioned studies differ in some important respects from the present work (e.g., different onset times of the face and context stimuli). A conclusive answer on the question whether similar perceptual processes are entailed by these different kinds of contexts should therefore be postponed.

The primary goal of the present study was to investigate how emotional scenes influence the recognition of facial expressions. A number of questions therefore remain unanswered as yet. We selected contrasting emotions (positive/happy vs. negative/fear or disgust) in order to introduce strong polarity between the scenes, because the relative distance between emotional values of the stimuli may affect how they are categorized (Tanaka-Matsumi et al., 1995). Isolated facial expressions or neutral scenes were not used here (but see chapter 2). A problem of selecting a baseline context is that it is unclear whether neutral contexts are truly non-emotional. Therefore, conclusions whether the processing of facial expressions was facilitated by congruent contexts, or whether the processing of facial expressions was inhibited by incongruent contexts, cannot be drawn.

It is worth noting that the physical properties of the scene stimuli are unlikely to explain the observed interaction between facial expression and context emotion. First, and most importantly, each context served as a background for facial expressions of fear, disgust and happiness, and therefore is congruent with some facial expressions, and incongruent with others. Second, the main effect found for facial expressions has been observed repeatedly in the behavioral literature and there is strong evidence that this does not depend on low-level features (Kirita & Endo, 1995). A similar line of reasoning holds for scenes. De Houwer and Hermans (1994) used line-drawings and found similar congruency effects, in their case for naming the affective value of a word (negative, positive) that was overlaid on congruent line-drawings as compared to incongruent ones. This suggests, as shown by the line-drawings, that color is not, at least not entirely, explaining this interaction between facial expression and context emotion (for a similar reasoning about perceptual complexity, see Blanchette, 2006).

This is not to say that the physical properties of face and context are unimportant. Colors (Spence, Wong, Rusan, & Rastegar, 2006) and spatial frequencies (Bar, 2004; Schyns & Oliva, 1994) play a key role in the early stages of visual encoding of scenes and may be an integral property of emotion perception.

*Conclusions*

The present study shows that the categorization of facial expressions differs depending on whether it figures in the context of an emotional congruent or incongruent scene. This effect was not confined to a simple negative-positive dichotomy, in that positive facial expressions (i.e., happy) were better and faster recognized in positive contexts than in negative contexts (i.e. both fear and disgust) and vice versa, but specific effects were obtained for each facial expression (i.e. fearful facial expressions were faster recognized in fearful than in disgusted contexts and vice versa). The effects were also obtained for short presentation times allowing only a single fixation, and resisted the negative effects of a concurrent task that competes for attentional resources. Taken together, our study clearly indicated that the gist of the emotional scene plays an important role in rapidly recognizing facial expressions.

Chapter 4

# The recognition of facial expressions is influenced by emotional scenes An ERP study

*Abstract*

The context of faces importantly affects how we interpret facial expressions of emotion. Event-related potential (ERP) studies have recently shown that early stages of face perception are affected by the accompanying context. The aim of the present study was to examine how these early stages are affected when individuals categorize facial expressions (fear, happy) that are accompanied by congruent, incongruent and neutral contexts. A general effect of context information was found, in that the N170/VPP amplitudes were smaller when faces were accompanied by meaningful scenes as compared to meaningless scenes (i.e., scrambled images). A specific emotion effect was found for the N170, as N170 amplitudes were larger for faces in fearful scenes as compared to faces in happy and faces in neutral scenes. Planned comparisons showed that N170 amplitudes were significantly increased for fearful faces in fearful scenes as compared to fearful faces in happy scenes. These findings suggest that the information that is derived from the context scene is combined with facial expressions on an early stage of encoding.

This chapter will be submitted as: Righart, R. & de Gelder, B. (2007). Recognition of facial expressions in contexts. An ERP study.

## 4.1 Introduction

For several decades, it has been examined how context information affects perception (Bar, 2004; Davenport & Potter, 2004). A number of studies have focused on the semantic consistency between objects and contexts, and found that objects are better identified when they are perceived in a congruent than incongruent context (Davenport & Potter, 2004). For instance, a chair is often perceived in the context of a living room, and the perception of this context may therefore set the expectation for this object. A similar kind of expectation may be valid for perceiving facial expressions in emotional salient contexts (see Kim et al., 2004; Mobbs et al., 2006).

Facial expressions can be rather ambiguous if seen in isolation from the evoking situation. But these expressions may become clear if the accompanying context is known. A recent study has shown that the human brain responds differently to facial expressions of emotion if knowledge is acquired about the relevant context information (Kim et al., 2004). Individuals may associate the expression with the context that possibly evoked this expression. Similar to this study, both unimodal and crossmodal studies have found that emotion-inducing contexts, for instance vocal and bodily expressions, affect how individuals perceive a facial expression (de Gelder et al., 2006).

Behavioral studies have shown that individuals evaluate a facial expression as more angry when they had heard a negative story in advance (e.g., a rejection in a restaurant) as compared to the same facial expression shown after a positive story (Carroll & Russell, 1996). In addition, recently it has been shown that individuals categorize a facial expression faster when seen in a congruent visual context (e.g., faster recognition of a face conveying disgust in front of a garbage area) than when these same expressions are seen in incongruent contexts (chapter 3). These behavioral data are however inconclusive as to what stage of processing is affected by contexts.

Event-related potential studies (ERPs) provide insight into the time-course of information processing. Different stages of face processing have been related to distinct ERP components and ERPs could therefore be of interest to study context effects. Most researchers have focussed their attention on the N170 component (Bentin et al., 1996; George, Evans, Fiori, Davidoff, & Renault, 1996; Itier & Taylor, 2004b) and some have also examined the VPP component (Bötzel & Grusser, 1989; Cauquil et al., 2000; George et al., 1996; Itier & Taylor, 2004b). The N170 has its peak at around 170 ms after stimulus onset and a maximal negative peak is observed on occipito-temporal sites. The VPP has its maximum positivity on central electrode sites and occurs at about the same latency as the N170. Both components are larger to faces as compared to other objects (Itier & Taylor, 2004b) and may reflect similar neural generators (Joyce & Rossion, 2005).

Much less emphasis has been put on the P1, a positive deflection occurring at occipital sites at about 100 ms after stimulus onset. Most studies have focused on the response properties of the P1 component as related to spatial attention and physical features, for example colour, motion and shape (Hillyard & Anllo-Vento, 1998). However, recent studies have found that the P1 is larger to faces than non-face objects (Herrmann, Ehlis, Ellgring et al., 2005; Itier & Taylor, 2004b), and that facial expressions affect the P1 as well (Batty & Taylor, 2003; Eger et al., 2003).

Most studies have used faces that were isolated from context information. Therefore, a relatively small number of studies have examined how contexts contribute to face perception. ERPs have been used to investigate the effects of neutral (Ganis & Kutas, 2003) and emotional scenes (Smith, Dolan et al., 2004) figuring in the context of objects. These studies have found long-latency ERP effects for perceiving objects in semantically congruent scenes (Ganis & Kutas, 2003) and recognition memory for objects that were studied in emotional salient scenes (Smith, Dolan et al., 2004). In contrast to objects, for face processing ERP studies suggest that early stages of encoding are affected by the presence of meaningful scenes. Although a clear N170 component is still found for faces when they are accompanied by a natural scene contexts (G. A. Rousselet, M. J. Mace, & M. Fabre-Thorpe, 2004a), the N170 amplitude is decreased for faces that are shown in the context of meaningful scenes as compared to faces isolated from scene information (chapter 2, and see also in recognition memory test Galli, Feurra, & Viggiano, 2006).

The P1 and N170 evoked by faces may relate to early stages of encoding (Itier & Taylor, 2004b). An important question is how context information in general and emotional context information in specific affect these stages of face processing. It has been found that bodily expressions that are presented in the context of facial expression affect the P1 amplitude, with the P1 amplitude being increased for incongruent combinations of facial expression and bodily expression (fear-anger or anger-fear) as compared to congruent combinations (fear-fear or anger-anger) (Meeren et al., 2005). In addition, it has been shown that emotional scenes affect the P1 amplitude, as P1 amplitudes were increased to fearful scenes as compared to neutral scenes (see Carretie et al., 2004 for scenes without a face; Smith et al., 2003). However, the P1 for faces presented in the context of emotional scenes is not affected differently than when these emotional contexts were presented without a face (chapter 2).

In contrast, for the N170 component different results have been obtained. The N170 was larger to faces in fear-inducing context scenes, especially if fearful faces were presented (chapter 2). This finding is consistent with earlier studies that have observed that the emotional content of socially relevant stimuli affects the N170. Although some studies had observed that the N170 is insensitive to facial expressions (Eimer & Holmes, 2002; Holmes et al., 2003), a number of recent studies have shown that facial expressions of emotion (Batty & Taylor, 2003; Caharel, Courtay, Bernard, Lalonde, & Rebai, 2005; Righart & de Gelder, 2006; Stekelenburg & de Gelder, 2004; Williams, Palmer, Liddell, Song, & Gordon, 2006) and social affective features (Pizzagalli et al., 2002) modify the N170 amplitude, and that the N170 amplitude to facial expressions interacts with the emotional scene that is shown in the context of faces (Righart & de Gelder, 2006).

In the present study, happy and fearful facial expressions were used and were overlaid on happy, fearful or neutral scene contexts. We used these opposite expressions to have clear congruent and incongruent face-context compounds. Participants were instructed to categorize facial expressions of fear and happiness. These instructions differ from our previous ERP study, in which an orientation-decision task was used (chapter 2). Our main question was whether the N170 response to faces in contexts will be affected by this task manipulation. Previous work by Caharel et al. (2005) predicts that the N170 will not change by these manipulations, because they found that the N170 was affected by facial expressions both in task conditions that required participants to categorize facial expressions, and conditions that required participants to make decisions on face familiarity.

We hypothesized that the ERP amplitudes will increase for the N170 component to fearful faces in fear-inducing contexts compared to happy and neutral contexts, but not for the P1 component (chapter 2). The colour properties were controlled for by using scrambled versions of the context stimuli. No differences are expected as a function of these scrambled contexts. Further, it was predicted that the N170 amplitude will be increased for faces in scrambled contexts compared to intact contexts, because attentional competition between face and context will be decreased for faces in scrambled contexts (Galli et al., 2006).

## 4.2 Method

### *4.2.1 Participants*

Eighteen participants (12 males) ranging from 21 to 50 years participated in the experiment (13 right handed). One participant was removed as it turned out that she had had visual problems (glaucoma). The remaining participants had normal or corrected-to-normal vision. None of them reported a history of neurological or psychiatric diseases and all had given informed consent. The study was performed in accordance with the ethical standards of the Declaration of Helsinki.

### *4.2.2 Stimuli*

Stimuli were faces centrally embedded on context pictures (Figure 4.1). Face stimuli were colour pictures of caucasian faces from the Karolinska Directed Emotional Faces set (Lundqvist et al., 1998). Facial identities were rated by a separate group of participants in order to select facial identities that expressed both fear and happy emotions validly. 24 male and 24 females conveying facial expressions of fear and happiness were selected from the set resulting in a total of 96 faces.

Contexts were colour pictures selected from the IAPS (Lang et al., 1999) and from the web. Contents were validated for emotions of fear, happiness or neutrality by a different group of participants (N=10) on arousal (7-point scale, from 1=unaffected to 7 extremely aroused) and emotion category using a collection of 401 picture stimuli. From this collection, stimuli were chosen and prepared for the experiment. In total, 24 scene stimuli were selected for each emotion category (Figure 4.1). For the final set that was used in the experiment, arousal rates for both fearful (4.84) and happy contexts (4.24) were significantly higher than neutral contexts (2.00) (both $p < 0.001$ but $p > 0.05$ for fearful-happy). On average, the intended label for happy contexts was chosen for 75% of the trials, for fearful contexts on 64% of the trials, and for neutral contexts on 87% of the trials. Contexts were as much as possible selected for being scenes conveying a situation instead of objects.

In order to control for low-level features (e.g., color), all scene pictures were scrambled by randomizing the position of pixels across the image (blocks of 8 x 8 pixels were randomized across the image measuring 768 x 572 pixels width and height respectively), which renders the pictures meaningless. The resulting pictures were inspected carefully in order to avoid that certain features would cue recognition.

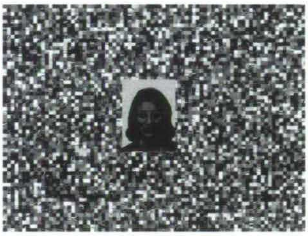
The height and width of the facial images was 7.9 cm x 5.9 cm (5.6° x 4.2°) and for context images 24.5 cm x 32 cm respectively (17.4° x 22.6°). Faces did not occlude the critical information in the context picture.

### *4.2.3 Procedure*

Participants were seated in dimly illuminated and a electrically shielded cabin with a monitor positioned at 80 cm distance. Participants were instructed and familiarized to the experiment by practice sessions.

The experiment was run in 16 blocks each containing 72 trials of face-context compounds, 8 blocks of faces with intact context scenes and 8 blocks of faces with scrambled scenes. Intact and scrambled blocks alternated (with order randomized across participants).

a) b)



c)

| Fearful (4.84) | Happy (4.25) | Neutral (2.00) |
|---|---|---|
| ○ Carcrash (4.08) | ○ Beach (4.65) | ○ Highway (2.60) |
| ○ Fire (5.45) | ○ Candles (4.98) | ○ House (1.63) |
| ○ Flood (4.78) | ○ Firework (4.45) | ○ Mill (2.00) |
| ○ Planecrash (5.38) | ○ Palmtrees (4.28) | ○ Room (1.90) |
| ○ Storm (4.43) | ○ Party (3.53) | ○ Street (2.23) |
| ○ Tornado (4.93) | ○ Swimmingpool (3.60) | ○ Train (1.65) |

**Figure 4.1 Stimuli.** Facial expressions of fear and happiness were combined with context scenes conveying fear, happiness or a neutral situation. (a). An example of a face-context compound showing a facial expression of fear and a carcrash and (b) the same facial expression shown in a scrambled version of the carcrash that was used as a control stimulus for the effects of color. (c). Six context categories of stimuli were selected, each category containing four different stimuli (i.e., in total 24 for each emotion) that were validated. Average arousal ratings of validation are shown in parentheses. Participants rated the stimuli on a scale from 1-7.

Each block took about three minutes. Facial expressions (fear, happy) were paired to a scene of each context emotion category (fear, happy, neutral) in order to balance the design factorially. Altogether, intact and scrambled blocks amounted to 12 conditions each containing 96 trials. After two blocks, identity and context pairs were changed.

The face-context compounds were presented for 200 ms and were prompted by a fixation cross. Participants performed a two-alternative forced choice task in which they categorized whether the facial expression was happy or fearful. Responses were recorded from stimulus-onset. No feedback was given. They were instructed to respond as accurately and fast as possible. Response-buttons were counterbalanced across participants. The intertrial interval was randomized between 1200 and 1600 ms.

After the main experiment, participants received a validation task in which they were required to judge the arousal of the scenes and to categorize scenes on emotion. In contrast to the stimulus validation, arousal was measured on a five-point scale now because of response box limitations in the ERP lab (five-point scale from 1=calm to 5=extremely arousing). Emotional categorization was measured by using three options (fear, happy, neutral). Arousal and valence were measured in separate blocks. The order of the blocks was counterbalanced across participants.

*4.2.4 EEG recording*

EEG was recorded from 49 electrode locations using active Ag-AgCl electrodes (BioSemi Active2) mounted in an elastic cap referenced to an additional active electrode (Common Mode Sense). EEG was bandpass filtered (0.1-30 Hz, 24 dB/Octave). The sampling rate was 512 Hz. All electrodes were offline referenced to an average reference. Horizontal electrooculographies (hEOG) and vertical electrooculographies (vEOG) were recorded. The raw data were segmented into epochs starting 200 ms before stimulus onset to 1000 ms after stimulus onset. The data were baseline corrected to the first 200 ms.

EEG was EOG corrected by using the algorithm of Gratton, Coles, & Donchin (1983). Epochs exceeding 100 μV amplitude difference at any channel were removed from analyses. No differences were observed across conditions for facial expressions and context emotions. On average 87.5 trials (range across conditions: 86.8 - 88.5) were left for faces in intact contexts and 85.1 trials (84.4 - 85.4) were left for faces in scrambled contexts after removal of the artifacts. After removal of trials containing inaccurate responses or responses below 200 ms, ERPs were averaged for conditions of facial expressions of fear in intact fearful, happy and neutral contexts, and for happy facial expressions in the same context categories. Similarly, averages were computed for scrambled blocks, resulting in a total of 12 conditions.

Electrode selection for P1 and N170/VPP analyses was based on previous studies. Based on grand average ERP inspection, peak detection windows for P1 and N170/VPP were measured using time-windows of 60-140 ms and 100-220 ms respectively. Peak latencies and amplitudes of P1 were analyzed at occipital sites (O1/2) and occipito-temporal sites (PO3/4, PO7/8) as the maximal positive peak amplitude. The N170 was analyzed on occipito-temporal sites (P5/6, P7/8 and PO7/8) as the maximal negative peak amplitude. The VPP was analyzed on midline sites (FCz, Cz and CPz) as the maximal positive amplitude.

*4.2.5 Data analyses*

Behavioral analyses were performed for error rate (percentage of incorrect responses) and response times (average for correct responded trials) for differences across conditions. Response-time data were inspected on outliers for each participant separately. Response times that were more than 2.5 SD from the mean of each condition were removed from analyses. Using these criteria, 2.6% of the trials were removed. Main- and interaction effects were analyzed by using multivariate repeated measures ANOVA containing the factors Image (intact, scrambled), Facial expression (fear, happy) and Context emotion (fear, happy, neutral). Planned comparisons were performed (Howell, 2002) to test our specific hypothesis for faces in context, that is, that fearful faces are faster recognized in fearful contexts (congruent) as compared to happy contexts (incongruent), and that happy faces are faster recognized in happy contexts (congruent) as compared to fearful contexts (incongruent). ($\alpha$ = .05, one-tailed test because of directional hypotheses). Statistics are indicated with original degrees of freedom (Picton et al., 2000).

P1 and N170/VPP latencies and amplitudes for faces in contexts were analyzed by using repeated measures ANOVA containing the within subject factors Image (intact, scrambled), Facial expression (fear, happy), Context emotion (fear, happy), Hemisphere (left, right) and Electrode position. Because of the specific hypothesis, planned comparisons were performed here as well for the same comparisons as in the behavioral analysis. Scalp topographic analysis and differences across the topographies for emotional scenes were analyzed by using repeated measures ANOVA. Mean amplitudes centering around the N170/VPP (140-160 ms) were calculated and differences across topographies were tested by t-tests on each electrode site (Rousselet et al., 2004a). In addition, amplitudes were vector normalized according to the method employed by McCarthy and Wood (1985). *p*-values were corrected by Greenhouse-Geisser epsilon, if appropriate.

## 4.3 Results

### *4.3.1 Behavioral Results*

The average error rate across all conditions was below 5%. Because of this relatively low rate, no planned comparisons were performed. The main effect for image was significant, $F(1,16) = 4.08$, $p < .05$, in that errors were slightly increased for faces in intact contexts ($M = 4.4\%$) compared to scrambled contexts ($M = 4.0\%$).

The analyses for response times showed a main effect for facial expression, $F(1,16) = 4.07$, $p < .05$) as reflected in faster responses to happy facial expressions ($M = 665$ ms) compared to fearful facial expressions ($M = 690$ ms). A main effect was also observed for image, $F(1,16) = 7.61$, $p < .01$, reflected by slower response times for faces in intact scenes ($M = 683$ ms) compared to scrambled scenes ($M = 672$ ms), $p < .05$. A three-way interaction was observed between facial expression, context emotion and image, $F(2,15) = 3.89$, $p < .05$. Planned comparisons showed that happy faces were recognized faster in intact happy contexts ($M = 667$ ms) and neutral ($M = 668$ ms) compared to fearful contexts ($M = 682$ ms), $t(16) = 2.54$, $p < .05$ and $t(16) = 2.16$, $p < .05$ respectively. The differences were not significant between happy faces in scrambled versions of happy- ($M = 661$ ms), neutral- ($M = 656$ ms) and fearful scenes ($M = 657$ ms).

In the evaluation task after the main experiment, a main effect was found for arousal, $F(2,32) = 59.41$, $p < .001$, in that fearful contexts ($M = 3.81$) and happy contexts ($M = 3.45$) were rated as being more arousing than the neutral contexts ($M = 1.84$), $t(16) = 11.99$, $p <$

.001 and $t(16) = 8.80$, $p < .001$. No significant effects were found for the categorization task, $F(2,32) = 2.22$, $p > .05$. However, post-hoc t-tests showed that the ratings for fearful and happy categories differed significantly, $t(16) = 2.15$, $p < .05$. The intended label for happy contexts was chosen on 87% of the trials, for fearful contexts on 77% of the trials, and for neutral contexts on 81% of the trials. Thus, these results are consistent with our previous ratings on arousal and category.

*4.3.2 ERP results*

*P1 component*

The ERPs of fourteen participants showed a distinctive P1 deflection. Peaks were scored at latencies ranging from 70 to 131 ms across participants. A main effect was observed for image, $F(1,13) = 25.07$, $p < .001$, in that latencies were somewhat shorter for faces in scrambled contexts ($M = 93$ ms) than intact contexts ($M = 98$ ms). A main effect was also observed for electrode position, $F(2,12) = 23.48$, $p < .001$, in that latencies were slightly longer for electrodes PO7/8 ($M = 99$ ms) than O1/2 ($M = 93$ ms) and PO3/4 ($M = 95$ ms), both $p < .001$. No main effects were observed for facial expression and context emotion, $p > .05$.

Significant two-way interactions were observed between image and context emotion, $F(2, 12) = 9.62$, $p < .01$, and electrode position and context emotion, $F(4,10) = 3.59$, $p < .05$, which were qualified by a marginal three-way interaction between image, context emotion and electrode position, $F(4,10) = 3.10$, $p = .07$. Latencies were shorter for faces in intact happy contexts ($M = 95$ ms) than intact fearful ($M = 98$ ms) and neutral contexts ($M = 98$ ms), an effect being significant on electrodes PO3/4, $p < .001$ and $p < .01$ respectively. Although the difference was also significant for scrambled contexts, the direction of the effect was reversed, as latencies were shortest for scrambled neutral contexts ($M = 91$ ms) than scrambled happy contexts ($M = 94$ ms) on pair PO3/4, and shorter for scrambled fearful ($M = 96$ ms) than scrambled happy contexts ($M = 98$ ms) on pair PO7/8.

P1 amplitude analyses showed a main effect for image, $F(1,13) = 6.57$, $p < .05$, in that amplitudes were increased for faces in intact contexts ($M = 7.11$ µV) than scrambled contexts ($M = 6.24$ µV). None of the remaining factors showed main effects. An interaction was observed between image and electrode position, $F(2,12) = 6.12$, $p < .05$, and a marginal interaction between image, electrode position and context emotion, $F(4,10) = 2.92$, $p = .08$. This three-way interaction is explained by amplitudes being smaller for faces in intact happy contexts ($M = 6.18$ µV) compared to fearful ($M = 6.97$ µV) and neutral context images ($M = 6.91$ µV), but on electrode pair O1/2 only, respectively $p < .01$ and $p < .05$ (Figure 4.2). None of the comparisons for scrambled contexts were significant. Additionally, an interaction was found between hemisphere and context emotion, $F(2,12) = 12.44$, $p < .001$, indicating that the difference between faces in happy contexts ($M$ = 6.68 μV) and fearful contexts ($M$ = 7.14 μV) and neutral contexts ($M$ = 6.90 μV) was significant on the right hemisphere, $p < .001$ and $p = .05$ respectively, but not on the left hemisphere, $p > .10$ and $p = .09$.

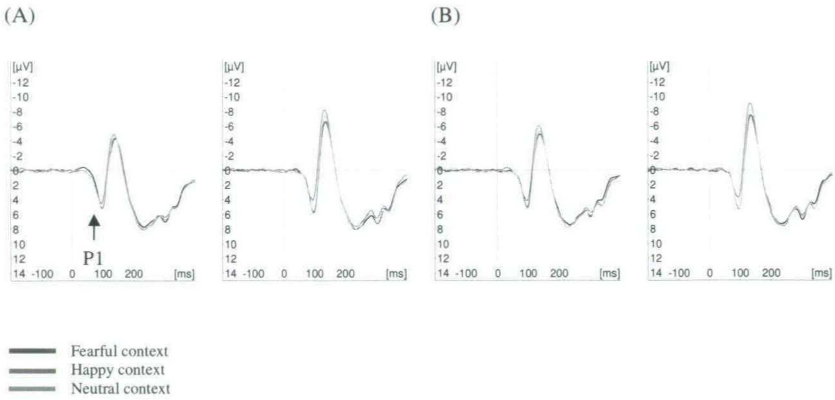



**Figure 4.2 Grand-averaged ERPs of the P1 component.** P1 for fearful (A) and happy faces (B) in fearful, happy and neutral contexts. The P1 is displayed for occipital electrode sites O1 (left) and O2 (right). Negative amplitudes are plotted upwards.

*N170 component*

The ERP of fifteen participants showed a distinctive N170. Peaks were scored at latencies ranging from 120 to 220 ms across participants. Main effects were found for the N170 latency of facial expression, $F(1,14) = 20.41$, $p < .001$ and context emotion, $F(2,13) = 15.47$, $p < .001$. Latencies for happy faces ($M$ = 148 ms) were slightly shorter than latencies for fearful faces ($M$ = 149 ms) and latencies were slightly shorter for faces in neutral ($M$ = 148 ms) than faces in fearful contexts ($M$ = 149 ms). A main effect was also found for electrode position, $F(2,13) = 19.20$, $p < .001$, as latencies were shortest for electrode-pair PO7/8 ($M$ = 146 ms) compared to electrode-pairs P7/8 ($M$ = 151 ms) and P5/6 ($M$ = 149 ms).

The analyses for the N170 amplitude showed a main effect for electrode position, $F(2,13) = 26.94$, $p < .001$, as amplitudes were most negative on electrode pair PO7/8 ($M$ = -5.96 μV), followed by P7/8 ($M$ = -4.67 μV) and P5/6 ($M$ = -2.72 μV). A main effect was also found for image, $F(1,14) = 19.87$, $p < .01$, in that amplitudes were more negative to faces in scrambled scenes ($M$ = -5.01 μV) compared to intact scenes ($M$ = -3.88 μV). A main effect was also found for facial expression, $F(1,14) = 6.36$, $p < .05$, in that amplitudes were more negative to facial expressions of happiness ($M$ = -4.64 μV) than fear ($M$ = -4.25 μV).

A three-way interaction was observed between image, electrode-position and context emotion, $F(4,11) = 6.90$, $p < .05$ (Figure 4.3). Amplitudes were more negative for faces in intact fearful contexts ($M$ = -4.51 μV ) than faces in happy contexts ($M$ = -3.94 μV) and faces in neutral contexts ($M$ = -4.05 μV), $p < .01$, $p < .05$ respectively. These results were significant on electrodes P7/8 only. It is unlikely that these effects are based on low-level features. The difference between faces in scrambled fearful ($M$ = -5.23 μV) and faces in scrambled happy contexts ($M$ = -5.03 μV) obtained marginal significance ($p$ = .06). The difference between faces in scrambled fearful contexts and faces in scrambled neutral contexts ($M$ = -5.26 μV) was not significant. However, on electrode P5/6, larger amplitudes were found for faces in scrambled neutral contexts ($M$ = -3.35 μV) than faces in scrambled happy contexts ($M$ = -3.09 μV), $p < .05$), but not faces in scrambled fearful contexts ($M$ = -3.26 μV). These differences were not found for faces in intact contexts, $p > .05$.

Planned comparisons were performed on N170 peak amplitudes for fearful and happy facial expressions separately as a function of context emotion. It was found that the N170 amplitudes were more negative for fearful facial expressions in fearful contexts compared to fearful facial expressions in happy contexts, for P7 (-4.18 μV; -3.19 μV), $t(14) = 3.74$, $p < .01$, P8 (-4.84 μV; -4.23 μV), $t(14) = 2.59$, $p < .05$ and compared to neutral contexts for P7 (-4.18 μV; -3.33 μV), $t(14) = 3.06$, $p < .01$. (Figure 4.3). For happy facial expressions, N170 amplitudes were more negative in fearful contexts compared to happy facial expressions in happy contexts for P7 (-3.90 μV; -3.52 μV), $t(14) = 2.61$, $p < .05$, and happy facial expressions in neutral contexts for P7 (-3.90 μV; 3.39 μV), $t(14) = 2.28$, $p < .05$. The comparisons for other electrodes on the occipito-temporal scalp are shown in Figure 4.4. T-tests on all electrodes for the calculated mean amplitudes (140-160 ms) were similar to peak analyses (Figure 4.4).

Planned comparisons for faces in scrambled scenes showed that the patterns for fearful faces could not be explained by low-level features. However, for happy faces, the difference on electrode site P7 for fearful scenes as compared to happy scenes was also found for scrambled scenes.

*Scalp topography analyses*

Scalp topography analyses were performed to examine further whether the implied hemispheric differences for fearful facial expressions in fearful contexts were reflected in topographical differences. For the scalp topographic distribution, mean amplitudes were calculated by the mean amplitude around the N170 peak, which occurred on ~150 ms. Figure 4.4 shows by planned comparisons for mean amplitudes at 140-160 ms that fearful faces were larger in fearful contexts than happy and neutral contexts on left occipito-temporal electrodes. The t-tests show that rather similar electrodes show differences as the planned comparisons that were performed on the N170 peaks (see N170 section). The mean amplitudes were then vector normalized according to McCarthy and Wood (1985). The topographic interaction between a large range of symmetrically positioned occipito-temporal sites (T7/8; TP7/8; CP3/4; CP5/6; P7/8; P5/6; PO7/8; PO3/4; O1/2) and context emotion was tested for facial expressions of fear and happiness separately by repeated measures ANOVA.

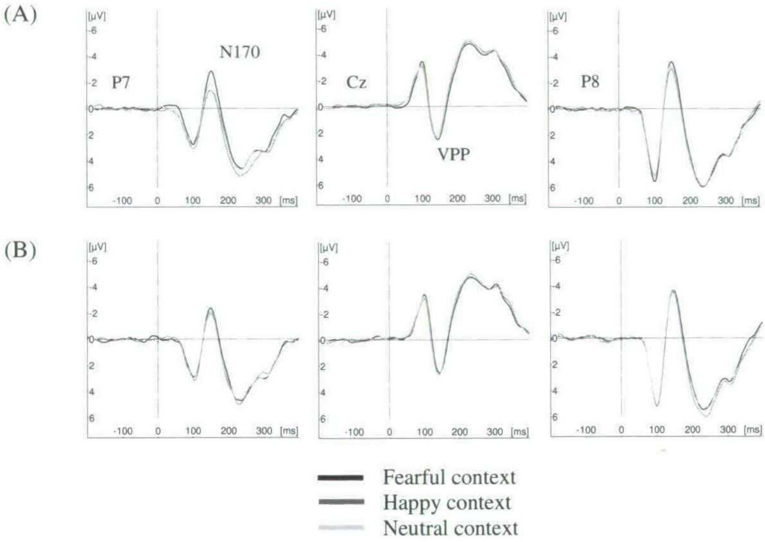



**Figure 4.3 Grand-averaged ERPs of the N170/VPP components**. N170/VPP for fearful (A) and happy faces (B) in fearful, happy and neutral contexts. The N170 is displayed for occipito-temporal electrode sites (P7/8) and the VPP is displayed for the vertex electrode (Cz). Negative amplitudes are plotted upwards.

A significant topography difference was found between fearful facial expressions in intact fearful contexts as compared to fearful facial expressions in intact happy contexts, $F(19,304) = 4.10$, $p < .05$, but it was not different from fearful facial expressions in neutral contexts, $F(19,304) = 1.28$, $p = .29$. The difference was neither significant between fearful facial expressions in happy contexts and fearful facial expressions in neutral contexts, $F(19,304) = 1.61$, $p = .18$. The comparisons across scrambled contexts were also non-significant. For happy facial expressions, no topography differences were found between faces that were presented in fearful contexts compared to happy contexts, $F(19,304) = 1.92$, $p = .13$, and neutral contexts, $F(19,304) = 0.51$, $p = .74$, and between happy and neutral contexts, $F(19,304) = 0.93$, $p =.44$. The comparisons across scrambled contexts were non significant. However, a significant topography difference was found between happy facial expressions in scrambled happy contexts and happy facial expressions in scrambled neutral contexts, $F(19,304) = 3.30$, $p < .05$.

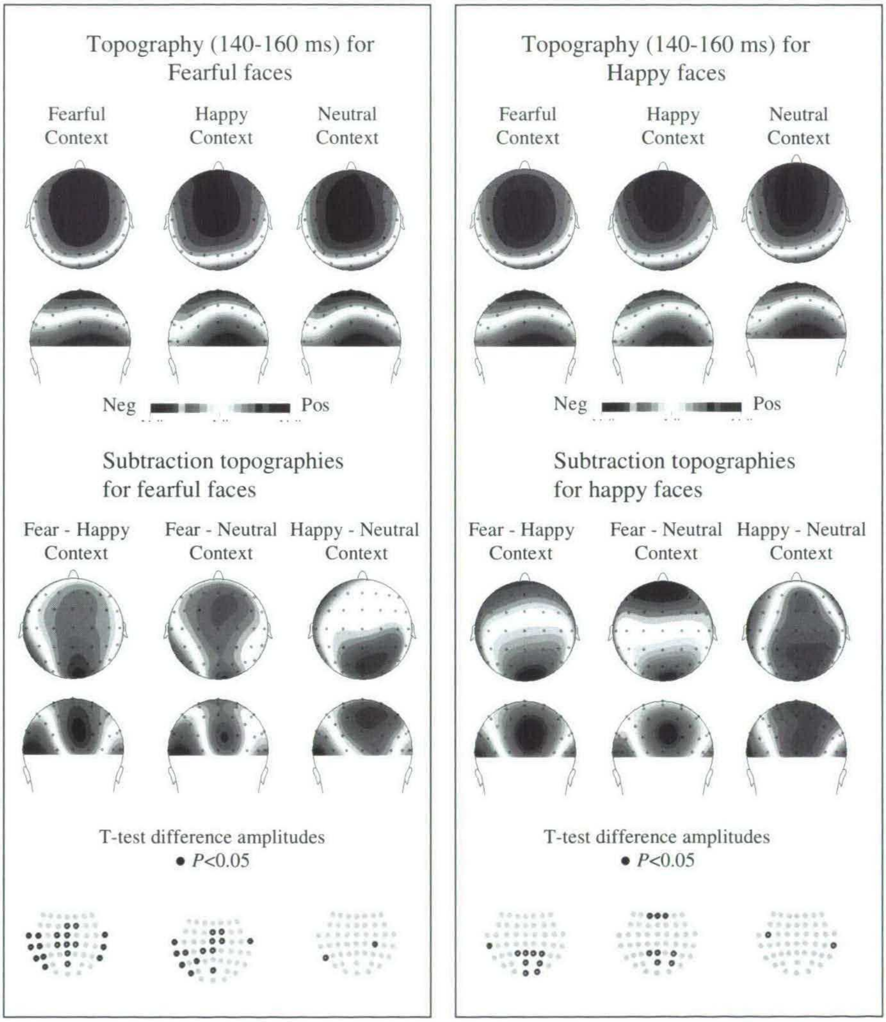



**Figure 4.4 Scalp topographies.** Scalp topographies for faces in context based on a window (140-160 ms) around the N170. Upper panel left: scalp topography of fearful faces in fearful, happy and neutral contexts show the typical central positivity of the VPP and the occipito-temporal negativities of the N170. Upper panel right: happy faces in fearful, happy and neutral contexts. Lower panel left: scalp topographies based on difference waves show the prominent left occipito-temporal response to faces in fearful contexts but neither for fearful faces in happy or neutral contexts nor for happy faces (lower panel right). Black dots in electrode-map depict sites on which difference is significant (t-test, $P < 0.05$, uncorrected).

*VPP component*

The same group of participants as in the analysis of the N170 showed a distinctive VPP. The main effect for the VPP latency reached marginal significance for facial expression, $F(1,14) = 4.14$, $p = .06$, as latencies were, like the N170 component, slightly shorter for happy facial expressions ($M$ = 144 ms) as compared to fearful facial expressions ($M$ = 146 ms). The main effect for context emotion was significant, $F(2,13) = 4.82$, $p < .05$, showing that latencies were shortest for happy and neutral contexts (both $M$ = 145 ms) as compared to fearful contexts ($M$ = 146 ms), $t(14) = 2.67$, $p < .05$ and $t(14) = 1.95$, $p < .05$ respectively.

The analyses for the VPP amplitude showed a main effect for image, $F(1,14) = 5.10$, $p < .05$, showing more positive amplitudes for scrambled contexts ($M$ = 3.51 μV) than intact contexts ($M$ = 3.20 μV). The main effect for facial expression was also significant, $F(1,14) = 7.26$, $p < .05$ (Figure 4.3), showing more positive amplitudes for happy facial expressions ($M$ = 3.48 μV) than fearful facial expressions ($M$ = 3.24 μV). This main effect was qualified by an interaction with electrode position, showing that these differences were significant for positions FCz and Cz, but not for CPz. The other interactions were not significant.

## 4.4 Discussion

In the present study, event-related potentials for facial expressions were investigated as a function of context information in the background. The contexts were scenes depicting fear-, happy- or neutral scenes. Color information was controlled for by including conditions containing scrambled contexts. The participants were required to recognize the facial expression. Several important findings were obtained. First, we found a general effect of context, as P1 amplitudes were increased for faces in intact contexts compared to faces in scrambled contexts, and N170/VPP amplitudes were smaller to faces in intact contexts than faces in scrambled contexts. Second, we observed an increased N170 amplitude for faces in fearful contexts compared to faces in happy- and neutral contexts, which was particularly increased for fearful faces on the left hemisphere electrodes.

### *4.4.1 General context effect*

The general effect of a meaningful context was analyzed by comparison of intact and scrambled contexts. We hypothesized that intact context scenes impose an additional load on face processing because contexts are meaningful to the observer and may for this reason distract attention from faces. It was found that faces were processed differently in intact contexts than scrambled contexts as was observed for both P1 and N170/VPP and on response times. The scrambled contexts were used to provide a context that has similar colour properties but in effect is meaningless. The increased P1 amplitudes for faces in intact scenes as compared to faces in scrambled scenes is consistent with a study by Ganis and Kutas

(2003). The increased N170 amplitude for faces in scrambled scenes as compared to faces in intact scenes was also expected, in that amplitudes were increased like the N170 results for faces in grey backgrounds that were used in previous work (chapter 2). In a recent study testing recognition memory for faces, it was observed that the N170 amplitude was smaller for retrieving faces that had initially been encoded in a meaningful context as compared to faces that had been studied on a white background (Galli et al., 2006).

The difference in physical properties of the context stimuli (and the control stimuli) indicates that across these studies physical properties *per se* do not explain these effects. It suggests that semantic information in the intact contexts decreases the N170 amplitude rather than that the physical information in the scrambled scenes increases N170 amplitudes. The N170 may be decreased by intact context because attention is distracted to these semantic meaningful contexts. It has been shown that the N170 amplitude can be affected by spatial attention (Holmes et al., 2003; but see also Cauquil et al., 2000), and for this reason the N170 amplitude may decrease when attention is distracted away from faces. This interpretation is consistent with the behavioral data that showed that response times were slowed for faces in intact contexts. Other behavioral work using different paradigms add to this finding by showing that context information may interfere with target processing (Lavie, 1995). Davenport and Potter (2004) reported that a briefly presented object was recognized more accurately when it was isolated from any context as compared to objects accompanied by contexts. An interesting question may be how the complexity of context information affects face processing. Several studies have already shown how the complexity scenes affects perception (Blanchette, 2006; Carretie et al., 2004).

### *4.4.2 Specific emotion effects of contexts*

In addition to the general effects of meaningful scenes, the emotional valence of the scenes is important for face processing. Specific emotion effects of contexts were observed in the behavioral and ERP data. The behavioral data showed that recognition of happy facial expressions was significantly faster in the happiness- than fear-inducing contexts, whereas the faster recognition of fearful facial expressions in fear-inducing contexts compared to happiness-inducing contexts did not reach significance. These data replicate our previous findings (chapter 3). The differences for scrambled scenes were not significant and therefore it is unlikely that low-level features explain these results. A novel finding in this study is that response times for happy faces in neutral contexts were also faster than fearful contexts. This suggests that the effect represents a delay by incongruent contexts rather than facilitation by congruent contexts.

The ERP data showed different patterns for the P1 component compared to the N170/VPP component. P1 amplitudes were larger to faces in fearful and happy contexts compared to neutral contexts. This is consistent with literature that has shown P1 effect for emotional scenes without faces (Carretie et al., 2004; Smith et al., 2003). No main effects were observed for facial expression while these effects have been observed before on the P1 component (Batty & Taylor, 2003; Eger et al., 2003). However, Meeren et al. (2005) have shown that the P1 amplitude was smaller for facial expressions that were congruent with the accompaning bodily expressions as compared to incongruent face-body pairs. However, it should be noted that fear and angry expressions were used, and for this reason it may not be excluded that similar emotions for faces in the context of visual scenes would result in similar patterns.

However, in contrast to the P1, increased N170 amplitudes were found on the left hemispheric electrode sites when fearful faces were accompanied by fearful contexts rather than happy or neutral contexts. This is in agreement with earlier findings (chapter 2). The differential amplitudes of the N170 on left hemisphere electrodes may correspond to left hemispheric responses for contextualized faces as has been found in an fMRI study by Kim et al. (2004). In this study, participants were presented with surprised faces that were cued by negative sentences (e.g., about losing money) or positive sentences (e.g., about winning money). In a contrast analysis of negatively cued faces against positively cued ones, a left hemispheric response of the amygdala and fusiform gyrus was found. It has been suggested that the N170 may be generated by the fusiform gyrus (Herrmann, Ehlis, Ellgring et al., 2005; Pizzagalli et al., 2002). The amplitude of the N170 to faces has been related to the BOLD response in the fusiform gyrus (Iidaka et al., 2006). Future fMRI studies may investigate whether these brain regions are also involved for faces in visual contexts or whether it is confined to faces in verbal contexts.

The present study used distinct emotions in face and context. The comparison with other emotions is important to test whether emotions by face and context integrate for specific emotions. The results are consistent with a role of the N170 in analyzing emotions from face and context. As fearful contexts increase the N170 amplitude for fearful faces significantly but not for happy faces, emotions may integrate specifically for fear at this stage of encoding. Previous work had already been shown that the N170 amplitude is modified by facial expressions of fear (Batty & Taylor, 2003; Stekelenburg & de Gelder, 2004; Williams et al., 2006) and that fearful contexts increase the N170 amplitude for fearful faces even more (chapter 2). The results suggest that the differences occur irrespective of the given task instruction, but are inconclusive as to whether the magnitude of these differences changes as a function of task instructions (see also Caharel et al., 2005).

It should be noted that we used a different set of face and context stimuli than used in our previous work (chapter 2), which leaves an explanation by low-level features rather unlikely. This is not to say that low-level features (e.g., color information) do not affect emotion perception. The peak latencies of both the P1 (~100 ms) and the N170 (~150 ms) components were considerably shorter than in our previous study (~130 ms and ~180 ms respectively) in which gray scale stimuli were used (chapter 2). This may relate to faster processing of faces in contexts when colour stimuli are used. However, shorter latencies have also been found by other groups using gray scale stimuli (Batty & Taylor, 2003; Campanella, Quinet, Bruyer, Crommelinck, & Guerit, 2002), which suggest that other factors, like task demands, are important as well. For instance, Gazzaley et al. (2005) have used task conditions in which participants were required to remember faces in one experimental block, and ignore faces in another block. The results showed that the N170 latency was shorter if participants were instructed to remember the face. These data suggest that the latency is modified by attentional factors, although other studies have not found effects of attention (Cauquil et al., 2000).

Some low-level features may be integral parts of the perceived emotion. Previous work has shown that for instance color may be important for early stages of encoding (Spence et al., 2006) and for this reason it may be potentially important for ERP studies examining early components. Intracranial recording studies have shown that early ERP components are affected by colored checkerboards (Allison et al., 1993). We found that scrambled contexts exerted some influence on the N170. However, the effects of the intact context cannot be explained by low-level features entirely. First, the N170 responses to faces in scrambled contexts were larger than faces in intact contexts (see also VPP). Second, the differences across conditions were smaller than for the intact images, and reached only marginal significance. Third, and most importantly, the direction of the amplitude differences were not similar for conditions of intact and scrambled contexts.

It may be interesting to investigate how individuals use context information of facial expressions that are more ambiguous and thus harder to recognize (de Gelder & Vroomen, 2000). In the present study, participants could recognize the facial expressions relatively easily. It has been shown that ambiguous expressions are more difficult to recognize, and more importantly, that the auditory context can bias the response to the emotion present in the context (de Gelder et al., 2006). Similarly, visual contexts may shift the response towards the emotion present in the scene, and may give more insight into the function of the N170 in encoding face and context.

### *Conclusions*

The present study examined how recognition of facial expressions is affected by the accompanying context. Behavioral and ERP measures were used in an experiment in which participants were required to recognize the facial expression. First, it was found that N170 amplitudes were smaller to faces in intact contexts as compared to faces in scrambled contexts. The results suggest that meaningful contexts may distract attention away from face processing as. This finding extends our previous study in which gray backgrounds were used. Second and more important, the N170 amplitudes were larger for faces in fearful context compared to happy and neutral contexts. The N170 amplitudes were even more increased for fearful faces in fearful contexts as compared to fearful faces in happy contexts. Future studies should determine whether the magnitude of the observed differences depends on the type of task participants perform. Studies using neuropsychological groups (e.g., patients defective on emotion recognition) may in addition reveal whether this increased N170 response to fearful faces in fear-inducing contexts depends on the capacity to identify emotions.

*Acknowledgement*

We are grateful to Thessa Caus for help and assistance in the EEG experiment.

# Chapter 5

# The role of contexts in the processing of facial expressions in Schizophrenia. An ERP study

*Abstract*

Behavioral and neuroimaging studies have shown face recognition deficits in schizophrenia patients. Traditionally, most experiments have studied face recognition isolated from the context. Yet, in populations with poor recognition, the context may provide critical information of how a facial expression is interpreted. It has as yet not been tested how schizophrenia patients recognize facial expressions in context. Seventeen patients with schizophrenia and eighteen healthy control participants were presented with stimuli of facial expressions (fear, happy) in natural scenes (fear, happy, neutral). Event-related potentials (ERPs) were recorded in order to assess whether the emotional gist of scenes has an influence on the N170 response. Schizophrenia patients showed like healthy control participants larger N170 amplitudes to fearful faces in fearful contexts compared to fearful faces in happy contexts on left occipito-temporal electrodes. However, in contrast to the healthy control group, it was found that the N170 amplitudes were not larger for fearful faces in fearful contexts as compared to fearful faces in neutral contexts. The results suggest that the perception of facial expressions in schizophrenia is largely influenced by the emotional valence (positive-negative) of the emotional scene in context.

This chapter will be submitted as: Righart, R., de Jong, S., Caus, T., Hodiamont, P., & de Gelder, B. (2007). The role of contexts in the processing of facial expressions in schizophrenia. An ERP study.

## 5.1 Introduction

Schizophrenia is a major psychiatric disorder mainly defined by hallucinatory-delusional symptoms, thought disorder, social withdrawal and emotional flattening. The latter two symptoms importantly affect emotional functioning and social communication and may associate with the ability to perceive facial expressions and social cues from the environment (Keltner & Kring, 1998).

Deficits of facial expression recognition in schizophrenia has been shown by a number of experimental studies (Dougherty, Bartlett, & Izard, 1974; Feinberg et al., 1986; Kohler et al., 2003; Walker, Marwit, & Emory, 1980). However, it has been found that this deficit not necessarily reflects specific emotion recognition deficits, but that poor performance in schizophrenia may generalize to other non-emotional recognition tests as well (Feinberg et al., 1986; Whittaker, Deakin, & Tomenson, 2001). For example, it was shown that schizophrenia patients perform equally poor on emotion recognition and identity recognition tests, and that overall performance was poorer for schizophrenia patients than normal healthy control participants (Kerr & Neale, 1993; see also Kohler, Bilker, Hagendoorn, Gur, & Gur, 2000; Salem, Kring, & Kerr, 1996). Others have found, in contrast, that schizophrenia patients showed a relative larger impairment on emotion recognition than non-emotional face recognition (Kosmidis et al., 2007; Penn et al., 2000; Schneider et al., 2006).

Another issue that has not been raised, to our knowledge, is how emotional salient contexts influence the way facial expressions are recognized. Traditionally, investigations of face recognition have not paid attention to the context in which faces are seen. However, investigating face recognition as a function of the accompanying context may enhance the ecological validity. In daily life, perception of facial expressions does not occur separately from the social situation, but facial expressions reflect emotional reactions of a person that actively responds to its situation (Frijda & Tcherkassof, 1997).

A number of experiments have shown that context information importantly contributes to how normal healthy individuals judge facial expressions. In a behavioral experiment, facial expressions of fear, happy and disgust were better and faster recognized when shown in congruent contexts as compared to incongruent contexts (chapter 3). An important question is how emotional contexts contribute to face recognition in schizophrenia. In the present study, it was investigated whether emotional information in context, by presenting emotional scenes, affects the recognition of facial expressions in schizophrenia.

Several studies have provided insight into how schizophrenic patients perceive emotional scenes. Behavioral data have shown that evaluations by schizophrenic patients were not different from evaluations by healthy control participants. Pleasant pictures were rated higher on valence than neutral pictures, and unpleasant pictures were rated lower on valence (Bigelow et al., 2006). In addition, unpleasant and pleasant pictures were rated as being more arousing than neutral pictures (Hempel et al., 2005; Rockstroh, Junghofer, Elbert, Buodo, & Miller, 2006).

In contrast to these behavioral data, a number of studies have shown that schizophrenia patients differ from healthy control participants in their physiological (Hempel et al., 2005) and neural response (Rockstroh et al., 2006) to emotional scenes. In contrast to healthy control participants, who showed higher responses to emotional stimuli in the visual regions of the brain, schizophrenics did not show significant different responses in these regions (Takahashi et al., 2004; Taylor, Phan, Britton, & Liberzon, 2005). These differential responses between patients and control participants may relate to early stages of stimulus processing (~100 ms) as shown by magnetoencephalography (MEG) (Rockstroh et al., 2006). Similar to MEG, Event-related potentials (ERPs) have shown differential processing of emotional and neutral scenes from 100 ms onwards (i.e., the P1 component) in healthy control participants (Carretie et al., 2004; Smith et al., 2003).

To our knowledge, these early differences have not been tested for schizophrenia patients as yet in research using ERPs. ERPs provide an excellent temporal resolution about when the neural processing between two or more experimental conditions differs significantly. It has been shown that the N170 ERP component, which occurs ~170 ms after stimulus presentation, is related to face processing. The N170 component is maximally negative at occipito-temporal sites on the scalp. The N170 is larger to faces as compared to non-face objects (Bentin et al., 1996; Itier & Taylor, 2004b). This stage of processing may relate to the structural encoding of identity (Bentin et al., 1996; Eimer, 2000). It has, however, been observed that the N170 amplitude is sensitive to facial expressions of emotion (Batty & Taylor, 2003; Caharel et al., 2005; Stekelenburg & de Gelder, 2004; Williams et al., 2006; but see also Eimer & Holmes, 2002; Eimer et al., 2003). It has also been observed that the N170 amplitude for faces is influenced by the emotional context (chapter 2 and 4).

Since face recognition involves different stages of processing, face recognition problems in schizophrenia patients may relate to deficits in these stages. Several studies have used ERPs to study face processing in schizophrenia. It has been shown that schizophrenia patients showed less differentiation between faces and images of buildings on the N170 amplitude as compared to healthy control participants (Herrmann et al., 2004). Another study reported that the N170 amplitude to faces was not significantly larger than images of cars and hands in schizophrenia patients, in contrast to the results obtained for healthy control participants (Onitsuka et al., 2006). These studies suggest that face recognition deficits in schizophrenia patients relate to early stages of encoding. It is however unclear whether this deficit is exclusively related to identity processing or whether this is also related to the processing of facial expressions. In an experiment that recorded ERPs to fearful, happy and sad facial expressions in both schizophrenia patients and healthy control participants, it was found that the N170 amplitude to fearful faces was increased over happy- and sad faces for healthy control participants, but not for schizophrenia patients (Campanella et al., 2006). This latter study thus suggests that the deficits on the N170 relate not only to the encoding of identity but also to the encoding of facial expressions.

The present study investigated the categorization of facial expressions (fear, happy) presented in emotional scenes (fear, happy, neutral) using behavioral and ERP measures. Participants were required to categorize facial expressions. It was hypothesized that the categorization of facial expressions would not be influenced by the emotional scenes in schizophrenia patients. Specifically, given that schizophrenia patients show emotion recognition deficits, we expect that the N170 component will not differentiate between facial expressions and emotional contexts, which would contrast to the results obtained for healthy control participants (chapter 2 and 4).

## 5.2 Method

### *5.2.1 Participants*

Seventeen patients with schizophrenia (2 female), 14 were right handed, ranging from 22 to 58 years ($M$ = 39 years) participated in the experiment. All patients were under treatment at the local day care hospital GGZ Psychiatric hospital, Midden Brabant, Tilburg, and were clinically stable at the time of testing. Only patients that fulfilled the criteria for schizophrenia or schizoaffective disorder according to the DSM-IV were selected for participation. Diagnosis was established with the Schedules for Clinical Assessment in Neuropsychiatry (SCAN 2.1), a standardized interview for diagnosing axis I disorder, conducted by a trained investigator. Fourteen patients were using antipsychotic medication (four typical antipsychotica) at the time of testing. Six patients were receiving antidepressants and four were receiving anxiolytics.

The control group were eighteen (6 female) age-matched healthy participants, 13 were right handed, ranging from 21 to 50 years ($M$ = 32 years) that were studied in the previous experiment (chapter 4). Both groups had a higher proportion of male than female participants ($X^2 = 10.31, p < .001$). The proportion of male participants did not differ significantly across groups ($X^2 = 2.31, p = .13$). One participant was removed as she reported to have had visual problems (glaucoma). The remaining participants had normal or corrected-to-normal vision. None of them reported a history of neurological or psychiatric disease and were not receiving medications at the time of testing. All had given informed consent. The study was performed in accordance with the ethical standards of the Declaration of Helsinki. All participants of the healthy control group and the patient group were paid for their participation. Education was translated into average years of the highest accomplished scholarship. As expected, patients attained less years of education ($M = 5.2$ years) than did healthy control participants ($M = 6.3$ years), but the difference was marginally significant, $t(33) = 1.93$, $p = .06$.

### *5.2.2 Stimuli*

Stimuli consisted of faces centrally embedded on scene pictures (For an example, see Figure 5.1). Face stimuli were color pictures of caucasian faces from the Karolinska Directed Emotional Faces set (Lundqvist et al., 1998). In a pilot experiment, face pictures were rated by a separate group of participants in order to select facial identities that expressed both fear and happy emotions validly. 24 male and 24 females conveying facial expressions of fear and happiness were selected from the set resulting in a total of 96 faces.

Scene pictures serving as contexts consisted of color pictures selected from the IAPS (Lang et al., 1999) and from the web. (See for arousal and categorization procedure chapter 2 and 4). Potential artefacts introduced by low-level features were countered by preparing scrambled versions of the scene stimuli. Scene pictures were scrambled by randomizing the pixels across the image (blocks of 8 x 8 pixels were randomized across the image measuring 768 x 572 pixels width and height respectively), thereby rendering the pictures meaningless. The resulting pictures were inspected carefully in order to avoid that certain features would cue recognition.

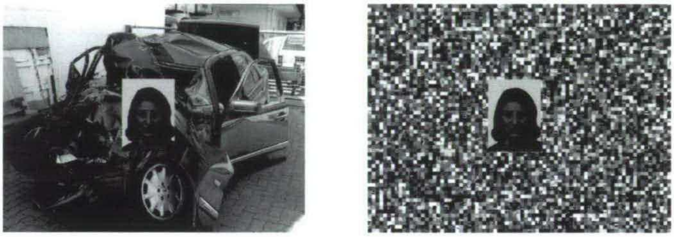
Faces were overlaid on the center of each scene. The height and width of the facial images was 7.9 cm x 5.9 cm (5.6° x 4.2°) and for context images 24.5 cm x 32 cm respectively (17.4° x 22.6°). Faces did not occlude the critical information in the context picture.

### *5.2.3 Procedure*

Participants were seated in an electrically shielded cabin with dimmed illumination with a monitor positioned at 80 cm distance. They were instructed and familiarized with the experiment by practice sessions. The stimuli were presented for 200 ms. Stimuli were prompted by a fixation cross. Participants performed a two-alternative forced choice task in which they indicated whether the facial expression was happy or fearful. Responses were recorded from stimulus-onset. No feedback was given. They were instructed to respond as accurately and fast as possible. Response-buttons were counterbalanced across participants. The intertrial interval was randomized between 1200 and 1600 ms.

The experiment was run in 16 blocks each containing 72 trials of face-context compounds, 8 blocks of faces with intact scenes and 8 blocks of faces with scrambled scenes. Blocks containing intact and scrambled scenes alternated (with order randomized across subjects). Each block took about three minutes. Facial expressions (fear, happy) were paired to a scene of each context emotion category (fear, happy, neutral) in order to balance the design factorially. Altogether, intact and scrambled blocks amounted to 12 conditions each containing 96 trials. After two blocks, facial identity and context pairs were changed.

a) b)



c)

| | Arousal | | | Categorization | | |
|---|---|---|---|---|---|---|
| | Fearful | Happy | Neutral | Fearful | Happy | Neutral |
| HC | 3.81 (.53) | 3.45 (.74) | 1.84 (.56) | 77% (23) | 87% (11) | 81% (13) |
| Sch | 3.46 (.74) | 3.30 (.78) | 2.27 (.86) | 56% (35) | 78% (16) | 73% (24) |

**Figure 5.1**. **Stimuli.** Face-context stimuli. Facial expressions of fear and happiness were combined with context scenes conveying a fearful, happy or neutral situation. (a). An example of a face-context compound showing a facial expression of fear and a carcrash and (b) the same facial expression shown in a scrambled version of the carcrash that was used as a control stimulus. (c). Arousal rating and categorization of scenes for healthy controls (HC) and schizophrenia patients (Sch). Participants rated arousal on five designated button from 1 (calm) to 5 (excited). The average of arousal ratings is depicted. Categorization of scenes was performed by using three designated buttons. Percentage indicates the proportion of scenes that was categorized correctly. Standard deviations are in parentheses.

After the main experiment, participants received an evaluation task in which they were required to evaluate the scenes on arousal and emotion category. Arousal was measured on a five-point scale because of response box limitations in the ERP lab (five-point scale from 1=calm to 5=extremely arousing). Categorization required a response between three alternatives (fear, happy, neutral). Arousal and valence were measured in separate blocks. Scene stimuli were presented without faces and were presented as long as the participant needed to respond. However, participants were encouraged to base their evaluation of the scenes on their first impression. The order of the blocks was counterbalanced across participants.

### *5.2.4 EEG recording*

EEG was recorded from 49 electrode locations using active Ag-AgCl electrodes (BioSemi Active2) mounted in an elastic cap referenced to an additional active electrode (Common Mode Sense). EEG was bandpass filtered (0.1-30 Hz, 24 dB/Octave). The sampling rate was 512 Hz. All electrodes were offline referenced to an average reference. Horizontal electrooculographies (hEOG) and vertical electrooculographies (vEOG) were recorded. Raw data were segmented into epochs starting 200 ms before stimulus onset to 1000 ms after stimulus onset. The data were baseline corrected to the first 200 ms.

EEG was EOG corrected by using the algorithm of (Gratton et al., 1983). Epochs exceeding 100 μV amplitude difference at any channel were removed from analyses. On average, 83.6 trials (range across conditions: 82.3-84.9) and 84.9 trials (range 81.7-86.1) were left for faces in intact contexts and faces in scrambled contexts respectively after removal of the artifacts for schizophrenia patients. 87.5 trials (range across conditions: 86.8-88.5) and 85.1 trials (84.4-85.4) were left for analyses of ERPs for faces in intact and scrambled contexts for healthy control participants. After removal of trials containing inaccurate responses or responses below 200 ms, ERPs were averaged for conditions of facial expressions of fear in intact fearful, happy and neutral contexts, and for happy facial expressions in the same context categories. Similarly, averages were computed for faces in scrambled contexts, resulting in a total of 12 conditions.

Electrode selection for the N170/VPP analyses was based on previous studies. The N170 was analyzed on occipito-temporal sites (P5/6, P7/8 and PO7/8). The VPP was analyzed on midline sites (FCz, Cz, CPz). Based on grand average ERP inspection, peaks were identified using a time-window of 100-200 ms.

### *5.2.5 Data analyses*

Behavioral analyses were performed for error rate (percentage of incorrect responses) and response times (average of correct responded trials). Three schizophrenia patients performed poorly (accuracies below 0.60 on one or more conditions) in the experiment and were therefore excluded from analyses. Response time data were inspected on outliers for each participant separately. Response times that exceeded 2.5 SD from the mean of each condition were removed from analyses. Using these criteria, on average 2.6% of the trials were removed for the healthy control participants, and 2.9% were removed for schizophrenia patients. ERP analyses were performed for the N170 and VPP component.

Main- and interaction effects were analyzed between groups by using multivariate repeated measures ANOVA containing the factors Image (intact, scrambled), Facial expression (fear, happy) and Context emotion (fear, happy, neutral). Firstly, the general effect of contexts was analyzed by the factor Image, which compares the effects of faces in intact contexts to faces in scrambled contexts. For ERP analyses, the factors Hemisphere (left, right) and Electrode position (P5/6; P7/8; PO7/PO8) were included. Previous work with healthy participants has shown repeatedly that meaningful context information influences face processing (chapter 2, 4 and Galli et al., 2006).

Secondly, specific effects of context emotions on face processing were analyzed by using multivariate repeated measures ANOVA and planned comparisons by using one-tailed t-tests ($\alpha = .05$, directional hypothesis) (Howell, 2002). For the behavioral data, the specific hypothesis for healthy control participants that fearful faces are faster recognized in fearful contexts (congruent) as compared to happy contexts (incongruent), and that happy faces are faster recognized in happy contexts (congruent) as compared to fearful contexts (incongruent) was also tested for schizophrenia patients. For the ERP data, the specific hypothesis for healthy control participants that the N170 amplitude would be more negative for fearful faces in fearful contexts as compared to fearful faces in happy and neutral contexts was also tested for schizophrenia patients.

## 5.3 Results

### *5.3.1 Evaluation of scenes*

In the evaluation task after the main experiment, a main effect was found for arousal, $F(1.96;56.84) = 69.79$, $p < .001$, in that fearful contexts ($M = 3.64$) and happy contexts ($M = 3.38$) were more arousing than neutral context ($M = 2.06$), $t(30) = 10.04$, $p < .001$ and $t(30) = 9.62$, $p < .001$ respectively (the difference between fear-happy, $t(30) = 1.84$, $p = .08$). The interaction between emotion and group reached significance, $F(1.96;56.84) = 3.97$, $p < .05$. However, the effects for schizophrenic patients were similar to the effects of the control group. Schizophrenia patients rated fearful scenes ($M = 3.46$, $SD = 0.74$) and happy scenes ($M = 3.30$, $SD = 0.78$) as more arousing than neutral scenes ($M = 2.27$, $SD = 0.86$), $t(13) = 4.61$, $p < .001$ and $t(13) = 5.41$, $p < .001$ respectively. The difference between fearful scenes and happy scenes was not significant ($p > .05$). Similarly for healthy control participants, it was found that fearful scenes ($M = 3.81$, $SD = 0.53$) and happy scenes ($M = 3.45$, $SD = 0.74$) were rated as more arousing than neutral scenes ($M = 1.84$, $SD = 0.56$), $t(16) = 11.99$, $p < .001$ and $t(16) = 8.80$, $p < .001$. The healthy control participants evaluated the fearful and happy scenes as slightly more arousing than the schizophrenia patients, while the evaluation was reversed for neutral scenes. The groups were compared for each of the emotions but none of the comparisons was significant (all $p > .05$).

The analyses of the categorization task showed a main effect for emotion, $F(1.76; 50.89) = 5.20$, $p < .05$, in that the percentage of categorizations differed between fearful scenes (average across groups $M = 67\%$) and happy scenes (83%). The neutral category fell in between (77%), but was not significant different from the other categories. No significant difference was found between groups ($p > .05$).

### *5.3.2 Behavioral results*

#### *5.3.2.1 General context effects*

For the error rates, no main effects were found for image. However, a significant interaction was found between group and image, $F(1,29) = 4.03$, $p < .05$. The error rate for faces in scrambled scenes was slightly higher for schizophrenia patients ($M = 6.3\%$) than healthy control participants ($M = 4.0\%$), $t(29) = 2.19$, $p < .05$; the error rate for faces in intact scenes was also higher for schizophrenia patients ($M = 5.9\%$) than healthy control participants ($M = 4.4\%$), but this difference was non-significant, $t(29) = 1.26$, $p > .05$. No further interactions or main effects were found for group.

Although response times were slower for schizophrenia patients ($M = 749$ ms) than healthy control participants ($M = 678$ ms), this difference did not reach significance in the between-group comparisons, $F(1,29) = 2.10$, $p > .05$. A main effect was found for image, $F(1,13) = 4.59$, $p < .05$, in that schizophrenics had, like the control group, slower response times for faces in intact scenes ($M = 759$ ms) compared to faces in scrambled scenes ($M = 740$ ms).

#### *5.3.2.2 Specific emotion effects of contexts*

For the error rate, no significant effects were found of context emotion or facial expression. For response times, a main effect was observed for facial expression, $F(1,13) = 5.29$, $p < .05$, in that schizophrenics had faster response times for facial expressions of happiness ($M = 738$ ms) as compared to facial expressions of fear ($M = 779$ ms). Unlike the control group, no interaction effect was found between facial expression and context emotion. Whereas faster response times were found for happy faces in happy contexts ($M = 667$ ms) and neutral contexts ($M = 668$ ms) compared to fearful contexts ($M = 682$ ms) in the control group, this effect was not found for schizophrenic patients, as happy faces in happy contexts ($M = 736$ ms), and happy faces in neutral contexts ($M = 736$ ms) differed non-significantly from happy faces in fearful contexts ($M = 742$ ms).

### *5.3.3 ERP Results*

#### *5.3.3.1 General context effects*

*N170 component*

The ERPs of eleven schizophrenia patients showed a distinctive N170. Peaks were identified at latencies ranging from 117 ms to 197 ms. General context effects were analyzed by comparison of faces in intact scenes to faces in scrambled scenes.

For the N170 latency, schizophrenia patients showed a main effect for image, $F(1,10) = 5.39$, $p < .05$, in that N170 latencies were slightly shorter for faces in intact scenes (150 ms) than faces in scrambled scenes (151 ms). A main effect was observed for electrode position, $F(2,9) = 7.86$, $p < .05$, in that latencies were shorter for electrode-pair PO7/8 ($M = 149$ ms) as compared to electrode-pair P5/6 ($M = 152$ ms).

For the N170 amplitude, schizophrenia patients showed a main effect for electrode position, $F(2,9) = 16.54$, $p < .001$, in that N170 amplitudes were most negative on electrode pair PO7/8 ($M = -7.65$ μV), followed by P7/8 (-6.21 μV) and P5/6 (-3.21 μV). A similar pattern was obtained for healthy control participants.

A main effect was found for image, $F(1,10) = 12.21$, $p < .01$, reflected by larger N170 amplitudes for faces in scrambled scenes ($M = -6.27$ μV) as compared to faces in intact scenes ($M = -5.11$ μV). An interaction was observed between image and hemisphere, $F(1,10) = 3.40$, $p < .05$, as the effect of image was significant on right hemisphere electrodes only. The effect of image was also found for the healthy control participants but there was no interaction with hemisphere.

*VPP component*

For the VPP latency, unlike the control group, no main effect was found for facial expression. A main effect was observed for electrode position, $F(4,7) = 3.02$, $p < .05$, which was explained by significant different latencies between electrode FCz ($M = 147$ ms) and Cz ($M = 145$ ms), $t(10) = 2.49$, $p < .05$. For the VPP amplitude, no main effect was found for image in schizophrenics.

*5.3.3.2 Specific emotion effects of contexts*

*N170 component*

For the N170 latency, schizophrenia patients showed a main effect for context emotion, $F(2,9) = 4.94$, $p < .05$, which was reflected by latencies that were slightly shorter for faces in neutral ($M = 150$ ms) compared to faces in fearful contexts ($M = 151$ ms). This effect is consistent with the result found for the control group.

Schizophrenia patients showed larger N170 amplitudes for facial expressions of happiness (-5.23 μV) than facial expressions of fear (-4.98 μV), but unlike the healthy control participants, this difference did not reach statistical significance, $F(1,10) = 2.19$, $p > .05$. An interaction was found between context emotion and electrode position, $F(4,7) = 5.77$, $p < .05$, which was qualified by an interaction with facial expression, $F(4,7) = 3.70$, $p < .05$ (Figure 5.2, 5.3 and 5.4). This effect was reflected by N170 amplitudes that were more negative on electrode-pair PO7/8 for happy faces in happy contexts (-7.43 μV) and neutral contexts (-7.39 μV) as compared to happy faces in fearful contexts (-6.49 μV), $t(10) = 2.74$, $p < .05$ and $t(10) = 2.42$, $p < .05$ respectively.

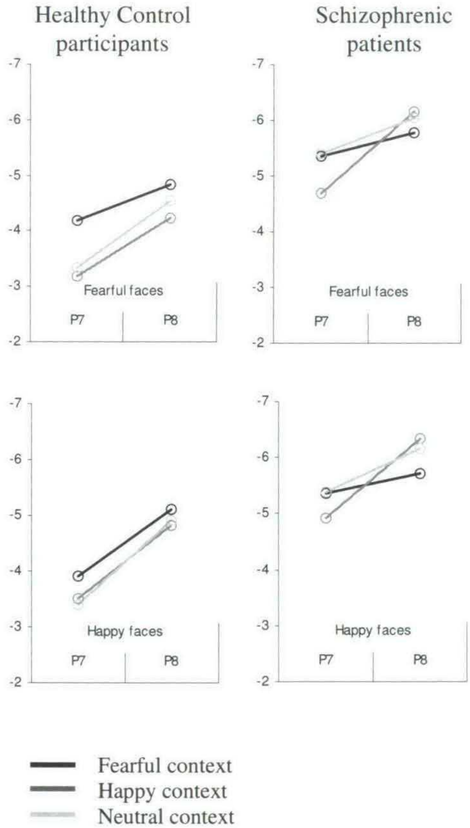



**Figure 5.2 N170 plots for fearful and happy facial expressions as a function of context emotion.** N170 peak amplitudes for schizophrenia patients and healthy control participants for fearful and happy facial expressions accompanied by fearful, happy or neutral scene contexts. The X-axis displays the left occipito-temporal electrode site (P7) and the right occipito-temporal electrode site (P8). The Y-axis displays the peak amplitude (μV). Planned comparisons show that the N170 amplitudes are more negative for fearful faces in fearful contexts as compared to fearful faces in happy and neutral contexts. These effects are most clear for healthy control participants. For schizophrenia patients, N170 amplitudes were increased on left occipito-temporal sites for fearful faces in neutral and fearful contexts in comparison to happy contexts.

In order to investigate whether the effects differed for fearful facial expressions and happy facial expressions, planned comparisons were performed for N170 peak amplitudes (see Figure 5.2). It was found that N170 amplitudes were more negative for fearful facial expressions in fearful contexts compared to fearful facial expressions in happy contexts, for P7 (fearful contexts = -5.36 μV; happy contexts = -4.70 μV), $t(10) = 2.79$, $p < .05$. This is consistent with the results obtained for healthy control participants (Figure 5.3). In contrast to the healthy control participants, N170 amplitudes were more negative for fearful facial expressions in neutral contexts as compared to fearful facial expressions in happy contexts, for P7 (neutral contexts = -5.40 μV; happy contexts = -4.70 μV)[4].

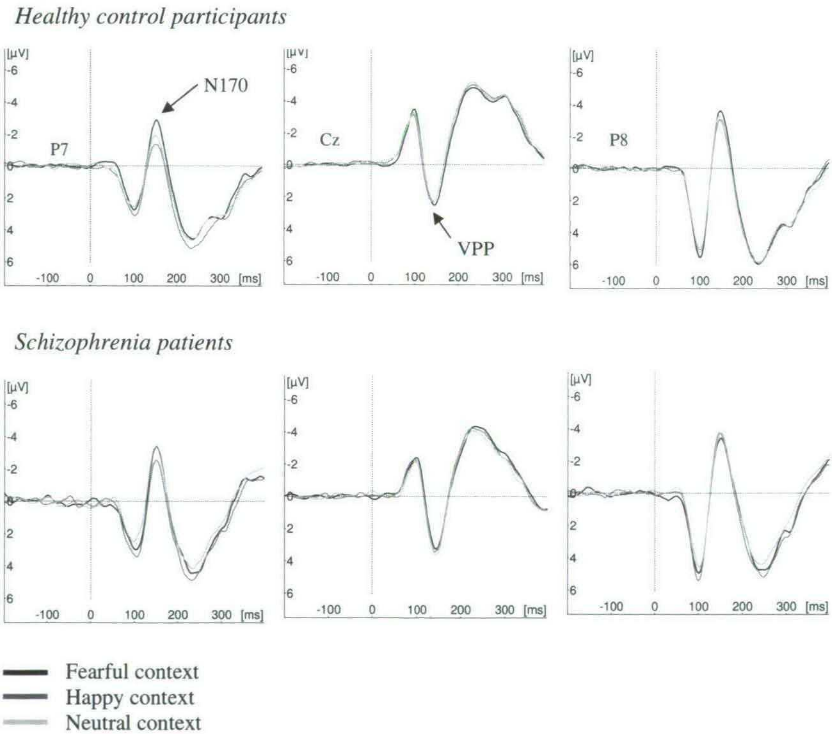



**Figure 5.3 Grand-averaged ERPs of the N170/VPP for fearful facial expressions.** Fearful facial expressions in fearful, happy and neutral contexts for healthy control participants and schizophrenia patients. Negative amplitudes are plotted upwards.

[4] Mean amplitudes using a time-window 140-160 ms were analyzed for fourteen schizophrenia patients (i.e., including three patients lacking a clear N170 peak). Similar results were found for the planned comparisons, showing more negative N170 amplitudes for fearful faces in fearful contexts as compared to fearful faces in happy contexts, $t(13) = 4.20$, $p < .01$ (11 out of 14 patients showed similar pattern, binomial $P(11/14) < .05$), but not if compared to fearful faces in neutral contexts. N170 amplitudes were more negative for fearful faces in neutral contexts as compared to fearful faces in happy contexts, $t(13) = 2.74$, $p < .05$ ($P(10/14) = .09$).

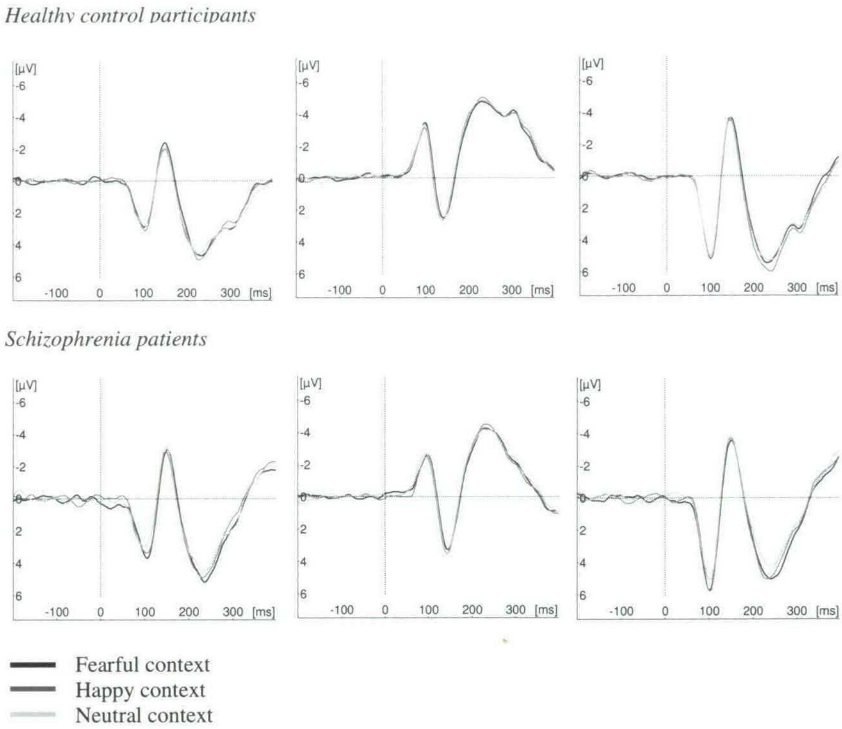



**Figure 5.4 Grand-averaged ERPs of the N170/VPP for happy facial expressions**. Happy facial expressions in fearful, happy and neutral contexts for healthy control participants and schizophrenia patients. Negative amplitudes are plotted upwards.

For happy facial expressions, N170 amplitudes were more negative if accompanied by neutral contexts as compared to happy contexts, for P7 (neutral contexts =-5.39 μV; happy contexts = -4.93 μV), $t(10) = 2.17$, $p < .05$ (Figure 5.4). N170 amplitudes were more negative for happy faces in neutral contexts compared to happy faces in fearful contexts, for P8 (neutral contexts = -6.17 μV; fearful contexts = -5.73 μV), $t(10) = 2.47$, $p < .05$. Therefore, amplitudes were increased for both fearful and happy facial expressions in neutral contexts.

The results that were obtained for fearful faces were not attributable to the effects of low-level features as was controlled for by using scrambled versions of the scenes. Although a significant difference was found, the effect was in the reversed direction. The N170 amplitude for fearful faces in scrambled fearful scenes was less negative (-5.63 μV) as compared fearful faces in scrambled neutral scenes (-6.25 μV), $t(10) = 2.92$, $p < .05$. However, the effects for happy faces in scrambled scenes were in a similar direction as the observed effects for intact scenes and could therefore have explained these results.

*VPP component*

Unlike the control group, VPP latency analyses showed no main effect for facial expression or context emotion. VPP amplitude analyses showed a main effect for context emotion, $F(2,9) = 3.79$, $p < .05$, which was qualified by an interaction with electrode position, $F(4,7) = 5.11$, $p < .05$. This interaction was explained by more positive amplitudes for faces in neutral contexts ($M = 5.59$ μV) than faces in happy contexts ($M = 5.12$ μV) and faces in fearful contexts ($M = 4.83$ μV) on electrode FCz, $t(10) = 3.52$, $p < .01$ and $t(10) = 3.01$, $p < .05$ respectively, and more positive amplitudes for faces in neutral contexts ($M = 5.35$ μV) and for faces in happy contexts ($M = 5.20$ μV) as compared to faces in fearful contexts ($M = 4.92$ μV) on electrode Cz, $t(10) = 2.04$, $p < .05$ and $t(10) = 2.34$, $p < .05$ respectively.

No main effects were found for facial expression.

## 5.4 Discussion

We investigated the recognition of facial expressions in emotional scenes for schizophrenia patients, and recorded ERPs to study whether the N170 is sensitive to facial expressions of emotion and the accompanying context emotions.

The behavioral data do not reveal significant context effects for schizophrenic patients, which contrast to the healthy control participants, who had faster response times to happy facial expression in happy and neutral contexts as compared to fearful contexts. The evaluation task showed that schizophrenics were capable to identify the category and rate the arousal of scenes.

The ERP data showed general effects of context in the increased N170 amplitudes for faces in scrambled scenes as compared to faces in intact scenes, which were significant only for right hemisphere electrodes. Although effects on the N170 often are also reflected in the VPP, this effect was not significant for the VPP. Most importantly, specific emotion effects were found, as it was found that N170 amplitudes were more negative to happy faces in happy and neutral contexts compared to happy faces fearful contexts. This effect was mirrored by an effect on the VPP, but was here not specific for happy faces. Planned comparisons showed that fearful facial expressions had larger N170 amplitudes when shown in fearful scenes as compared to happy scenes, but in contrast to control participants, fearful faces in fearful contexts were not increased over fearful faces in neutral contexts.

### *5.4.1 General context effect*

The general context effects were analyzed by comparison of faces in intact scenes with faces in scrambled scenes. The general context effects were similar for schizophrenia patients and healthy control participants. Similar to healthy control participants, schizophrenics had longer response times to faces in intact scenes as compared to faces in scrambled scenes.

The ERP data show that schizophrenics had, like control participants, smaller N170 amplitudes for faces in intact scenes compared to faces in scrambled scenes. This is consistent with earlier findings in which the N170 amplitude was smaller for faces that were accompanied by intact scenes as compared to grey background (chapter 2). Since the N170 amplitude is decreased when attention is distracted away from faces to non-face objects (Holmes et al., 2003), we previously suggested that intact meaningful scenes may impose an additional load on face processing. Consequently, the N170 may be smaller to faces in intact scenes (chapter 2 and 4). The effect a meaningful context imposes on encoding is even reflected in the N170 for faces that are tested subsequently for recognition memory (Galli et al., 2006). The longer response times for faces in intact contexts are consistent with the notion that attention was distracted by meaningful information in context. The difference in physical properties of the context stimuli (and the control stimuli) across studies suggests that the physical properties do not explain this difference found for intact and scrambled scenes (see chapter 2 and 4, Galli et al., 2006).

*5.4.2 Specific emotion effect of context*

The behavioral results show that the arousal ratings and categorization of emotional content of the scenes were similar for schizophrenics and controls, which confirms earlier findings (Bigelow et al., 2006; Hempel et al., 2005; Rockstroh et al., 2006). Schizophrenia patients were rather accurate at emotion categorization of facial expressions of fear and happiness. Error patterns were equal to healthy control participants. This result is inconsistent with several previous studies, in which schizophrenics were poor at recognition of facial expressions (Dougherty et al., 1974; Feinberg et al., 1986; but see Gur et al., 2002; Kerr & Neale, 1993; Kosmidis et al., 2007; Salem et al., 1996; Schneider et al., 2006). However, this inconsistency may be explained partly by the task difficulty. In the present task it was necessary to use a two-alternative forced choice task, because multiple response alternatives would have introduced excessive eye-movements and motor responses. The behavioral performance may therefore appear relatively intact (see also Gur et al., 2002), but deficits may be found when a more demanding tasks is used.

Different effects of context emotion were however found for the response times to facial expressions. The significant interaction effect between facial expression and context emotion for the control participants was not found for schizophrenics. In the control group, happy faces were faster recognized in intact happy and neutral contexts as compared to fearful contexts. These effects were not found for schizophrenics. This suggests that emotional salient context do not influence face processing on a behavioral level, which contrasts to the effects found for the ERPs results, in which contexts do affect face processing.

The N170 amplitudes for faces in fearful contexts were more negative than the N170 amplitudes for faces in happy contexts for healthy control participants. It was found that fearful facial expressions had larger N170 amplitudes on left hemisphere electrodes when shown in fearful scenes as compared to happy scenes, but in contrast to control participants, not when fearful facial expressions in fearful scenes were compared to fearful facial expressions in neutral scenes. The effects on left hemisphere electrodes is replicated from previous studies (Stekelenburg & de Gelder, 2004 and see chapter 2 and 4).

The results are inconsistent with earlier neuroimaging studies, which suggest that the processing of emotional stimuli is impaired in schizophrenia patients. Experiments that presented facial expressions (Campanella et al., 2006; Gur et al., 2002) and emotional scenes (Rockstroh et al., 2006; Takahashi et al., 2004; Taylor et al., 2005) found that schizophrenia patients did not discriminate between the different emotions, perhaps related to early stages of visual processing (Rockstroh et al., 2006). However, the differences that were obtained with these studies may relate to the different methods that were used, as the brain activation represented by ERP data is not necessarily similar to the activation measured by fMRI/MEG. Another important source of these differences may be the combination of faces and contexts that was used in the present study. The combination of faces and context may provide more ecologically valid information, which may have critical impact on the perception of facial expressions. The results suggest that schizophrenic patients are sensitive to emotions in context, but differ from healthy controls with respect to facial expressions of fear in neutral context. It is not entirely clear which factors explain these N170 amplitude shifts for neutral contexts, but certainly a number of factors may have contributed to these effects.

One of these factors is attention, as it has been shown that the degree of attention that was paid to faces may modify the N170 amplitude (Holmes et al., 2003). It has been shown that the N170 amplitude increases if attention is cued to faces instead of houses. The larger N170 amplitudes to fearful faces in fearful and neutral contexts as compared to happy contexts may therefore be explained by attentional factors. Experimental studies have already shown that attention is attracted faster by fear-relevant scenes (Blanchette, 2006; Öhman, Flykt et al., 2001), which may explain the increase of the N170 to facial expressions in fearful contexts compared to happy contexts. However, it is unclear how attentional factors may explain the increased N170 amplitude to neutral scenes. To our knowledge, no study has reported as yet increased attention for neutral stimuli.

Previous error pattern analysis of behavioral responses have shown that schizophrenic patients tend to misidentify neutral facial expressions as negatively valenced (Kohler et al., 2003). During an acute episode, schizophrenics are inclined to misinterpret neutral occurrences that are in fact personally irrelevant (Kohler et al., 2003). If clinically stable patients also impute negative valence to neutral contexts, the N170 amplitudes may have been increased for fearful faces in neutral contexts for schizophrenics because neutral contexts were not perceived as non-emotional, but were imputed with negative valence. In contrast to fearful scenes, which are congruent to fearful faces, and happy scenes, which are incongruent with fearful faces, neutral scenes are rather ambiguous. This is consistent with the proposal that schizophrenics may be more vigilant for social scenes depicting ambiguous rather than overtly threat information (Phillips, Drevets, Rauch, & Lane, 2003).

In addition to the question whether the ambiguity of scenes influences perception in schizophrenia, another question is whether effects of contexts on face processing generalize to other contexts that have been used in face recognition research in schizophrenia (de Gelder et al., 2006). In a multisensory experiment that had used emotional auditory stimuli (voice prosody, happy or sad), it was investigated how the emotional tone of voice affects the categorization of facial expressions. Compared to healthy control participants, it was found that the emotional tone of voice had a reduced impact on the categorization of facial expressions in schizophrenics (de Gelder et al., 2005). These results are consistent with the behavioral data in which we found that emotional scenes did not significantly affect the response times for categorizing facial expressions.

Another factor that is often discussed in relation to face recognition deficits in schizophrenia is antipsychotic medication. A systematic comparison of medicated and unmedicated schizophrenia patients has not yet been performed for face recognition studies. This comparison is complicated by the inherent differences between patients who can and cannot be withdrawn from medication. However, data suggest that emotion recognition is not affected by medication (Kerr & Neale, 1993; Salem et al., 1996). No firm conclusion can be drawn as to whether antipsychotic medication influences the N170 for faces. By using auditory and visual ERP paradigms in the study of schizophrenia patients and age-matched controls it was shown that the N1 to simple visual stimuli (i.e., target detection among plus and minus characters) did not change after treatment with medication (Ford et al., 1994). As this N1 was measured for simple visual stimuli, it cannot be predicted whether this result holds for the N170 for faces (see discussion N1 and N170 by Itier & Taylor, 2004b).

It is rather unlikely that the results for fearful faces are based on low-level features of the scenes. This was investigated by using scrambled versions of the scene stimuli. However, for happy facial expressions, significant differences were found for scrambled scenes. The scrambled versions of the scenes were carefully inspected in that no recognizable features of the scene were left after the scrambling procedure. It is well documented that the color features of stimuli may affect early stages of processing (Allison et al., 1993). Color may be important and for some scene stimuli an inherent property of the emotional salience of the stimulus. The present results may suggest that certain color elements influence the perception of facial expressions differently.

*5.4.3 Conclusions*

Previous work has shown that schizophrenia patients showed less (Herrmann et al., 2004) or no (Onitsuka et al., 2006) distinctive increase of the N170 for faces over non-face objects. In addition, schizophrenics showed no sensitivity of the N170 to facial expressions (Campanella et al., 2006). The present data have shown that schizophrenics are sensitive to emotional information that is shown in the context of faces. This sensitivity suggests that at a certain stage of processing emotional perception is intact.

The results show that contexts influence face processing in schizophrenia patients. First, general context effects, resembling the effects that were observed in the healthy control participants, were found. The behavioral data showed increased response times for faces in intact scenes as compared to faces in scrambled scenes. The ERP data showed smaller N170/VPP amplitudes for faces in intact scenes than scrambled scenes. Second, specific emotion effects for contexts were found on the N170. Schizophrenia patients were, unlike healthy control participants, not sensitive to the difference of presenting fearful facial expressions in fearful scenes compared to fearful facial expressions in neutral scenes. A difference was, however, observed on the N170 between fearful facial expressions in fearful contexts compared to fearful facial expressions in happy contexts.

Acknowledgement

We thank W.A.C. van de Riet for helping with the Matlab procedures for scrambling scene stimuli.

# Chapter 6

# Recognition memory for faces encoded in emotional contexts

# An ERP study

*Abstract*

Emotional events are strongly remembered but do we also remember the persons we met during these events? Memory for faces and scenes has been studied relatively isolated. The present study centers on the question how recognition memory for a face is related to the emotional context in which it had been encoded. Neutral faces were presented in fearful and neutral contexts in a study-phase, and subsequently, in the test-phase, these faces were presented isolated from context along with novel faces in an old/new discrimination task. ERPs were recorded during both study-phase and test-phase. In the study-phase it was found that the N170 was increased for faces that were encoded in a fearful context as compared to a neutral context. As an effect of recognition memory, it was found that the N170 was less negative for hits than correct rejections for intact scenes (experiment 1) but not for scrambled scenes (experiment 2). The long-latency ERP component (LPC) showed an increased positive amplitude from ~400 ms onwards for faces that were recognized correctly as old (Hits). Additional effects were found as function of the context emotion, as from ~600 ms onwards correctly recognized faces that were studied in a fearful context introduced smaller parietal positivity than faces that had been studied in a neutral context. The results suggest that multiple stages of face processing are modified by the accompanying context during face encoding and recognition memory.

This chapter will be submitted as: Righart, R., & de Gelder, B. (2007). Recognition memory for faces in emotional contexts. An ERP study.

## 6.1 Introduction

Recognition memory for others may be profoundly affected by emotional factors like the context in which we saw others. Cases of eyewitness memory are for example one of the most intriguing, but also controversial, issues in which memory for others is affected by emotional salient events (see Christianson, 1992). Experimental work in humans and animals has shown that emotional salient events affect memory (McGaugh, 2003). A number of researchers have shown better memory for negative and positive valenced scene pictures (Bradley, Greenwald, Petry, & Lang, 1992; Calvo & Lang, 2005; Hamann, Ely, Grafton, & Kilts, 1999), video-clips (Cahill et al., 1996), and objects that were encoded in positive contexts (Smith, Dolan et al., 2004). Although emotional memory has been studied for scenes, and for objects encoded in scenes, it is not known how recognition memory for faces is affected by emotional scenes.

Recognition memory for various stimuli (e.g. words, faces) is influenced by the level of processing during encoding (Craik & Lockhart, 1972). Task conditions may affect the encoding strategy and, as a consequence, recognition memory for faces (e.g., rating faces for likeability, friendliness, Beales & Parkin, 1984; Bower & Karlin, 1974; Memon & Bruce, 1983; Patterson & Baddeley, 1977). In addition, the contexts may also influence the depth of encoding. A number of behavioral studies have investigated context effects by presenting them in both study- and test-phase. It has been shown that face recognition is better for faces in which the context was similar across study- and test-phase (Beales & Parkin, 1984; Davies & Milne, 1982; Winograd & Rivers-Bulkeley, 1977), a finding that has also been reported for objects (Tsivilis et al., 2001). These studies have progressed our understanding of how contexts affect face recognition performance. However, the role contexts fulfill during the study-phase and test-phase separately cannot be answered as yet. We have examined this question for the study-phase by presenting neutral faces in different categories of emotional salient scenes, which are present only during the first time the face is encoded (i.e., the study-phase).

On the one hand, it may be predicted that emotional contexts affect the depth of face encoding negatively, because emotional salient scenes may distract attention away from the target stimulus (Calvo & Lang, 2005). It has been shown in a number of studies that attentional modulation may affect subsequent face recognition (Jenkins et al., 2005; Reinitz et al., 1994). On the other hand, enhanced memory has been observed for emotional scenes in several experiments (Bradley et al., 1992; Calvo & Lang, 2005; Hamann et al., 1999). Emotional scenes may enhance face recognition memory if faces are perceived as an integral part of the scene.

To investigate this question we used event-related potentials in order to acquire ERP correlates of both encoding and retrieval. The P1 and N170 component have been associated with stages of face encoding. The P1 occurs at ~100 ms after stimulus onset and is the first positive deflection after stimulus onset (Hillyard & Anllo-Vento, 1998). The N170 occurs at ~170 ms and is the first negative deflection after the P1 component, and is maximal at occipito-temporal electrodes (Bentin et al., 1996). Both components are increased for faces as compared to objects (Bentin et al., 1996; Itier & Taylor, 2004b). The N170 is also increased for faces that are encoded in the context of fearful scenes relative to neutral (chapter 2) or happy scenes (chapter 4). However, it is unclear whether this increase is related to emotion perception *perse* or whether it is a correlate of encoding identity in an emotion-inducing context. Previous studies have found that the N170 amplitude increases for facial expressions of emotion, eye gaze direction and biological motion (Batty & Taylor, 2003; Caharel et al., 2005; Puce et al., 2000; Stekelenburg & de Gelder, 2004; Williams et al., 2006), and thus the N170 may also relate to the perception of socially important stimuli (Allison et al., 2000). However, the increase of the N170 amplitude may also relate to enhanced encoding of identity. However, to our knowledge, a correlation between the N170 amplitude and subsequent recognition memory has not been reported as yet.

Several ERP components have been studied for the retrieval of facial identity from memory. For the N170, face repetition effects have been found, as the N170 amplitude decreased to repeated presentations of facial identity (Heisz et al., 2006; Itier & Taylor, 2004a). In contrast, others did not find any effects of repetition for unfamiliar faces (Schweinberger et al., 2002) and famous faces (Bentin & Deouell, 2000; Henson et al., 2003; Schweinberger et al., 2002). More consistent results have been observed for the long latency ERP components that have been related to outcome measures of recognition memory for faces. A positive going amplitude shift from ~400 ms onwards, lasting several hundreds of milliseconds, has traditionally been found for various stimulus materials that are correctly recognized (pictures Curran, 2004; words: Rugg et al., 1998; objects in contexts Tsivilis et al., 2001). A similar positive-going ERP has been related to recognition memory for faces (Hannula, Federmeier, & Cohen, 2006; Itier & Taylor, 2004a; Johansson et al., 2004; Joyce & Kutas, 2005; Paller et al., 2000; Schweinberger, 1995; Schweinberger et al., 2002; Yovel & Paller, 2004). Hereafter, we refer to the term late positive complex (LPC), which has been used as a label for the positive-going ERPs in this time-course (Hannula et al., 2006; Joyce & Kutas, 2005).

The positivity of the LPC that is measured during the retrieval of the facial identity from memory is affected by the association that had originally been established between face and context during the study-phase. For instance, the association of biographical information to the presentation of a face, by pairing faces to names (Paller et al., 2000) or faces to occupations (Yovel & Paller, 2004) increased the LPC for faces that were successfully retrieved during the test-phase. The positivity of the LPC provides therefore a well-defined measure of how subsequent memory for faces is influenced by the context that accompanied the face during the study-phase.

In addition, it has been shown that emotion conveyed by face and context separately can modify the positivity of the LPC. Correctly recognized negative facial expressions increase the positive amplitude of the LPC relative to both positive and neutral expressions (Johansson et al., 2004). The LPC that was acquired for correctly recognized objects during the test-phase was sensitive to the emotional context in which the object had been encoded, as increased LPC amplitudes were found for objects that were encoded in emotional scenes (Smith, Dolan et al., 2004). In previous work, increased LPC amplitudes had already been found for emotional arousing scenes (Dolcos & Cabeza, 2002), but the study by Smith, Dolan et al. (2004) suggest that these effects are also found for the emotional scenes that are retrieved from memory.

If faces are presented in emotional scenes during the study-phase, on the one hand, it may be predicted that emotional scenes attract attention away from face processing. As attentional factors during encoding (Curran, 2004) and the subsequent quality of recollection (Wilding, 2000) both contribute to the amplitude of the LPC, the positivity of the ERP amplitude may decrease when faces were encoded in negative contexts. On the other hand, if emotional scenes enhance face encoding, or if emotional scenes are associated with the faces during retrieval, the LPC amplitudes may increase (see for similar account for objects, Smith, Dolan et al., 2004).

In the present study we investigated recognition memory for faces by using a study-test paradigm. In the study-phase, faces were accompanied by an emotional or a neutral context and participants were instructed to memorize the faces. In the test-phase, faces were presented without any context information, and participants performed an old/new discrimination task. An important goal of the present study was to investigate how context information that accompanies faces during the study-phase influences subsequent recognition memory. We hypothesized that the N170 would increase for faces encoded in emotional contexts compared to faces in neutral contexts, as has previously been reported (chapter 2). It was further hypothesized that if faces are distracted by emotional relative to neutral contexts, a less positive-going LPC would be recorded to correctly recognized faces that had been encoded in emotional scenes relative to neutral scenes.

## 6.2 Experiment 1

### *6.2.1 Method*

#### *6.2.1.1 Participants*

A total of 24 participants (9 males) all right handed (mean age = 22.2 years) participated in the study and had normal or corrected-to-normal vision. None of the participants reported a neurological history. Informed consent was given and the study was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki. Three participants were removed from analyses because their overall accuracy for recognition memory was very low (below 60% correct recognition of old faces). An additional three participants were removed from analyses because of artifacts and alpha-waves.

#### *6.2.1.2 Stimuli*

Faces were 591 color photos in frontal view found on the web and were unfamiliar faces (~90% caucasian). To prevent that the sex of the face stimuli would influence face recognition, only male faces were used. Faces were validated for facial expression (7 point scale: 1. very negative; 2. negative; 3. slightly negative; 4. neutral; 5. Slightly positive; 6. Positive; 7. Very positive). Only faces that were evaluated as neutral or slightly positive or slightly negative by at least 80% of the participants were selected for the experiments. After this validation, 320 face stimuli were selected for the experiment. A large number of faces was used in order to collect a sufficient number of trials to measure the P1 and the N170 for encoding in the study-phase and for hits and correct rejections during the test-phase. Original photo backgrounds and shoulders were removed and faces were centered on a grey rectangular background. Faces containing specific cues that influence recognition memory unintentionally were removed (facial hair, scars, shades or paraphernalia like earrings, spectacles) were removed.

Fearful and neutral scene contexts were color photos taken from the International Affective Picture System (IAPS, Lang et al., 1999) and the web, and they were validated for emotion. In a pilot experiment, fifteen participants were required to categorize each scene for emotional content (i.e., neutral or fear-, happy-, disgust inducing). Happiness and disgust evoking scenes were added as fillers in the validation but not used in the experiment. After the validation, 72 scenes were selected and figured as context of faces in the study-phase; half of the scenes were fearful, half were neutral. The contents of fear-inducing scene categories were collapsed building, fire, volcano, carcrash, flood, planecrash, storm, military tank, tornado. Non-emotional scene categories were car, van, highway, house, living room, mill, street, train, lamppost. In each of these categories, four scene pictures were selected. On average, participants correctly identified neutral scenes as being neutral in 82.73% of the presented scenes, which differs significantly from the percentage of responses that were given to fear ($M$ = 1.30%), disgust ($M$ = 2.24%) and happiness ($M$ = 13.73%), $F(3,12) = 58.00$, $p < .001$ (t-tests comparisons with response to neutral, all $p < .05$). Fearful scenes were correctly identified in 59.6% of the presented scenes, which differs significantly from the percentage of responses that were given to neutral ($M$ = 28.33%), disgust ($M$ = 6.11%) and happiness (5.93%), $F(3,12) = 450.60$, $p < .001$ (t-tests comparisons with response to fearful, all $p < .05$).

The height and width of the face was 8.3 cm x 5.7 cm (5.9° x 4.1°), and were overlaid on the center of scenes that measured 24.5 cm x 32 cm respectively (17.4° x 22.6°) (Figure 6.1). Faces did not occlude the critical information in the context picture.

*6.2.1.3 Procedure*

Participants were seated in a darkened electrically shielded cabin with a monitor positioned on a distance of 80 cm. They were instructed and familiarized to the experiment by practice sessions.

The total of 320 faces was divided into 20 sets of 16 faces. For each participant, half of the face sets served as old faces, the other half as new faces. The old faces were counterbalanced across the fearful and neutral contexts for one half of the participants, and served as new faces for the other half of participants, and vice versa for the new faces. Response buttons for old/new discrimination in the test phase were counterbalanced across participants as well. Context scenes were repeated across blocks but paired with novel faces.

The face-context experiment constituted of ten blocks all containing a study-phase and a test-phase. A test-phase directly followed its study-phase. Each novel block contained only faces that had not been used in preceding blocks. In the study phase, faces in context were presented randomly for 2000 ms, prompted by a fixation cross of 500 ms (Figure 6.1).

Each block contained 16 faces, each paired to a context. The contexts were from 18 different categories. Nine fearful contexts, the remaining were neutral contexts (i.e., after 10 blocks, 80 faces were studied in fearful contexts, and 80 faces were studied in neutral contexts, 20 faces were presented as target but did not return in the tes-phase). The intertrial interval was randomized between 1200 and 1600 ms. A fixation cross was shown during this interval.

In the study-phase, participants were instructed to study the faces carefully, because faces would be tested in a subsequent recognition memory session. Attention was controlled for by adding a detection task. Two catchtrials (an asterisk centered on the face) were included in each study phase randomly. Participants were instructed to press a button each time a face contained an asterisk. In order to control for motor-related artifacts, participants were allowed to respond only after disappearance of the face-context stimulus.

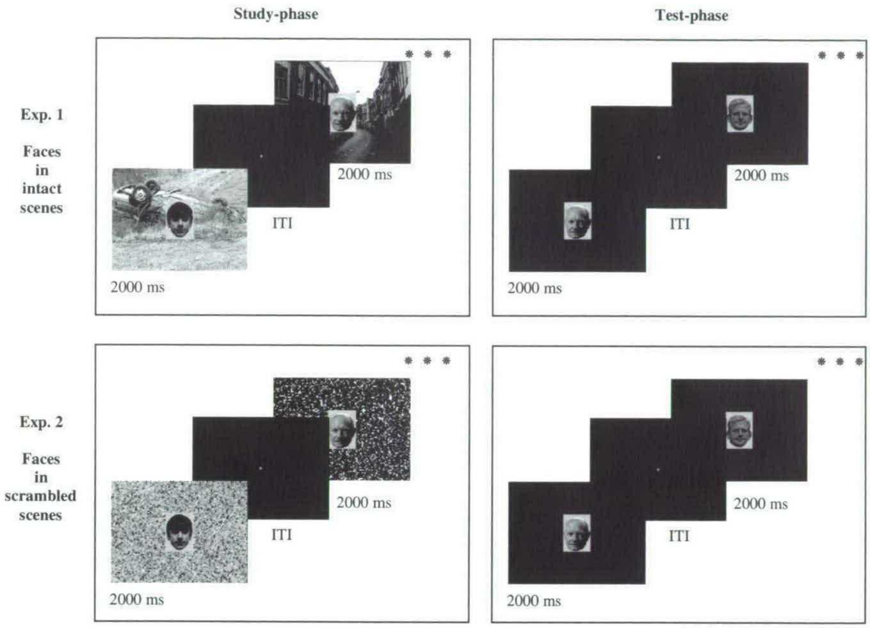



**Figure 6.1. Stimuli.** A stimulus example of the experimental design shows the study-test paradigm used for the face recognition memory study. In experiment 1, Faces overlaid natural scenes that could be neutral or fearful. Participants were instructed to memorize the faces and to detect an asterisk that was positioned in the center of some faces. In the test-phase, faces overlaid black backgrounds. Participants were instructed to indicate by pressing two buttons whether a face was shown before (old) or whether the face was not shown before (new). The intertrial interval (ITI) contains a black screen with a fixation cross in which participants were allowed to respond. The procedures were similar for experiment 2, in which a different group of participants memorized faces in scrambled versions of the same scene stimuli.

In the test-phase, the 16 old faces (faces containing catchtrials were excluded) and 16 new faces were presented on a black background for 2000 ms. Participants indicated whether they had seen the face before by using two buttons (yes/no). Again, responses were only allowed after disappearance of the stimulus. The intertrial interval was randomized between 1200 and 1600 ms

A practice session was used to familiarize participants to the task conditions of the main experiment. After this practice session, scenes that were used as contexts of faces in the main experiment were presented in advance of the main experiment in order to prevent novelty effects of the contexts. By using this procedure, great differences in performance between first and subsequent blocks could be avoided. Scenes were covered by a circle at the same position as the faces in the main experiment. The circle was used to cover information from scenes that was neither visible in the main experiment for faces in contexts.

*6.2.1.4 EEG recording*

EEG was recorded during both study-phase and test-phase, from 49 locations using active Ag-AgCl electrodes (Biosemi Active 2) mounted in an elastic cap. Four additional electrodes were used to monitor horizontal and vertical eye-movements. EEG was referenced to an additional active electrode (Common Mode Sense) during recording and was sampled at 512 Hz. All electrodes were off-line referenced to an average reference. The average reference has been used for both short-latency ERP components like the P1 and N170 (Itier & Taylor, 2004a; Schweinberger et al., 2002) and long-latency ERPs (Curran, 2004; Itier & Taylor, 2004a; Schweinberger et al., 2002). EEG signals were band-pass filtered (0.1 - 30 Hz, 24 dB/octave). Raw EEG data were segmented into epochs starting 200 ms before stimulus onset to 1000 ms after stimulus onset. EOG correction was performed by using the procedure of Gratton et al. (1983). Artifacts exceeding 100 μV at any channel were removed from analyses.

ERPs were averaged with respect to a 200 ms pre-stimulus baseline. Averages were made for the study-phase, for faces that were encoded in neutral contexts (NEU) and fearful contexts (FEAR), and for the test-phase, for faces that were correctly recognized (HITS) and correctly rejected (CR). In addition, separate averages were made for correctly recognized faces that were studied in neutral contexts (nHITS) and for correctly recognized faces that were studied in fearful contexts (fHITS). Catchtrials were excluded from analyses.

Peak latencies and amplitudes of the P1 were analyzed at occipital (O1/2) and parieto-occipital sites (PO3/4 and PO7/8) for the time window 80-130 ms (Hillyard & Anllo-Vento, 1998; Itier & Taylor, 2004b), and for the N170 at occipito-temporal sites (P5/6, P7/8, PO7/8, O1/2) for the time window 140-200 ms (Bentin et al., 1996; Itier & Taylor, 2004b). Mean amplitudes were analyzed at fronto-central, central, centro-parietal and parietal sites for consecutive time segments of 100 ms starting from 200 ms to 800 ms after stimulus-onset. Mean amplitude analyses in previous memory research has encompassed similar time courses as well (Johansson et al., 2004; Schweinberger et al., 2002; Smith, Dolan et al., 2004).

*6.2.1.5 Data analyses*

Accuracy for nHITS, fHITS and CR were analyzed for each participant. Only participants performing over 80% correct on the catchtrials and over 60% correct for recognition memory (HITS and CR) were analyzed. Recognition memory was analyzed by multivariate repeated measures ANOVA testing the within-subjects factor response type (fHITS, nHITS, CR).

In order to correct for false alarm rates, that is, incorrectly attributing a new item as being old, additional analyses were performed by using indexes of old-new discrimination (*Pr*) and response-bias (*Br*) as proposed by Snodgrass & Corwin (1988) [$Pr = p(\text{hit}) - p(\text{false alarm})$; $Br = p(\text{false alarm}) / p(1-Pr)$], and were also analyzed by ANOVA. Response times could not be analyzed because of the delayed response paradigm.

For ERP analyses, separate multivariate repeated measures of ANOVA were used for the study-phase and test-phase. The ANOVA for P1 analyses for the study-phase contained the within-subject factors context emotion (fearful, neutral), electrode position (O1/2, PO3/4, PO7/8) and hemisphere (left, right). The N170 analyses used similar factors but other electrode positions (P5/6, P7/8, PO7/8). For the test-phase, the P1 and N170 were analyzed for the factors response type (HIT, CR) and context emotion (nHIT, fHIT). The factors electrode position and hemisphere were similar to the factors analyzed for the study-phase. In addition, long latency ERPs were analyzed for each 100 ms time segment from 200 ms to 800 ms using electrode positions (fronto-central: FC3/z/4; central: C3/z/4; centro-parietal: CP3/z/4; Parietal: P3/z/4) and hemisphere (left, midline, right) as additional factors. Only the main effects or interactions of the factors response and context emotion will be reported. The selection of these electrodes and the time-segments of ERP analyses correspond closely with the segments chosen in previous studies (Johansson et al., 2004; Schweinberger et al., 2002; Smith, Dolan et al., 2004; Yovel & Paller, 2004). For each time-segment, separate ANOVAs were performed for response and context emotion. Subsidiary t-test analyses were performed for a larger selection of electrodes (Figure 6.4 and 6.5).

### *6.2.2 Behavioral results*

All participants performed accurately on the detection task that was presented in the study-phase (all participants >95% of catchtrials detected). A main effect was observed for response type, $F(2,19) = 4.03$, $p < .05$, which was explained by probabilities that were higher for faces that were correctly recognized as new ($CR = .85$, $SEM = .023$) compared to faces that were correctly recognized as old, for both faces that were studied in fearful contexts, $t(20) = 2.11$, $p < .05$ ($fHIT = .79$, $SEM = .016$), and for faces that were studied in neutral contexts, $t(20) = 2.83$, $p < .05$ ($nHIT = .79$, $SEM = .019$). No significant differences were found between correctly recognized faces that had been studied in fearful contexts and neutral contexts ($p > .05$).

Old-new discrimination (*Pr*) did not significantly differ between faces that had been studied in neutral context ($Pr = .64$, $SEM = .035$) and fearful context ($Pr = .65$, $SEM = .027$), $F(1, 20) = 0.11$, $p > .05$. Nor did the response-bias differ significantly between faces that had been studied neutral ($Br = .39$, $SEM = .037$) and fearful contexts ($Br = .39$, $SEM = .043$), $F(1, 20) = 0.08$, $p > .05$.

*6.2.3 ERP results*

On average 74 and 73 trials were left after artifact rejection for faces that had been studied in fearful and neutral contexts respectively. The number of trials left for correctly recognized faces that had been studied in neutral contexts was on average 56, and for correctly recognized faces that had been studied in fearful contexts 57. The number of trials left for correct rejections was on average 121 trials.

*6.2.3.1 Study-phase: face encoding in emotional contexts*

ERPs measured in the study-phase for faces that were encoded in fearful contexts were compared to faces that were encoded in neutral contexts. Sixteen participants showed a distinctive peak for the P1 component. P1 latency analyses did not show effects as a function of context emotion. It was shown that latencies differed as a function of electrode position, $F(1,15) = 10.23$, $p < .01$, as latencies were shortest for ERPs recorded from O1/2 ($M = 108$ ms) as compared to PO3/4 ($M = 112$ ms) and PO7/8 ($M = 115$ ms) (both comparisons $p < .01$). Peak amplitude analyses showed a main effect for hemisphere, $F(1,15) = 8.42$, $p < .05$, in that P1 amplitudes were larger for electrodes on the right hemisphere ($M = 11.40$ μV) than on the left hemisphere ($M = 9.10$ μV). A main effect was also found for context emotion, $F(1,15) = 5.39$, $p < .05$) as P1 amplitudes were smaller for faces studied in fearful contexts ($M = 9.97$ μV) than neutral contexts ($M = 10.53$ μV). However, these main effects were explained by a significant interaction between context emotion and hemisphere, $F(1,15) = 5.86$, $p < .05$). P1 amplitudes were smaller for faces studied in fearful contexts ($M = 8.68$ μV) than faces in neutral contexts ($M = 9.53$ μV) on the left hemisphere ($p < .05$), whereas the difference between faces studied in fearful contexts ($M = 11.26$ μV) and neutral contexts ($M = 11.53$ μV) did not differ significantly on the right hemisphere. Hemisphere in addition showed an interaction with electrode position, $F(2,14) = 5.86$, $p < .05$, in that amplitudes differed on the left hemisphere significantly between O1 and PO3 ($p < .05$), but on the right hemisphere between PO8 and O2 ($p < .05$), and PO8 and PO4 ($p < .01$).

Seventeen participants showed a distinctive peak for the N170 component. The N170 latency analyses did not show a significant effect as a function of the context emotion. A significant main effect was found for electrode position, $F(2,13) = 3.82$, $p < .05$), as latencies were shortest for electrodes P5/6 ($M = 161$ ms) compared to electrodes P7/8 ($M = 165$ ms) (comparison $p < .05$).

For the N170 amplitude analyses, importantly, a significant interaction was found between context emotion and hemisphere, $F(1,16) = 5.50$, $p < .05$), which was explained by amplitudes being more negative for faces in fearful contexts ($M = -1.78$ μV) compared to faces in neutral contexts ($M = -1.17$ μV) on the right hemisphere (comparison $p < .05$) but not between faces in fearful contexts ($M$ = -2.26 μV) and neutral contexts ($M$ = -2.21 μV) on the left hemisphere ($p > .05$) (Figure 6.2). A main effect for electrode position was found in addition, $F(2,15) = 67.73$, $p < .001$, showing that amplitudes were more negative to electrode pair P7/8 ($M$ = -3.52 μV) than PO7/8 ($M$ = -1.44 μV) and P5/6 ($M$ = -0.60 μV) ($p < .001$ and $p < .01$ respectively).

*6.2.3.2 Test-phase: Recognition memory*

ERPs were measured for hits and compared to correct rejections. No significant effects were observed for the P1 latency. Analyses for the P1 amplitude did not show a significant difference as a function of hits compared to correct rejection. A main effect was found for hemisphere, $F(1,15) = 5.85$, $p < .05$, showing larger amplitudes on the right hemisphere (9.59 μV) than left hemisphere (7.67 μV).

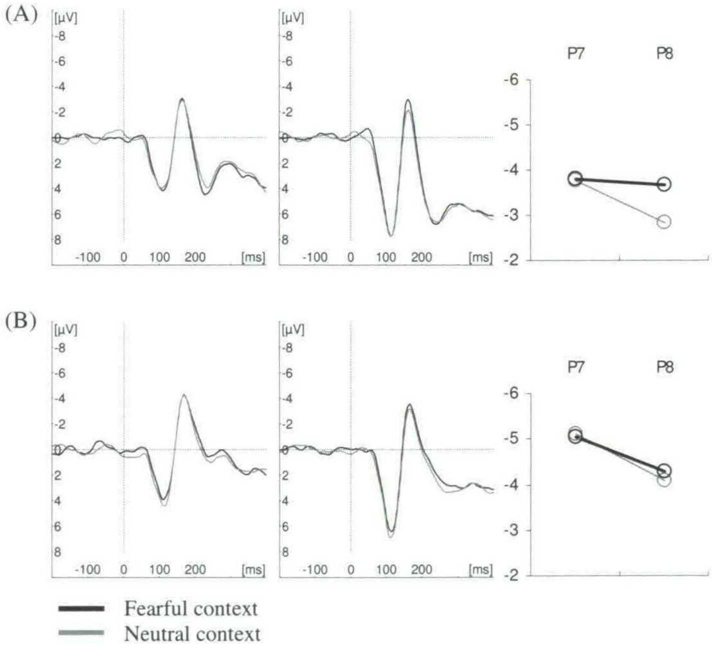



**Figure 6.2. Grand average ERP waveforms of the N170 for faces encoded in emotional scenes**. N170 in the study-phase (A) for faces that were encoded in a fearful context (red line) and in a neutral context (green line). Increased N170 amplitudes were observed for faces that were encoded in fearful contexts compared to neutral contexts on right hemisphere electrodes. N170 in the test-phase (B) for correctly recognized faces (hits) that were studied in fearful contexts (red line) and neutral contexts (green line).

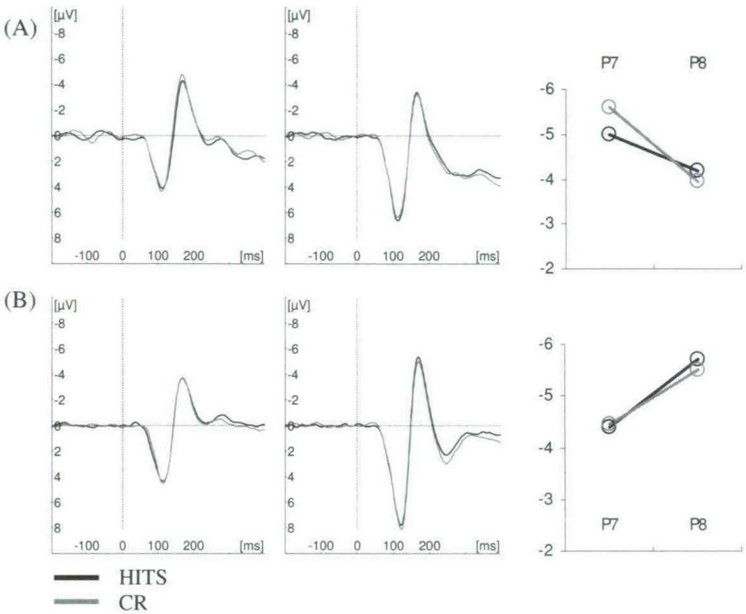



**Figure 6.3. Grand Average ERP waveforms of the N170 for old/new discrimination.** N170 in the test-phase for faces recognized as old or new. Faces were studied in intact scenes (A) and in scrambled scenes (B). The N170 peak is shown for faces correctly recognized (HITS, red) and correct rejected (CR, green). The N170 is larger for faces that were rejected as compared to faces that were recognized as old on left occipito-temporal electrode sites.

This main effect for hemisphere was qualified by an interaction with electrode position, $F(2,14) = 5.53$, $p < .05$, as left-right comparisons were significant for the comparisons of electrode PO7-PO8 only ($p < .05$). A main effect was in addition found for electrode position, $F(2,14) = 12.37$, $p < .001$, showing larger amplitudes for the P1 on both electrode sets O1/2 ($M = 9.58$ μV) and PO7/8 ($M = 9.50$ μV) than PO3/4 ($M = 6.82$ μV) (both comparisons $p < .01$).

No significant effects were found for the N170 latency. For the N170 amplitude, an important interaction between response type and hemisphere was found, $F(1,16) = 4.70$, $p < .05$, in that amplitudes were less negative for hits ($M = -3.67$ μV) than correct rejections ($M = -4.16$ μV) on left hemisphere electrodes ($p = .07$), but the difference between correct rejections ($M = -2.29$ μV) and hits ($M = -2.48$ μV) was not significant on right hemisphere electrodes ($p > .05$). In addition, N170 amplitudes for correct rejections tend to be more negative on left hemisphere electrodes than right hemisphere electrodes (comparison $p = .08$), whereas this amplitudes difference between hemispheres was not significant for hits ($p > .05$). A main effect was also observed for electrode position, showing largest amplitudes on P7/8, intermediate on PO7/8, and smallest on P5/6.

The long-latency analyses were performed for all participants and showed a main effect for response from 400 milliseconds onwards. Significant more positive-going amplitudes were found for hits than correct rejections (400-500 ms: $F(1,17) = 5.01$, $p < .05$; 500-600 ms: $F(1,17) = 12.40$, $p < .01$; 600-700 ms: $F(1,17) = 19.17$, $p < .001$; 700-800 ms: $F(1,17) = 10.89$, $p < .01$) (Figure 6.4). The interaction between response and hemisphere reached marginal significance for 500-600 ms ($F(2,16) = 2.78$, $p = .092$) and 600-700 ms ($F(2,16) = 3.40$, $p = .059$). The interactions for both time segments were explained by differences between hits and correct rejections being significant on midline electrodes (both $p < .001$) but not on left or right hemisphere electrodes (both $p > .05$).

*6.2.3.3 Test-phase: Recognition memory as a function of context emotion*

ERPs were measured for correctly recognized faces (HITS) that were encoded in fearful contexts (fHITS) and compared to correctly recognized faces that were encoded in neutral contexts (nHITS). No significant effects were observed for the P1 latency. For the P1 amplitude, no effect was found for hits as a function of the emotion of the context in which the faces were studied. A main effect was found for hemisphere, $F(1,15) = 6.66$, $p < .05$. This main effect was qualified by an interaction with electrode position, $F(2,14) = 6.40$, $p < 0.01$, showing that amplitudes were larger for right hemisphere electrodes than left hemisphere electrodes for pairwise comparisons of PO3-PO4 and PO7-PO8 but not for O1-O2. A main effect was also found for electrode position, showing larger amplitudes for electrodes O1/2 ($M = 9.67$ µV) and PO7/8 ($M = 9.57$ µV) than PO3/4 ($M = 6.95$ µV) (both $p < .001$).

The N170 latency showed a significant interaction between context emotion and hemisphere, $F(1,16) = 6.02$, $p < .05$, which was explained by latencies being shorter on the right hemisphere ($M = 165$ ms) than the left hemisphere ($M = 172$ ms) for faces that were studied in fearful contexts. This difference was not significant for neutral contexts (166 ms and 168 ms respectively). For the N170 amplitude, no effect was observed as function of the context emotion in which faces were studied. A main effect was observed for electrode position, showing the largest amplitudes for P7/8 ($M = -4.63$ µV), intermediate for PO7/8 ($M = -2.88$ µV), and smallest for P5/6 ($M = -1.93$ µV).

No significant effects were found for the long-latency ERPs as a function of context emotion in which faces had been studied (all $p > .05$). However, the pairwise t-tests depicted in Figure 6.5 show focal effects for CP5 (700-800 ms), P1 (700-800 ms), and Pz (600-800 ms), in that the ERP is less positive-going for faces that were studied in fearful contexts compared to neutral contexts.

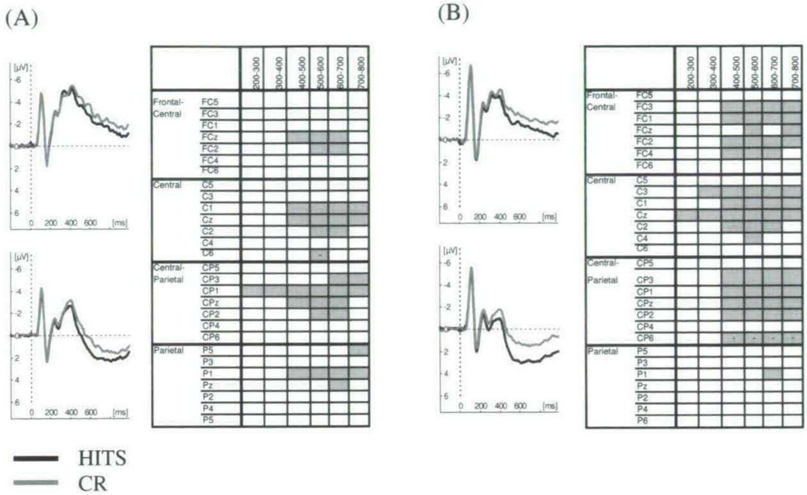



**Figure 6.4 ERP waveforms of recognition memory for faces encoded in intact scenes (A) and scrambled scenes (B).** ERPs at electrodes FCz (top) and Cz (bottom) to faces that were correctly recognized (HITS, red line) and correctly rejected (CR, green line). The ERPs were acquired at the test-phase after a study-phase in which faces were encoded in meaningful contexts. Mean amplitude analyses (grids) indicate time segments in which comparisons were significant by pairwise t-tests (gray: $p < 0.05$, uncorrected).

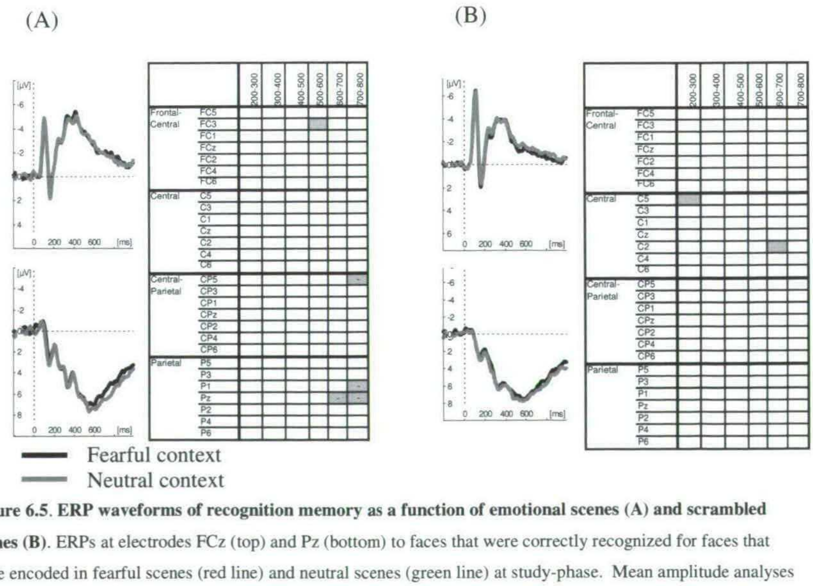



**Figure 6.5. ERP waveforms of recognition memory as a function of emotional scenes (A) and scrambled scenes (B).** ERPs at electrodes FCz (top) and Pz (bottom) to faces that were correctly recognized for faces that were encoded in fearful scenes (red line) and neutral scenes (green line) at study-phase. Mean amplitude analyses (grids) indicate time segments in which comparisons were significant by pairwise t-tests (gray: $p < 0.05$, uncorrected).

### *6.2.4 Discussion*

Neutral faces were studied in fearful and neutral contexts and were tested along with new faces in an old/new discrimination task. A number of important effects were observed for faces as a function of contexts in general, and for faces as a function of emotional contexts in specific.

For contexts in general, the most important findings were related to the recognition memory for the faces. First, the N170 amplitude was larger to new faces that were correctly rejected than old faces that were correctly recognized (i.e., being seen in the study-phase). Several factors complicate the interpretation of these results. Importantly, as the old faces differ from the new faces not only because they have been presented before, but also because they were accompanied by meaningful contexts, an important question is whether this effect is related to the semantic context information, possibly by associative memory for these contexts, or the depth of processing during encoding, or whether the effect is a correlate of memory for faces *per se*. The observed N170 amplitude difference is consistent with studies that have used a study-test paradigm (Itier & Taylor, 2004a) and repetition paradigm (Heisz et al., 2006).

Second, the positive-going ERP from 400 ms for faces that were correctly recognized as old compared to new was consistent with previous work (Johansson et al., 2004; Schweinberger, 1995; Yovel & Paller, 2004). The data are, however, as yet inconclusive whether there is a general effect of meaningful context information during the study-phase. In experiment 2, it is therefore tested whether these effects also occur for the same but scrambled versions of the fearful and neutral scenes.

Specific effects were found for context emotion during the study-phase and as an effect for correctly recognized faces in the test-phase. First, the P1 amplitude recorded in the study-phase was smaller to faces in fearful contexts than neutral contexts on left hemisphere electrodes. Further, the N170 was larger to faces in fearful contexts than neutral contexts on right occipito-temporal electrodes. These findings contrast to previous work, in which the differences were observed at electrode positions on the left hemisphere (chapter 2 and 4). The P1 effect was not found before. In fact, in work that investigated the neural responses to scenes without faces, it was observed that the P1 was larger to fearful than neutral contexts (Carretie et al., 2004; de Gelder et al., 2006; Smith et al., 2003). To exclude the possibility that these effects were caused by low-level features of the used scenes (e.g., color), in experiment 2 the same but scrambled contexts were used.

Second, late positive (LPC) memory-related ERPs recorded during the test-phase were also affected by the emotional scene that had accompanied faces during the study-phase. These long-latency ERPs have been related to recognition memory (Johansson et al., 2004; Schweinberger, 1995; Schweinberger et al., 2002; Yovel & Paller, 2004). It was found that the parietal positive-going LPC was less positive to correctly recognized faces that were studied in fearful contexts than neutral contexts. However, the effects were confined to a few electrode positions. Experiment 2 was used to rule out that low-level stimulus features had caused this effect.

## 6.3 Experiment 2

Faces were overlaid on meaningful scenes in experiment 1 and were compared to the same faces encoded in the same but scrambled contexts in experiment 2 (Figure 6.1). Scrambling contexts yields the informative content meaningless but keeps color properties intact. A between-subjects design was chosen because of the large number of trials needed to yield reliable P1 and N170 components. In addition, a problem of a within-subject design is that experience with scrambled contexts may result in carry-over effects as the color properties may prime the intact scenes (and this effect may also obtain for presentations of scrambled- and intact contexts in reverse order). As a consequence of scrambling contexts, it is expected that the specific effects observed for context emotion in experiment 1, will not be observed in experiment 2 anymore. However, the general effects of recognition memory for faces are still expected.

### *6.3.1 Method*

#### *6.3.1.1 Participants*

A total of sixteen participants (7 males) all right handed (mean age = 20.0 years) participated in the study and all had normal or corrected-to-normal vision. None of the participants reported a neurological history. Informed consent was given and the study was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki. Two participants were removed from the analyses because their overall accuracy for recognition memory was very low (below 60% correct recognition of old faces).

#### *6.3.1.2 Stimuli, procedure, EEG recording and data-analyses*

The same faces and contexts of experiment 1 were used, accept that scene pictures were scrambled. Scenes were scrambled by randomizing the position of the pixels across the image. Squares of 8 x 8 pixels were randomized across the image measuring 768 x 572 pixels width and height respectively, rendering the pictures meaningless. The resulting pictures were inspected carefully in order to avoid that recognition could be cued by certain features. All procedures and the settings for EEG recording were similar to experiment 1.
The within-subject factors that were used were similar to experiment 1. Between-group analyses were performed if appropriate.

*6.3.2 Behavioral results*

All participants performed accurately on the detection task that was presented in the study-phase (all participants >95% of catchtrials detected). The proportion of faces that were correctly recognized as old ($fHIT = .87$, $SEM = .018$; $nHIT = .88$, $SEM = .013$) was not significantly different from the proportion of faces that was correctly recognized as being new ($CR = 0.88$, $SEM = 0.013$), $F(2,26) = 1.05$, $p > .05$.

Old-new discrimination accuracy was not significantly different for faces that had been studied in neutral contexts ($Pr = .72$, SEM = .034) and fearful contexts ($Pr = .71$, SEM = .029), $F(1,13) = 0.66$, $p > .05$. Nor were significant differences observed at response bias measures for faces that were studied in neutral contexts ($Br = .53$, $SEM = .053$) and fearful contexts ($Br = .52$, $SEM = .52$), $F(1, 13) = 0.13$, $p > .05$).

Between-group analyses were performed between experiment 1 and 2. An interaction was found between the factor response type and group, $F(2,32) = 4.78$, $p < .05$. Correct recognition was also better for faces that had been studied in scrambled versions of both fearful and neutral contexts ($fHIT = .87$, $SEM = .018$; $nHIT = .88$, $SEM = .013$) as compared to intact fearful and neutral contexts ($fHIT = .79$, $SEM = .016$; $nHIT = .79$, $SEM = .019$), $t(33) = 3.14$, $p < .01$ and $t(33) = 3.60$, $p < .01$ respectively. Differences were not significant for correct rejections in experiment 1 ($CR = .85$, $SEM = .023$) and experiment 2 ($CR = .83$, $SEM = .031$). Similar patterns were found for old-new discrimination ($Pr$), but the difference between experiment 2 ($Pr = .71$, $SEM = .031$) and experiment 1 ($Pr = .64$, $SEM = .031$) did not reach significance, $F(1,33) = 2.32$, $p > .05$. However, the response bias in experiment 2 ($Br = .52$, $SEM = .057$) was also higher as compared to experiment 1 ($Br = .39$, $SEM = .039$), which reached marginal significance, $F(1,33) = 3.63$, $p = .07$.

*6.3.3 ERP results*

On average 77 and 75 trials were left after artifact rejection for faces that had been studied in fearful and neutral contexts respectively. The number of trials left for correctly recognized faces that had been studied in neutral contexts was on average 69, and for correctly recognized faces that had been studied in fearful contexts 67. The number of trials left for correct rejections was on average 128 trials.

*6.3.3.1 Study-phase: Face encoding in scrambled contexts*

Twelve participants showed a distinctive peak for the P1 component. For the P1 latency no effects were found for scrambled versions of context emotion. A marginal significant main effect was found for electrode position $F(2,10) = 3.63$, $p = 0.07$, as latencies were shortest on electrodes O1/2 (109 ms) compared to PO3/4 (117 ms) and PO7/8 (117 ms), $t(11) = 2.81$, $p < .05$ and $t(11) = 2.58$, $p < .05$ respectively. Analyses for the P1 amplitude did not show an effect of the scrambled versions of context emotion. A significant interaction between electrode position and hemisphere was found, $F(2,10) = 10.00$, $p < 0.01$, reflected by amplitudes that were larger on the right hemisphere than left hemisphere for electrode pairs PO3-PO4 and PO7-PO8, $t(11) = 2.43$, $p < .05$ and $t(11) = 2.58$, $p < .05$, but not for O1 and O2 ($p > 0.05$).

Thirteen participants showed a distinctive peak for the N170 component. Effects of faces in scrambled versions of emotional contexts were found neither for the N170 latency nor for the N170 amplitude. A main effect was found for the electrode position on the N170 amplitude, $F(2,11) = 21.14$, $p < 0.001$, in that amplitudes were largest on P7/8, intermediate PO7/8, and smallest on P5/6. Differences were not significant between P5/6 and PO7/8. An interaction was found between hemisphere and electrode position, $F(2,11) = 7.27$, $p < 0.05$.

*6.3.3.2 Test-phase: recognition memory*

No effects were found for the P1 latency and amplitude as a function of hits compared to correct rejections. A main effect was found for hemisphere, $F(1,11) = 11.40$, $p < 0.01$, which was qualified by an interaction with electrode position, $F(2,10) = 14.06$, $p < 0.001$, showing significant higher P1 amplitudes for electrodes on the right hemisphere than the left hemisphere for electrode PO3-PO4 and PO7-PO8 but not for O1-O2.

No effects were found for the N170 latency. For the N170 amplitude an interaction was found between electrode position and hemisphere, $F(2,11) = 10.87$, $p < 0.01$. This interaction was explained by N170 amplitudes being larger for P7/8 than PO7/8 and P5/6, and that these differences were more pronounced on the right hemisphere. In experiment 1, an important interaction was found between response type and hemisphere (Figure 6.3). This interaction was not found in experiment 2, $F(1,12) < 1$, $p > 0.05$).

The LPC analyses showed a main effect for ERPs from 400 ms onwards, reflected in hits being more positive-going than correct rejections (400-500 ms: $F(1,12) = 11.96$, $p < 0.01$; 500-600 ms: $F(1,12) = 25.21$, $p < 0.001$; 600-700 ms: $F(1,12) = 23.59$, $p < 0.001$; 700-800 ms: $F(1,12) = 15.67$, $p < 0.01$). An interaction between response type and hemisphere was found from 300 ms onwards, in that amplitudes were more positive on electrodes on the left hemisphere and the midline electrodes but not the right hemisphere electrodes for hits than correct rejections (300-400 ms: $F(2,11) = 5.22$, $p < 0.05$; 400-500 ms: $F(2,11) = 11.00$, $p < 0.01$; 500-600 ms: $F(2,11) = 7.84$, $p < 0.01$; 600-700 ms: $F(2,11) = 3.56$, $p = 0.064$), reaching marginal significance for the last two time-segments (700-800 ms: $F(2,11) = 3.92$, $p = 0.052$) (see Figure 6.4). An interaction was also found between response type and electrode position (i.e., anterior-posterior) from 400 ms onwards, in that amplitudes for hits were more positive than correct rejections for frontocentral, central, centroparietal, but not for parietal electrode positions (400-500 ms: $F(3,10) = 3.81$, $p < 0.05$; 500-600 ms: $F(3,10) = 5.73$, $p < 0.05$; 600-

700 ms: $F(3,10) = 4.34$, $p < 0.05$) and marginal significant for the segment 700-800 ms, $F(3,10) = 3.00$, $p = 0.082$.

*6.3.3.3 Test-phase: recognition memory as a function of context emotion*

No effects were found for the P1 latency as a function of scrambled versions of context emotions. A main effect was found for hemisphere, $F(1,11) = 13.67$, $p < 0.01$, which was qualified by an interaction with electrode position, $F(2,10) = 14.06$, $p < 0.01$, showing significant higher amplitudes for electrodes on the right hemisphere than the left hemisphere for electrode PO3-PO4 and PO7-PO8 but not for O1-O2.

Nor for the N170 latency neither for the N170 amplitude an effect was found as a function of scrambled versions of context emotion. For the N170 amplitude, a main effect was observed for electrode position, amplitudes being largest for P7/8, followed by P5/6 and PO7/8. This was qualified by an interaction with hemisphere, $F(2,11) = 10.96$, $p < 0.01$, and the direction of the effect was similar to the effects reported for HITS-CR in the previous section.

The long-latency ERP analyses showed a main effect for scrambled versions of context emotion in the time segment 300-400 ms, $F(1,12) = 5.10$, $p < 0.05$), as reflected by a more positive-going amplitude for correctly recognized faces that had been encoded in scrambled neutral than scrambled fearful context.

*6.3.4 Discussion*

In order to control for physical properties (luminance, color) as a factor of the ERP differences obtained for faces encoded in fear-inducing versus neutral scenes, we tested with a different group of participants the same faces in scrambled versions of these scenes. An advantage of presenting scrambled scenes to another group of participants is that experience with color properties of intact images does not prime processing of the scrambled ones in different blocks (and the other way around). It is known that color properties in scenes contribute to early stages of visual encoding (Spence et al., 2006). The second experiment showed that these properties do not explain the differences between face encoding in fearful and neutral scenes as observed for the N170 in experiment 1, and this finding is consistent with previous ERP data for intact and scrambled scene stimuli (chapter 4). In addition, it is unlikely that the divergence on the LPC (600-800 ms) between correctly recognized faces that were studied in fearful contexts and neutral contexts is explained by color features.

## 6.4 General Discussion

The present study investigated the neural correlates of encoding unfamiliar faces in emotional salient contexts, and subsequent face recognition memory. A study-test paradigm was used in which participants were required to memorize faces. Several short-latency (P1 and N170) and long-latency ERPs (LPC) were modified by the presence of meaningful scenes in general, and emotional scenes in specific. The results of the second experiment suggest that it is unlikely that these effects are caused by low-level features of the stimuli (e.g., color properties). The effects of each ERP component will be discussed for study-phase and test-phase separately.

### *6.4.1 Study-phase*

For the study-phase, important results were obtained for the P1 and the N170. It was observed that the P1 amplitude was smaller for faces that were studied in a fearful context compared to a neutral context. This finding is inconsistent with previous studies (Carretie et al., 2004; Smith et al., 2003) and is inconsistent with the findings in chapter 2 and 4. However, the task conditions may explain why different results were obtained, as participants were required to memorize the face. The N170 amplitude was increased for faces that were studied in a fearful context as compared to a neutral context. The differences were, however, found on electrodes on the right hemisphere. This contrasts to our previous work in which we observed larger amplitudes on left hemispheric electrodes (chapter 2 and 4). In previous work, facial expressions were used instead of neutral faces. These facial expressions could convey an emotion similar to the scenes, and different from the present study, an associative value could be perceived between the faces and the scenes. In addition, task differences across studies may have caused these asymmetries as well. It should be noted that lateralization effects for the N170 amplitude for facial expressions varies across studies (bilateral effects: Batty & Taylor, 2003; left hemisphere effects: chapter 2 and 4 and Stekelenburg & de Gelder, 2004) (Caharel et al., 2005; right hemisphere effects: Williams et al., 2006). The question which factors determine these lateralization effects is not resolved as yet.

### *6.4.2 Test-phase*

During the test-phase, the N170 for correct rejections showed more negative amplitudes on the left hemisphere electrodes than correct recognized faces. To our knowledge, this is the first study that has observed these effects by using a study-test paradigm. In a few studies, more negative N170 amplitudes have been found for new faces as compared to old faces (see Heisz et al., 2006; Itier & Taylor, 2004a), while in other studies no significant changes of the N170 amplitude were observed for unfamiliar faces (Schweinberger et al., 2002) and famous faces (Bentin & Deouell, 2000; Henson et al., 2003; Schweinberger et al., 2002). Interestingly, no effects were observed on the N170 as a function of recognition of old versus new stimuli in experiment 2, which is dissimilar to the findings in experiment 1. Several factors may have contributed to this effect.

First, intact scenes provide, in addition to the facial identity, meaningful information during face encoding. In the test-phase, in which no context was physically presented anymore, the contextual association that was formed with the face during encoding may generate differential activation during retrieval (Tulving & Thomson, 1973). This association is not formed for faces in scrambled scenes.

A second factor that may account for the effects is attention. Attentional differences during encoding in experiment 1 (faces in intact scenes) and 2 (faces in scrambled scenes) may have caused differential encoding, because intact scenes compete for attention (see for attentional effects on N170 Holmes et al., 2003). However, attentional distraction by intact scenes during encoding cannot directly explain the effects observed during the test-phase.

Thirdly, and perhaps related to the former point, the perceptual load that is introduced by intact scenes during encoding in the study-phase may have influenced the N170 amplitude for old faces compared to new faces in the test-phase. The intact scenes may have influenced the perceptual load because the information that is provided the face and the context exceeds the capacity for processing (see Yi et al., 2004). For this reason, the intact contexts may affect the depth of processing during encoding (Craik & Lockhart, 1972), although a firm conclusion cannot be drawn on this issue because it is unclear what the N170 amplitude difference between correctly recognize old and new faces means.

It should be noted that the data are inconclusive as to whether N170 amplitudes for new faces that were correctly rejected increased the negativity of the N170, or whether the N170 amplitudes for correctly recognized old faces decreased negativity of the N170. A between-subject comparison is inappropriate here because of potential signal-to-noise ratio differences between the groups.

We also analyzed the N170 in the test-phase for correctly recognized faces as a function of faces that were studied in emotional contexts, but no significant effects were observed. A recent study by Galli, Feurra, & Viggiano (2006) has found an effect of emotional context. Participants were presented with neutral faces that were superimposed on emotional negative and positive newspaper messages during the study-phase. During the test-phase, faces were framed on a white background. It was found that the N170 was decreased for correctly recognized faces that had been studied in a negative context. This suggests that context stimuli that were presented during the study-phase exert a differential effect on face recognition during the test-phase, possibly as an effect of face encoding during the study-phase or context retrieval during the test-phase.

For the LPC, positive-going amplitudes were replicated from previous studies that have investigated recognition memory for faces in repetition paradigms (i.e., experimental paradigm in which study- and test items are intermixed in a single experimental block: Heisz et al., 2006; Itier & Taylor, 2004a; Schweinberger, 1995; Schweinberger et al., 2002) and in study-test paradigms ( faces are repeated in a test-phase after a relatively long study-list: Johansson et al., 2004; Joyce & Kutas, 2005; Yovel & Paller, 2004). The left hemispheric effect found from 300 ms onwards may correspond to the usual left parietal effect found for recollection (Maratos et al., 2000; Smith, Dolan et al., 2004). This effect was found only in experiment 2 for faces that were encoded in scrambled contexts. The current paradigm could not dissociate between familiarity and recollection, which may explain the significant old/new discrimination effects across a wide range of electrodes (this issue will be discussed further below). It is interesting here that the effect was more widespread for faces that had been encoded in scrambled contexts (i.e., more electrodes showed significant differences as indicated by the number of gray grids in Figure 6.5). In addition, the behavioral data showed that accuracies for old faces were higher when studied in scrambled contexts (experiment 2) than intact contexts (experiment 1), although the effects disappeared when analyzed for response-bias. A tentative explanation is that this relates to the perceptual load during encoding. Face processing in intact contexts may correspond to a higher perceptual load than face processing in scrambled contexts, because meaningful contexts competed for attention in experiment 1 (chapter 2 and 4).

We observed differential LPC amplitudes between correctly recognized faces that had been encoded in fearful and neutral contexts. Amplitudes were less positive-going for faces that had been encoded in fearful contexts than neutral contexts, consistent with the study by Maratos et al. (2000), but inconsistent with other studies (Dolcos & Cabeza, 2002; Galli et al., 2006; Johansson et al., 2004; Smith, Dolan et al., 2004). Importantly, the effects were observed for faces in the test-phase in which they were shown without any context information around, which made reinstatement of contexts impossible. This is comparable to the study by Galli et al. (2006). However, Galli et al. (2006) used verbal contexts whereas the present study used natural scenes. Negative emotional scenes may have diverted attention away from face processing during the study-phase, although this effect was only significant on a few electrode positions (Figure 6.5) compared to the general recognition memory effects (Figure 6.4).

It should be noted that the used paradigm may have obscured differences between familiarity and recollection processes (Mandler, 1980; Yonelinas, 2002). Previous studies have observed a distinction between processes related to familiarity, showing an effect ~300-500 ms on mid-frontal sites, and recollection, showing a difference between ~400-800 ms maximal over left parietal sites (Johansson et al., 2004; Rugg et al., 1998), and an interesting issue is whether this distinction holds for faces in contexts. We removed the context information in the test-phase intentionally to confine the effects of recognition memory to the context emotion that accompanied faces during encoding in the study-phase. A number of studies suggest that context information that is present during retrieval greatly affects identification accuracies (Beales & Parkin, 1984; Davies & Milne, 1982; Memon & Bruce, 1983; Tulving & Thomson, 1973), which may be an intriguing issue to study further in behavioral and ERP experiments.

*Conclusions*

The present study sought to determine whether emotional information in context modifies memory for faces. A long tradition of research has shown that memory is modified for emotional events (McGaugh, 2003). It has been shown that memory is increased for emotional events (Cahill et al., 1996; Calvo & Lang, 2005). This seems not only the case for negative events, but also for positive events (Hamann et al., 1999). No behavioral effects of face recognition memory were found as an effect of emotional information in context in the present study. However, on electrophysiological measures during the study-phase, the N170, an ERP component related to encoding, was increased for faces in fearful contexts. Besides the usual positive-going ERP waveform for old-new discrimination, a less positive-going wave was found for faces that were studied in fearful compared to neutral contexts.

# Chapter 7

# Summary and Discussion

## 7.1 Summary of Results

Face processing was investigated as a function of the general presence of context information and as a function of the specific emotional content of the context by using behavioral and ERP methods. General effects of contexts on face processing were found by comparing intact scene contexts to grey backgrounds (chapter 2), and scrambled versions of scene contexts (chapter 4, 5). Behavioral data showed poorer performance for face recognition in intact scenes as compared to scrambled scenes. The N170 was smaller for faces in intact scenes as compared to faces in grey backgrounds or scrambled scenes. The effects may rely on attentional distraction by intact scenes or perceptual load of intact scenes, which will be discussed below.

Specific emotion effects for recognition of facial expressions were found as a function of emotional scenes. A categorization task for facial expressions was used (fearful, happy, disgusted), in which facial expressions were presented centrally in the context of emotional scenes. Emotional scenes could be congruent or incongruent to the facial expressions. Three behavioral experiments showed that the recognition of facial expressions (fear, happy, disgust) was influenced by emotional contexts, in that response times were faster to facial expressions that were accompanied by emotional congruent scenes as compared to incongruent scenes. This effect was found for short presentation-times (200 ms) and was resistant against the negative effects of a concurrent task. These results suggest that emotional scenes influence the speed of processing of facial expressions.

Event-related potentials (ERPs) of the N170 amplitude showed that emotional salient contexts influence face processing on an early stage of perception. Participants performed an orientation-decision task in chapter 2 and a categorization task in chapter 4. Facial expressions were positioned centrally in emotional salient scenes. In both experiments, effects were observed for emotional contexts, in that N170 amplitudes were increased for faces in fearful scenes as compared to neutral (chapter 2 and 4) and happy scenes (chapter 4), particularly for fearful faces on left occipito-temporal sites (chapter 2 and 4). Emotion processing was also investigated for schizophrenia patients by using the categorization task used in the behavioral experiments of chapter 3 and 4. Schizophrenia patients were like control participants sensitive to the emotional scenes, as the N170 amplitude was larger for fearful faces in fearful scenes as compared to fearful faces in happy scenes. However, they differed from the healthy control group in that N170 amplitudes were not larger for fearful faces in fearful contexts as compared to neutral contexts.

Similar effects of the emotional context on the N170 for face encoding were observed in a recognition memory task (chapter 6). A study-test paradigm was used in which faces were presented within emotional and neutral scenes. Participants were instructed to memorize the presented faces. In the test-phase, old and new faces were shown without any scene, and participants performed an old/new discrimination task. For the ERPs in the study-phase, the N170 amplitude was increased for neutral faces that were encoded in fearful scenes as compared to faces encoded in neutral scenes. The increases were found on right occipito-temporal electrode sites (chapter 6).

Recognition memory performance was not affected by the emotional contexts in which faces were encoded. For ERPs that were acquired in the test-phase, it was found that the N170 amplitude was smaller for hits as compared to correct rejections of faces. This difference was found for faces that had been encoded in intact scenes, but not for faces that had been encoded in scrambled scenes, which suggests that scenes had an important effect during face encoding (i.e., study-phase) or during retrieval of faces from memory (i.e., the test-phase), probably by the contextual association that was established with the presented facial identity. In addition, the LPC was more positive from ~400 ms onwards for correctly recognized faces compared to correct rejections, which was from ~600 ms onwards less positive for faces that were studied in negative compared to neutral contexts.

The results altogether suggest that the surrounding scene is an important factor in recognizing facial expressions, and that the effect occurs on an early stage of processing. The increase of the N170 amplitude to faces in fearful contexts was found irrespective of task conditions. The results will be discussed further in the following sections.

First, I will show how face and object recognition are affected by the semantic value of contexts. The effects of emotional salient contexts on face processing will be compared to the semantic effects that have been found for objects. The results will be discussed within the perspective of top-down modulation by contexts.

## 7.2 General effects of context

Numerous experimental studies have shown that faces can be detected fast, even if surrounded by meaningful context information (e.g., Lewis & Edmonds, 2003). It is as yet unknown how meaningful contexts affect face processing if compared to faces that are presented without any context information.

We found general effects of contexts across several experiments in which different stimulus-sets were used. The behavioral data showed slightly better performance for categorizing facial expressions in scrambled scenes as compared to facial expressions in intact scenes (chapter 4). Recognition memory was better for faces that had been encoded in scrambled scenes as compared to faces that had been encoded in intact scenes. However, although old-new discrimination (i.e., the proportion of hits corrected by false alarms, see Snodgrass & Corwin, 1988) was still higher for faces that were studied in scrambled scenes, the effects were not significant anymore (chapter 6). Previous studies have also reported better recognition for non-face objects that were isolated from scenes as compared to objects that were accompanied by scenes and are thus consistent with the observed recognition patterns for faces (Davenport & Potter, 2004; Ganis & Kutas, 2003; Murphy & Wisniewski, 1989).

The ERP data showed that P1 latencies were somewhat shorter for faces in scrambled contexts as compared to faces in intact contexts (chapter 4). The P1 amplitude was however inconsistent across studies. Whereas the P1 amplitude was marginally increased for faces without a context (i.e., faces presented on a gray background) as compared to faces in context in a study requiring participants to make an orientation-decision to the face (chapter 2), this effect was reversed in a study in which participants were asked to categorize facial expressions, as P1 amplitudes were decreased for faces in scrambled contexts as compared to faces in intact context (chapter 4). It should be noted, however, that these experiments did not only differ with respect to task conditions, but also the scene stimuli that were used (colored intact vs scrambled scenes in chapter 4, gray-scale intact vs gray contexts in chapter 2). Previous work has shown that the color properties of stimuli may also modulate the P1 amplitude (Allison et al., 1993).

The results were more consistent for the N170 component. The N170 amplitude was decreased for faces in intact scenes as compared to gray (chapter 2) and scrambled scenes (chapter 4, 5). It is unlikely that the observed effects are based on low-level features (color of scene), because different sets of face- and scene stimuli were used in the successive experiments. The data are consistent with previous studies (see for retrieval of faces Galli et al., 2006).

A number of factors may explain why the N170 amplitude was decreased for faces in intact scenes. One of these factors is the amount of attention that is dedicated to faces that are accompanied by meaningful contexts. If a meaningful scene is presented in the context of faces, this may distract attention from face processing. The effects of attention on the N170 amplitude for faces has been studied by instructing participant to direct their attention to two vertically aligned or two horizontally aligned positions. At these positions, faces or houses could be presented. It was found that the N170 amplitude was affected as to whether attention was cued to faces or houses, as the N170 amplitude was larger when faces were attended (Holmes et al., 2003). However, in contrast to these results, in another experiment it was found that the N170 was not affected as to whether faces were the target of attention (Cauquil et al., 2000).

Another factor that may explain the decreased N170 amplitude for faces in intact scenes is the perceptual load imposed by intact scenes. Lavie (1995; 2005) hypothesized that the load of the target-relevant stimulus (i.e., the face) determines the degree of processing of target-irrelevant information (i.e. the context). Therefore, it is conceived that if the face recognition task is relatively easy, context processing occurs optimally. In contrast, if the face recognition task is relatively difficult, context processing should be limitted. A firm conclusion regarding the perceptual load hypothesis cannot be drawn with the present ERP data, because a direct comparison between high and low load tasks was not made (see for an example, Yi et al., 2004).

In contrast to the general effects of contexts, for emotional salient contexts experiments were conducted to investigate the effect of high load on the processing of facial expression (chapter 3). Emotional salient contexts sorted an effect on facial expression recognition, even when a high task demand was imposed by using a concurrent task in which participant were required to categorize characters that were superimposed upon the facial expression (experiment 3 in chapter 3).

## 7.3 Semantic effects of context

### *7.3.1 Semantic effects of contexts in object recognition*

Semantic information can be extracted from scenes within a single glance (Loftus et al., 1983; Potter, 1975; Schyns & Oliva, 1994; Thorpe et al., 1996). Objects in the world around us are perceived rapidly, partly because of the rapid perception of the semantically related context that accompanies these objects. To study context effects experimentally, objects were presented that were accompanied by semantically congruent contexts, and in other trials with objects that were accompanied by incongruent contexts. An example of a congruent object-context combination is a loaf of bread in the context of a kitchen. In most experiments, participants are required to recognize the presented object. As the context generates expectations about what objects are likely to occur, it is hypothesized that congruent context facilitate object perception (Bar, 2004). Experimental studies have indeed found that participants recognize objects more accurately (Murphy & Wisniewski, 1989; Palmer, 1975; Biederman, Mezzanotte, & Rabinowitz, 1982; Boyce, Pollatsek, & Rayner, 1989; Davenport & Potter, 2004; Friedman, 1979; but see Henderson & Hollingworth, 1999; Hollingworth & Henderson, 1998) and faster (Ganis & Kutas, 2003) when they are presented in congruent contexts (e.g., loaf of bread in the kitchen) as compared to incongruent contexts (e.g., a loaf of bread in the bathroom). These studies show that object recognition is not independent of the accompanying context, but that the semantic relation between the object and its accompanying context predicts how objects are recognized.

fMRI studies have shown that such contextual associations evoke activity in the parahippocampal cortex (Bar & Aminoff, 2003; Goh et al., 2004) and the retrosplenial cortex (Bar & Aminoff, 2003). Activity in the parahippocampal cortex has been related to perception of places (Epstein & Kanwisher, 1998). However, fMRI activation in the parahippocampal cortex and retrosplenial cortex is not specific for places, but is also elicited by non-spatial contextual associations (Bar, 2004; Bar & Aminoff, 2003), and associations among meaningless visual patterns (Aminoff, Gronau, & Bar, 2006). This result puts forward the question whether the parahippocampal cortex also codes for the relation between faces and contexts (semantic / emotional associations, for instance persons that are usually seen in certain places). This issue will be discussed further in the section below about faces in contexts.

ERP studies have further found that the contextual association between objects and contexts relates to the negativity of the N400 component. The N400 has originally been studied in language paradigms, in which the negativity of the N400 relates to semantic violations between a target word and its sentential context. Thus, the negativity of the N400 amplitude increased if an unexpected word occurred in the sentence (Kutas & Hillyard, 1984). The negativity of the N400 amplitude was also increased for objects that were incongruent with the semantic expectancies provided by the context scene (Ganis & Kutas, 2003). The observed latencies for the N400 for objects show similar latencies and scalp topographies if compared to the language paradigms. Given this long-latency effects, it was proposed that the effects are based on a late stage of identification (a post-identification account by Ganis & Kutas, 2003; see also Henderson & Hollingworth, 1999) However, a number of researchers have proposed an early perception account for object identification in scenes (Bar, 2004; Schyns & Oliva, 1994), which will be discussed in section 7.3.3.

### *7.3.2 Semantic effects of contexts in face recognition*

ERP studies have also been concerned with face recognition memory and the relation between faces and the accompanying contexts. Traditionally, a positive-going LPC from ~400 ms onwards has been observed for faces that are correctly recognized (Johansson et al., 2004; Paller et al., 2000; Schweinberger, 1995; Schweinberger et al., 2002; Yovel & Paller, 2004). The effect of contexts on recognition memory has been studied in a number of ways.

First, the effect of contexts has been tested by presenting contexts during encoding (i.e., the study-phase) but not during retrieval (i.e., test-phase). We observed, similar to the aforementioned studies that presented faces without any context, a positive-going LPC for correctly recognized faces that had been encoded in scene contexts (chapter 6). These results agree with previous work that used a paradigm in which faces were studies in the context of newspaper messages (Galli et al., 2006).

Second, the effect of contexts has been investigated by modifying the semantic association between face and context. Yovel and Paller (2004) have investigated the well-known "the butcher-on-the-bus phenomenon". This phenomenon occurs when individuals realize that it takes more effort to recognize others in a different context. It was investigated how the LPC acquired during the test-phase for recognized faces related to the association that was established between faces and occupations during the study-phase. In the study-phase, faces were presented simultaneously with the occupation. Participants were instructed to memorize the face-occupation combination. In the test-phase, old faces were presented mixed with new faces and participants were required to indicate whether they had seen the faces before or whether the face was new. Correctly recognized faces including the associated occupation showed a more LPC from 500-700 ms than correctly recalled faces for which the associated occupation was not retrieved. This finding suggests that the context to which a face is associated changes how the face is retrieved from memory.

The aforementioned studies have applied a paradigm in which faces were placed in a certain context only during the study-phase, while the faces were isolated during the test-phase. Context-reinstatement may also be a factor of face recognition memory, in which the same contexts that were used during the study-phase are paired with the same old faces the participants had been exposed with. A number of studies have investigated face recognition memory as a function of context reinstatement. Individuals showed better memory for faces that were reinstated (Davies & Milne, 1982; Hannula et al., 2006; Memon & Bruce, 1983; Rainis, 2001).

Bruce and Young (1986) have proposed in their functional model of face recognition that these contextual effects occur on a stage in which semantic information for persons is processed, which is after the structural encoding stage (Bentin & Deouell, 2000). This proposal is consistent with results by Hannula et al. (2006), as they showed that faces that were tested in the same context compared to the context in the study-phase showed a smaller N400 negativity (see for objects Ganis & Kutas, 2003), and that the LPC amplitude for faces in the same contexts was increased over faces in new context (see for objects Tsivilis et al., 2001).

### *7.3.3 Top-down modulation in object recognition*

In addition to the late stages of processing the semantic relation between a target and its scene context, it has been shown that object processing is influenced on rather early stages of processing by its accompanying context (Bar, 2004; Schyns & Oliva, 1994; Torralba, Oliva, Castelhano, & Henderson, 2006). Bar (2004) has proposed a model of context processing which deals with the speed of object recognition. The global properties (the gist) of the scene are processed first, providing an initial representation of the scene by low spatial frequencies, followed by more detailed representations, carried by high spatial frequencies. The low spatial frequency representation of the scene is projected from early visual areas to the prefrontal cortex. The prefrontal cortex exerts top-down influence on object perception by limiting the possible object representations by providing an 'initial guess'. A recent magnetoencephalography (MEG) study suggests that such a top-down modulation indeed may be executed by the orbitofrontal cortex, showing earlier peak latencies (~130 ms) than the fusiform gyrus (~180 ms) (Bar et al., 2006). Future studies may show whether contexts that are congruent to the perceived objects show a similar kind of feedback projection between orbitofrontal cortex and fusiform gyrus.

*7.3.4 Top-down modulation in face recognition*

A study by Dolan et al. (1997) suggests that the segmentation of faces from contexts is facilitated by a priori knowledge. Two-tone images of a face (i.e., black-white) were shown in a pre-learning phase, in which the face was difficult to segment from its context. A gray-scale version of the image was shown in a learning phase, in which participants learnt that the two-tone image in fact represented a face in context. The same two-tone image was then presented in a post-learning phase. Participants showed a larger neural response in the fusiform gyrus to faces in the post-learning phase, which suggests that individuals were able to segment a face out of an impoverished context, by learning about the context (Dolan et al., 1997; see also contrast-reversed faces by George et al., 1999; and Rubin's vase-to-face illusion by Hasson, Hendler, Ben Bashat, & Malach, 2001).

A second example showing that individuals use context information to perceive degraded images of faces, is a study by Cox, Meyers & Sinha (2004). Faces and blurred faces were shown isolated or in the context of a body. A degraded face image in an appropriate context (i.e., a blurred face on a body) showed a higher response in the fusiform face area compared to when the degraded face was shown without this information. As neither the activation to the blurred face on its own nor the activation of the body on its own showed a similar response as the combination of face and body, this study suggests that faces are filled in at the place that was cued by the context. Consistent with this, behavioral data suggest that scene information cues attention to a place that likely contains the face (see for objects in scenes Biederman et al., 1982; Lewis & Edmonds, 2003).

It has been suggested that top-down modulation occurs on an early stage of object perception (Bar, 2004; Bar et al., 2006). For faces, the aforementioned studies by Dolan et al (1997) and Cox et al. (2004) suggest similarly early effects, although the activation of the fusiform gyrus found by fMRI could possibly also reflect late stages processing. However, ERP studies for face processing suggest that a priori knowledge about the presented face stimuli could modulate the N170 (Bentin & Golland, 2002; Bentin, Sagiv, Mecklinger, Friederici, & von Cramon, 2002), which is consistent with an early perception account. This is relevant for the effects we found by emotional context on face processing, which will be discussed in the next section.

## 7.4 Emotional effects of contexts

Faces may not only have a semantic association with the accompanying context (e.g., seeing the butcher in his butchers shop), but there may also be an important association if face and context evoke a similar emotion. Several important effects of emotional salient contexts on face processing were found. First, the N170 amplitude was increased for faces in fearful contexts compared to neutral (chapter 2 and 4) and happy contexts (chapter 4). Second, facial expressions were categorized faster when perceived in congruent scene contexts as compared to incongruent scene contexts. Consistent results were obtained across different task conditions that were used across the behavioral (presentation time, concurrent task) and ERP experiments (orientation-decision task, chapter 2; categorization task, chapter 4; memory task, chapter 6) which argues for the robustness of the effects.

Previous studies have treated the processing of facial expressions and emotional scenes as relatively separate domains of interest. Throughout the literature, however, the results show similar patterns for the study of facial expressions and emotional scenes. First, behavioral studies have observed early attentional capture for emotional scenes (Calvo & Lang, 2004, 2005; Öhman, Flykt et al., 2001) and facial expressions (Hansen & Hansen, 1988; Öhman, Lundqvist et al., 2001; Vuilleumier & Schwartz, 2001; White, 1995). Second, ERP studies have found that the P1 amplitude is modulated by the emotional content of scenes (Carretie et al., 2004; Smith et al., 2003) and facial expressions (Batty & Taylor, 2003; Eger et al., 2003; Meeren et al., 2005). The P1 has been related to attentional processing (Hillyard & Anllo-Vento, 1998; Mangun, 1995). However, to our knowledge, no effects have been reported for emotional scenes on the N1, whereas for facial expressions several studies have found effects on the N1 as reported in Table 1.2 (i.e., N170, see for discussion about N1/N170, Itier & Taylor, 2004). Third, another similarity has been found by fMRI studies, in which overlapping areas were found for facial expressions and emotional scenes in the amygdala (Britton, Taylor, Sudheimer, & Liberzon, 2006; Hariri et al., 2002). If emotions that are evoked by face stimuli and scene stimuli separately involve similar neural systems and evoke similar behavioral and electrophysiological responses, an additive effect may be observed when facial expressions are accompanied by emotional scenes.

### *7.4.1 Recognition of facial expressions preceded by contexts*

A number of experiments have studied the role of contexts in face processing by using storylines or images that preceded the presentation of facial expressions (see Table 1.1). It was found that emotion-inducing storylines read by the experimenter significantly modify how a subsequently presented facial expression is interpreted by the participant, for instance a story of not being helped in a restaurant (Carroll & Russell, 1996). Consistent with this, fMRI data of another study showed a greater amygdala response to ambiguous facial expressions of surprised faces that were preceded by a negative emotional sentence (about losing money) than a positive emotional sentence (Kim et al., 2004). Another way to create a context is by exposing participants to emotional pictures that precede the facial expressions. Contextual movie presentations have been applied in this way, by zooming into a static image for four seconds, followed by a facial expression that was evaluated on a negative-positive dimension by the participants (Mobbs et al., 2006). It was found that evaluations were more positive when facial expressions were presented after a positive context, whereas the evaluations were more negative when facial expressions were presented after a negative context. This study importantly showed that emotional images that preceded the facial expressions, significantly affected the pleasantness ratings of otherwise equal facial expressions.

### *7.4.2 Recognition of facial expressions presented within contexts*

When facial expressions are perceived, the facial expression is often a part of the person's interaction with the environment (Frijda & Tcherkassof, 1997), and therefore the occurrence of a facial expression on the one hand, and the occurrence of an emotional salient event that evoked the facial expression on the other hand, are closely related in time. We therefore have used similar onset times for face and context. Facial expressions (fear, disgust, happiness) were positioned centrally within emotional scenes. Participants were required to categorize facial expressions. It was found that facial expressions were better and faster categorized when they were perceived in congruent scenes compared to incongruent scenes (chapter 3). A number of findings are important in this study. First, the results do not show a plain negative-positive dichotomy, but show distinct effects for fear and disgust (fear being faster recognized in a fear evoking context, disgust being faster recognized in a disgust evoking context). Second, the effects were obtained with short presentation-times (200 ms), making possible only a single fixation. This finding that emotional scenes exert an effect even with short viewing times is consistent with previous work (Calvo & Lang, 2005). Given these short presentation-times, it is unlikely that this effect is based on local details in the scene. It rather suggests that the perception of the gist of the scene (Bar, 2004), the rapid extraction of the global meaning, was important for sorting these effects.

The result is consistent with other studies that have investigated the recognition of facial expressions in the context of auditory (de Gelder & Vroomen, 2000) and bodily expressions (Meeren et al., 2005; Van den Stock, Righart, & de Gelder, 2006; for a review see de Gelder et al., 2006). These studies have shown that participants performed better for congruent face-context pairs than incongruent pairs. This suggests that the interpretation of facial expressions importantly relies on the perception of diverse context backgrounds which we perceive by different senses. The effect of contexts may depend on the clarity of the facial expression that is perceived. The effect of context information is largest for ambiguous facial expressions, as was shown in different crossmodal studies that have paired facial expressions with emotional voices (de Gelder et al., 2006; de Gelder & Vroomen, 2000). It remains to be examined whether this extends to visual scene contexts (see Mobbs et al., 2006).

The faster response times to facial expressions that are perceived in congruent contexts (chapter 3) may have different functions for each emotion. Fast processing of facial expressions of fear in congruent contexts may optimize behavioral efforts to avoid dangerous situations (LeDoux, 1996). The perception of facial expressions of disgust may prevent contact with contaminated objects (Rozin & Fallon, 1987). Facial expressions of happiness signal that it is social rewarding to approach a situation (Keltner & Kring, 1998). The behavioral results make clear that perception is based on the interaction between facial expressions and context emotions.

The specific effects that facial expressions of emotions are processed faster when they are perceived in an emotional congruent context may be consistent with the proposal that separate neural systems are involved in emotion perception (Calder et al., 2001). The amygdala responds mainly to threat (LeDoux, 1996) and facial expressions of fear (Morris et al., 1998). The insula/basal ganglia system is mainly involved in processing disgust (Phillips et al., 1997). However, not all studies support this idea. Other studies suggest that these brain regions are involved for the processing of other emotions as well for (the amygdala: Britton et al., 2006; Fitzgerald et al., 2006; Hamann et al., 1999, the insula: Britton et al., 2006; Schienle et al., 2002). The activity in the amygdala also related to the intensity of the perceived facial expression (Morris et al., 1998). Increased amygdala activity was also found for facial expressions paired with fearful voices (Dolan, Morris, & de Gelder, 2001). An interesting question is whether the degree of activation in the aforementioned regions of interest corresponds to the intensity of the emotions conveyed by facial expressions in congruent contexts as opposed to incongruent contexts.

In addition to the aforementioned brain regions, contextual associations may also involve the parahippocampal cortex and retrosplenial cortex (Bar, 2004; Bar & Aminoff, 2003). Although these regions have previously been found to spatial relations between object and environment (Epstein & Kanwisher, 1998), it was found that the parahippocampal cortex is also involved in non-spatial contextual associations (e.g., confetti and champagne, Bar & Aminoff, 2003; Aminoff, Gronau, & Bar, 2006). Others have further found that these regions are sensitive to emotion. Smith, Henson, Dolan, & Rugg (2004) found increased parahippocampal cortex activation to neutral objects in emotional scenes as compared to neutral objects in neutral scenes. Sterpenich et al. (2006) acquired fMRI in the test-phase of a study-test recognition memory paradigm for faces that were encoded in aversive scenes during the study-phase, but were presented without scenes during the test-phase. They found increased activity of the parahippocampal cortex for successfully recognized faces that had been encoded in aversive contexts. It is unclear as yet whether this is related to different encoding of faces in emotional context, which has its effect on the activity acquired during the test-phase, or whether this activity is related to retrieval of face-context associations (similar effects are discussed in chapter 6 and by Smith, Dolan et al., 2004).

The retrosplenial cortex has been reported to be sensitive for emotional properties of stimuli as well (Maddock, 2000; Maddock, 1999), but see also (Vogt, Absher, & Bush, 2000). Increased activity in the retrosplenial cortex has been reported for fearful faces as contrasted with neutral faces, and in specific for fearful facial expressions characterized by low spatial frequencies as compared to high spatial frequencies (Vuilleumier, Armony, Driver, & Dolan, 2003). Thus, the aforementioned studies show that these areas may be of interest, not only in the study of objects in contexts (Bar, 2004), but also in the study of facial expressions in emotional contexts.

### *7.4.3 Top-down modulation of recognition for facial expressions*

The interaction between facial expression and emotional contexts appear to occur in a perceptual stage of processing. It was found that the N170 amplitude for faces in fearful contexts was larger than faces in neutral contexts. The N170 amplitude was additionally increased for fearful faces in fearful contexts compared to neutral contexts (chapter 2). An orientation-decision task was used for which emotion categorization was irrelevant for task performance. However, when participants were required to categorize the facial expression, similar increases of the N170 amplitude were found for faces in fearful contexts compared to faces in neutral and happy contexts. The results appeared most prominent for fearful facial expressions (chapter 4).

These findings are inconsistent with studies that suggest that the N170 amplitude is insensitive to facial expressions (See table 1.2). Thus, the interpretation that the N170 is related to the stage of structural encoding (Bentin & Deouell, 2000; Eimer et al., 2003) is not supported. Structural encoding is the encoding of the physical invariants of the face in order to match the perceived structure with the stored structural codes contained by the "face recognition units" (Bruce & Young, 1986). In fact, it seems that the N170 is sensitive to stimulus characteristics that are crucial for social perception (Allison et al., 2000). It has been shown that the N170 is not only affected by facial expressions (Batty & Taylor, 2003; Caharel et al., 2005; Stekelenburg & de Gelder, 2004; Williams et al., 2006), but also by expressional change (Miyoshi et al., 2004), social affective features (Pizzagalli et al., 2002), movement of eyes and mouth (Puce et al., 2000; Puce et al., 2003), direction of eye gaze (Watanabe et al., 2002), and biological motion (Jokisch et al., 2005). These results, together with the results obtained for facial expressions that were presented in emotional scenes, suggest that the N170 amplitude is not reducible to the perception of specific low-level features but is uniquely driven by emotional content.

The influence that was found for emotional salient contexts on face processing may rely on top-down modulation. Early activation based on coarse information derived from scenes may be helpful in extreme survival related situation where an immediate reaction is necessary (Bar, 2003). Several studies suggest that top-down modulation may occur on this early stage for face processing. First, a priori knowledge about face stimuli that were otherwise interpreted as schematic displays increases the N170 amplitude (Bentin & Golland, 2002; Bentin et al., 2002). A candidate anatomical connection for top-down modulation may be the orbitofrontal cortex and/or amygdala upon the fusiform gyrus (Aggleton, Burton, & Passingham, 1980; Ghashghaei & Barbas, 2002).

If the N170 is (in part) generated by the fusiform gyrus (Herrmann, Ehlis, Muehlberger et al., 2005; Pizzagalli et al., 2002), the orbitofrontal cortex and/or amygdala may exert a top-down modulation on the fusiform gyrus and thus affect the N170 amplitude. Although a direct relation between the orbitofrontal cortex and/or amygdala activity and the N170 amplitude has as yet not been shown, a number of arguments predict such a relation. First, it has been found that object recognition elicits activity in the orbitofrontal cortex (peak latency ~130 ms) that precedes activation in the fusiform gyrus (~180 ms), which is perhaps explained by a feedback modulation between orbitofrontal cortex and the fusiform gyrus (Bar et al., 2006). The occurrence of the N170 converges with the time-course of this activation. Second, the orbitofrontal cortex is not only important for optimal object perception, but is also important for emotion perception. fMRI data have shown an enhanced response in the orbitofrontal cortex to facial expressions and emotional scenes (Blair, Morris, Frith, Perrett, & Dolan, 1999; Smith, Stephan, Rugg, & Dolan, 2006; Taylor et al., 2003). Intracranial recordings have shown early neuronal responses (~120 ms) that diverge between aversive and neutral emotional scenes (Kawasaki et al., 2001). Therefore, activation in this region precedes the activity that underlies the N170 component.

The amygdala may also play a role in the feedback connection with the fusiform gyrus. First, the human amygdala has shown responses that precede the N170 (~120 ms, see Halgren et al., 1994). Second, fMRI studies have shown a relation between the activity of the amygdala and the activity of the fusiform gyrus to facial expressions of fear (Morris et al., 1998; Vuilleumier et al., 2004). The fMRI response in the fusiform gyrus increases to higher intensities of facial expressions (Morris et al., 1998; Surguladze et al., 2003). This is consistent with a significant correlation that has been observed between the fMRI BOLD response in the fusiform gyrus and the N170 amplitude (Iidaka et al., 2006).

In addition to the fusiform gyrus, the superior temporal sulcus has been proposed as a generator of the N170 (Henson et al., 2003; Itier & Taylor, 2004c). The amygdala may also have a feedback connection with the superior temporal sulcus (Aggleton et al., 1980; Allison et al., 2000). The superior temporal sulcus has been implied in social perception (Allison et al., 2000). The fusiform gyrus and the superior temporal sulcus may be involved in distinct aspects of face processing as has been shown for faces isolated from contexts (Haxby et al., 2000). Future studies may investigate how these different brain regions are involved in the perception of faces in contexts, and whether similar top-down modulations should be proposed for face processing as suggested for objects in contexts (Bar, 2004).

*7.4.4 Recognition memory for faces in contexts*

Contexts may influence recognition memory for faces on distinct stages of processing: the stage of encoding and the stage of retrieval. In the encoding stage, meaningful contexts may divert attention away from face processing. It has been shown that attentional factors influence performance on a face recognition memory task (Jenkins et al., 2005; Reinitz et al., 1994). Behavioral data showed that recognition memory (i.e., the proportion of hits) was better for faces that had been encoded in scrambled scenes as compared to faces encoded in intact scenes, which would suggest that meaningful scenes distract attention from face encoding. However, analyses that corrected for response-bias showed no significant difference for old-new discrimination (Pr). However, response bias was higher for faces that had been encoded in scrambled scenes than faces that had been encoded in intact scenes. These results suggest that participants who had received the task of experiment 1, where faces were shown in intact contexts, were more certain about the old-new status of faces, possibly because of the associative memory for the encoded context, whereas participants who had received the task of experiment 2, where faces were shown in scrambled contexts, tended to make more false alarms (chapter 6). The issue whether meaningful contexts during face encoding increases confidence during retrieval may be an interesting research domain. An important question is how meaningful contexts affect face processing during encoding (see for effects of emotional scenes on retrieval, Smith, Dolan et al., 2004; Sterpenich et al., 2006).

ERP data have shown that it is likely that face encoding is modulated by meaningful contexts. It has been shown that the N170 is decreased for faces in meaningful contexts as compared to faces on gray backgrounds (chapter 2) and scrambled contexts (chapter 4)[5].

---

Interestingly, Galli et al. (2006) have recently shown that the N170 is decreased for faces that had been encoded in meaningful contexts as compared to faces that had been presented on a white background. Importantly, the N170 was acquired during the test-phase, in which all faces were presented without any context (chapter 6). This suggests that contexts that were presented during face encoding (i.e., study-phase) had affected the retrieval of faces (i.e., test-phase). Perhaps, this effect may relate to the different encoding of faces in intact contexts or to the retrieval of the context associations with faces. However, no difference in behavioral performance on recognition memory was found for faces that had been studied in context compared to faces studied without context.

In addition to attention attracted by meaningful contexts, a different factor that plays a key role is whether the context is emotional salient. On the one hand, emotions may attract attention away from faces (see for comparable view for objects encoded in negative emotional contexts, Smith et al., 2004). Emotional scenes capture attention in early stage of processing (Calvo & Lang, 2005). On the other hand, an emotionally arousing event may increase memory performance for the event (Bradley et al., 1992; Hamann et al., 1999; Heuer & Reisberg, 1990), and may facilitate encoding of the association between face and context and therefore improve recognition memory for faces.

A recent study found better memory for faces that had been studied in verbal negative contexts (i.e., negative newspaper messages). In addition, they found a larger positive-going LPC for faces that had been studied in negative contexts compared to positive contexts (Galli et al., 2006). We found the reversed effect on the LPC. In contrast, no significant effect was found for faces that were encoded in emotional scenes (i.e., study-phase) on face recognition memory performance (chapter 6). This replicates previous results that emotional scenes do not affect face recognition memory (Sterpenich et al., 2006).

Others have observed the reversed effect of positive and negative contexts. Rainis (2001) has reported a decreased memory for faces in negative contexts, but an increased memory for negative contexts that were reinstated. Others have found better memory performances for non-face objects (Smith, Dolan et al., 2004) and words (Erk et al., 2003) that had been encoded in positive contexts. Thus, the results differ across stimulus materials, but may also depend on the task that was applied. Another factor that affects memory performance is the time period that is interpolated between study-phase and test-phase. It has been reported that the recall of highly arousing words increases over time, from 2 minutes interval to one week interval, whereas the recall shows a high immediate recall but rapid

[5] This comparison was complicated in chapter 6, where a different group of participants were presented with faces in scrambled context. The comparison is based on a single condition for each group (group 1: faces in intact contexts; group 2: faces in scrambled contexts). Therefore, potential differences on ERP amplitudes may rely on spurious factors that affect the absolute amplitude differently for both groups.

decline in recall for words that are not highly arousing (Kleinsmith & Kaplan, 1963). However, to our knowledge, this effect has not been studied as yet for facial expressions or faces that had been encoded in emotional salient contexts.

Sterpenich et al. (2006) have studied how recognition memory for faces is modulated by emotional context information by using a study-test paradigm. In the study-phase, faces were coupled to scenes (negative / neutral), whereas in the test-phase old/new faces were presented without any context information. Participants that showed a better retrieval of faces that were encoded in emotional contexts (16 out of 30 participants) were separated from the subjects that showed no emotional effect at retrieval (14 participants). It was found that the amygdala was involved in the recognition of faces that were encoded in negative relative to neutral contexts. The fMRI recordings were made during the test-phase. The amygdala activity was related to faces that were successfully recognized, and that had been studied in emotional compared to neutral contexts. This is interesting, because no context information was presented in the test-phase, a finding which has been observed before (Erk, Martin, & Walter, 2005). Unfortunately, the fMRI results for the study-phase were not reported. fMRI studies had hitherto shown that the amygdala response during the study-phase predicts recall of emotional scenes (Hamann et al., 1999) and videos at test (Cahill et al., 1996). Whether the amygdala response during the study-phase also predicts memory for faces is unresolved as yet. However, as mentioned before, it would be interesting to investigate the underlying neural systems and the time-course of neural activity related to face processing in semantic and emotional contexts.

In chapter 6 we used a study-test paradigm in which emotional scenes were only presented during the study-phase. An interesting question is how emotional contexts influence recognition memory when these contexts are reinstated. A response-bias may be expected given the false positive rates found for facial expressions (Johansson et al., 2004) and emotional words (Windmann & Kutas, 2001).

## 7.5 Conclusions

The results suggest that contexts importantly affect how faces are recognized. Firstly, the results suggest that meaningful scenes delay face processing. Meaningful contexts influence an early stage of face processing. Perceptual load or spatial attention were proposed as possible mechanisms for these effects.

Secondly, the results indicate that the contexts in which facial expressions are perceived have a key role in social perception. The emotional saliency of context information contributes to speed of face processing. The ERP data have shown that effects of emotional scene contexts occur on an early stage of face processing and show, in addition to the existing literature on facial expressions, that the N170 is sensitive to the perception of emotions from face and context. Top-down modulation based on prior knowledge about the emotional events may affect how facial expressions are perceived, and is consistent with the time courses that have been found in previous studies of face and object perception. The recognition memory studies for faces encoded in contexts suggest that the neural systems for memory for faces may be influenced by the context that is associated with the facial identity retrieved from memory. Future studies may investigate whether similar neural systems are involved for perceiving semantic and emotional relations between face and context.

## List of tables and figures

# References


Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion of Neurobiology, 12*, 169-177.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature, 372*, 669-672.

Aggleton, J. P., Burton, M. J., & Passingham, R. E. (1980). Cortical and subcortical afferents to the amygdala of the rhesus monkey (Macaca mulatta). *Brain Research, 190*, 347-368.

Albright, T. D., & Stoner, G. R. (2002). Contextual influences on visual processing. *Annual Review of Neuroscience, 25*, 339-379.

Algom, D., Chajut, E., & Lev, S. (2004). A rational look at the emotional stroop phenomenon: a generic slowdown, not a stroop effect. *Journal of Experimental Psychology: General, 133*, 323-338.

Allison, T., Begleiter, A., McCarthy, G., Roessler, E., Nobre, A. C., & Spencer, D. D. (1993). Electrophysiological studies of color processing in human visual cortex. *Electroencephalography and Clinical Neurophysiology, 88*, 343-355.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences, 4*, 267-278.

Aminoff, E., Gronau, N., & Bar, M. (2006). The Parahippocampal Cortex Mediates Spatial and Nonspatial Associations. *Cerebral Cortex, In Press.*

Ashley, V., Vuilleumier, P., & Swick, D. (2004). Time course and specificity of event-related potentials to emotional expressions. *Neuroreport, 15*, 211-216.

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience, 15*, 600-609.

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5*, 617-629.

Bar, M., & Aminoff, E. (2003). Cortical analysis of visual context. *Neuron, 38*, 347-358.

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., Hamalainen, M. S., Marinkovic, K., Schacter, D. L., Rosen, B. R., & Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the USA, 103*, 449-454.

Batty, M., & Taylor, M. J. (2003). Early processing of the six basic facial emotional expressions. *Cognitive Brain Research, 17*, 613-620.

Beales, S. A., & Parkin, A. J. (1984). Context and facial memory: the influence of different processing strategies. *Human Learning, 3*, 257-264.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience, 8*, 551-565.

Bentin, S., & Deouell, L. Y. (2000). Structural encoding and identification in face processing: ERP evidence for separate mechanisms. *Cognitive Neuropsychology, 17*, 35-54.

Bentin, S., & Golland, Y. (2002). Meaningful processing of meaningless stimuli: the influence of perceptual experience on early visual processing of faces. *Cognition, 86*, B1-14.

Bentin, S., Sagiv, N., Mecklinger, A., Friederici, A., & von Cramon, Y. D. (2002). Priming visual face-processing mechanisms: electrophysiological evidence. *Psychological Science, 13*, 190-193.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14*, 143-177.

Bigelow, N. O., Paradiso, S., Adolphs, R., Moser, D. J., Arndt, S., Heberlein, A., Nopoulos, P., & Andreasen, N. C. (2006). Perception of socially relevant stimuli in schizophrenia. *Schizophrenia Research, 83*, 257-267.

Blair, R. J., Morris, J. S., Frith, C. D., Perrett, D. I., & Dolan, R. J. (1999). Dissociable neural responses to facial expressions of sadness and anger. *Brain, 122*, 883-893.

Blanchette, I. (2006). Snakes, spiders, guns, and syringes: how specific are evolutionary constraints on the detection of threatening stimuli? *Quarterly Journal of Experimental Psychology, 59*, 1484-1504.

Bötzel, K., & Grusser, O. J. (1989). Electric brain potentials evoked by pictures of faces and non-faces: a search for "face-specific" EEG-potentials. *Experimental Brain Research, 77*, 349-360.

Bötzel, K., Schulze, S., & Stodieck, S. R. (1995). Scalp topography and analysis of intracranial sources of face-evoked potentials. *Experimental Brain Research, 104*, 135-143.

Bower, G. H., & Karlin, M. B. (1974). Depth of processing pictures of faces and recognition memory. *Journal of Experimental Psychology, 103*, 751-757.

Boyce, S. J., & Pollatsek, A. (1992). Identification of objects in scenes: the role of scene background in object naming. *Journal of Experimental Psychology: Learning Memory & Cognition, 18*, 531-543.

Boyce, S. J., Pollatsek, A., & Rayner, K. (1989). Effect of background information on object identification. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 556-566.

Bradley, M. M., Greenwald, M. K., Petry, M. C., & Lang, P. J. (1992). Remembering pictures: pleasure and arousal in memory. *Journal of Experimental Psychology: Learning Memory & Cognition, 18*, 379-390.

Britton, J. C., Taylor, S. F., Sudheimer, K. D., & Liberzon, I. (2006). Facial expressions and complex IAPS pictures: common and differential networks. *Neuroimage, 31*, 906-919.

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

Caharel, S., Courtay, N., Bernard, C., Lalonde, R., & Rebai, M. (2005). Familiarity and emotional expression influence an early stage of face processing: an electrophysiological study. *Brain and Cognition, 59*, 96-100.

Cahill, L., Haier, R. J., Fallon, J., Alkire, M. T., Tang, C., Keator, D., Wu, J., & McGaugh, J. L. (1996). Amygdala activity at encoding correlated with long-term, free recall of emotional information. *Proceedings of National Academy of Sciences of the USA, 93*, 8016-8021.

Calder, A. J., Lawrence, A. D., & Young, A. W. (2001). Neuropsychology of fear and loathing. *Nature Reviews Neuroscience, 2*, 352-363.

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 527-551.

Calvo, M. G., & Lang, P. J. (2004). Gaze patterns when looking at emotional pictures: Motivationally biased attention. *Motivation and Emotion, 28*, 221-243.

Calvo, M. G., & Lang, P. J. (2005). Parafoveal semantic processing of emotional visual scenes. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 502-519.

Campanella, S., Hanoteau, C., Depy, D., Rossion, B., Bruyer, R., Crommelinck, M., & Guerit, J. M. (2000). Right N170 modulation in a face discrimination task: an account for categorical perception of familiar faces. *Psychophysiology, 37*, 796-806.

Campanella, S., Montedoro, C., Streel, E., Verbanck, P., & Rosier, V. (2006). Early visual components (P100, N170) are disrupted in chronic schizophrenic patients: an event-related potentials study. *Neurophysiology Clinique / Clinical Neurophysiology, 36*, 71-78.

Campanella, S., Quinet, P., Bruyer, R., Crommelinck, M., & Guerit, J. M. (2002). Categorical perception of happiness and fear facial expressions: an ERP study. *Journal of Cognitive Neuroscience, 14*, 210-227.

Carretie, L., Hinojosa, J. A., Martin-Loeches, M., Mercado, F., & Tapia, M. (2004). Automatic attention to emotional stimuli: Neural correlates. *Human Brain Mapping, 22*, 290-299.

Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology, 70*, 205-218.

Cauquil, A. S., Edmonds, G. E., & Taylor, M. J. (2000). Is the face-sensitive N170 the only ERP not affected by selective attention? *Neuroreport, 11*, 2167-2171.

Christianson, S. A. (1992). Emotional stress and eyewitness memory: a critical review. *Psychological Bulletin, 112*, 284-309.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.*: Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Coles, M. G. H., Gratton, G., & Fabiani, M. (1990). Event-related brain potentials. In J. T. Cacioppo & L. G. Tassinary (Eds.), *Principles of Psychophysiology*: Cambridge University Press.

Cox, D., Meyers, E., & Sinha, P. (2004). Contextually evoked object-specific responses in human visual cortex. *Science, 304*, 115-117.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Nonverbal behavior, 11*, 671-684.

Curran, T. (2004). Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. *Neuropsychologia, 42*, 1088-1106.

Darwin, C. (1998). *The expression of emotion in man and animal.* New York: Oxford University Press. (original work published 1872).

Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science, 15*, 559-564.

Davies, G., & Milne, A. (1982). Recognizing faces in and out of context. *Current Psychological Research, 2*, 235-246.

de Gelder, B., Meeren, H. K. M., Righart, R., Van den Stock, J., van de Riet, W. A. C., & Tamietto, M. (2006). Chapter 3 Beyond the face: exploring rapid influences of context on face processing. *Progress in Brain Research, 155PB*, 37-48.

de Gelder, B., & Vroomen, J. (2000). The perception of emotion by ear and eye. *Cognition and Emotion, 14*, 289-311.

de Gelder, B., Vroomen, J., de Jong, S. J., Masthoff, E. D., Trompenaars, F. J., & Hodiamont, P. (2005). Multisensory integration of emotional faces and voices in schizophrenics. *Schizophrenia Research, 72*, 195-203.

de Houwer, J., & Hermans, D. (1994). Differences in affective processing of words and pictures. *Cognition and Emotion, 8*, 1-20.

Dolan, R. J., Fink, G. R., Rolls, E., Booth, M., Holmes, A., Frackowiak, R. S., & Friston, K. J. (1997). How the brain learns to see objects and faces in an impoverished context. *Nature, 389*, 596-599.

Dolan, R. J., Morris, J. S., & de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proceedings of the National Academy of Sciences of the USA, 98*, 10006-10010.

Dolcos, F., & Cabeza, R. (2002). Event-related potentials of emotional memory: encoding pleasant, unpleasant, and neutral pictures. *Cognitive Affective & Behavioral Neuroscience, 2*, 252-263.

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica, 30*, 412-431.

Dougherty, F. E., Bartlett, E. S., & Izard, C. E. (1974). Responses of schizophrenics to expressions of the fundamental emotions. *Journal of Clinical Psychology, 30*, 243-246.

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

Eger, E., Jedynak, A., Iwaki, T., & Skrandies, W. (2003). Rapid extraction of emotional expression: evidence from evoked potential fields during brief presentation of face stimuli. *Neuropsychologia, 41*, 808-817.

Eimer, M. (2000). The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport, 11*, 2319-2324.

Eimer, M., & Holmes, A. (2002). An ERP study on the time course of emotional face processing. *Neuroreport, 13*, 427-431.

Eimer, M., Holmes, A., & McGlone, F. P. (2003). The role of spatial attention in the processing of facial expression: an ERP study of rapid brain responses to six basic emotions. *Cognitive Affective & Behavioral Neuroscience, 3*, 97-110.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*, 169-200.

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

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature, 392*, 598-601.

Erk, S., Kiefer, M., Grothe, J., Wunderlich, A. P., Spitzer, M., & Walter, H. (2003). Emotional context modulates subsequent memory effect. *Neuroimage, 18*, 439-447.

Erk, S., Martin, S., & Walter, H. (2005). Emotional context during encoding of neutral items modulates brain activation not only during encoding but also during recognition. *Neuroimage, 26*, 829-838.

Erthal, F. S., de Oliveira, L., Mocaiber, I., Pereira, M. G., Machado-Pinheiro, W., Volchan, E., & Pessoa, L. (2005). Load-dependent modulation of affective picture processing. *Cognitive Affective & Behavioral Neuroscience, 5*, 388-395.

Feenan, K., & Snodgrass, J. G. (1990). The effect of context on discrimination and bias in recognition memory for pictures and words. *Memory & Cognition, 18*, 515-527.

Feinberg, T. E., Rifkin, A., Schaffer, C., & Walker, E. (1986). Facial discrimination and emotional recognition in schizophrenia and affective disorders. *Archives of General Psychiatry, 43*, 276-279.

Fernández-Dols, J. M., & Carroll, J. M. (1997). Is the meaning perceived in facial expression independent of its context? In J. A. Russell & J. M. Fernández-Dolls (Eds.), *The psychology of facial expression* (pp. 275-294): Cambridge University Press.

Fitzgerald, D. A., Angstadt, M., Jelsone, L. M., Nathan, P. J., & Phan, K. L. (2006). Beyond threat: amygdala reactivity across multiple expressions of facial affect. *Neuroimage, 30*, 1441-1448.

Ford, J. M., White, P. M., Csernansky, J. G., Faustman, W. O., Roth, W. T., & Pfefferbaum, A. (1994). ERPs in schizophrenia: effects of antipsychotic medication. *Biological Psychiatry, 36*, 153-170.

Fox, E., Lester, V., Russo, R., Bowles, R., Pichler, A., & Dutton, K. (2000). Facial expressions of emotion: are angry faces detected more efficiently? *Cognition and Emotion, 14*, 61-92.

Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General, 130*, 681-700.

Friedman, A. (1979). Framing pictures: the role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General, 108*, 316-355.

Frijda, N.H. (1961). Facial expressions and situational cues: a control. *Acta Psychologica, 18*, 239-244.

Frijda, N. H. (1986). *The emotions*. Cambridge: Cambridge University Press.

Frijda, N. H., & Tcherkassof, A. (1997). Facial expressions as modes of action readiness. In J. A. Russell & J. M. Fernández-Dolls (Eds.), *The psychology of facial expression*. Cambridge: Cambridge University Press.

Galli, G., Feurra, M., & Viggiano, M. P. (2006). "Did you see him in the newspaper?" Electrophysiological correlates of context and valence in face processing. *Brain Research, 1119*, 190-202.

Ganis, G., & Kutas, M. (2003). An electrophysiological study of scene effects on object identification. *Cognitive Brain Research, 16*, 123-144.

Gazzaley, A., Cooney, J. W., McEvoy, K., Knight, R. T., & D'Esposito, M. (2005). Top-down enhancement and suppression of the magnitude and speed of neural activity. *Journal of Cognitive Neuroscience, 17*, 507-517.

George, N., Dolan, R. J., Fink, G. R., Baylis, G. C., Russell, C., & Driver, J. (1999). Contrast polarity and face recognition in the human fusiform gyrus. *Nature Neuroscience, 2*, 574-580.

George, N., Evans, J., Fiori, N., Davidoff, J., & Renault, B. (1996). Brain events related to normal and moderately scrambled faces. *Cognitive Brain Research, 4*, 65-76.

Ghashghaei, H. T., & Barbas, H. (2002). Pathways for emotion: interactions of prefrontal and anterior temporal pathways in the amygdala of the rhesus monkey. *Neuroscience, 115*, 1261-1279.

Goh, J. O., Siong, S. C., Park, D., Gutchess, A., Hebrank, A., & Chee, M. W. (2004). Cortical areas involved in object, background, and object-background processing revealed with functional magnetic resonance adaptation. *Journal of Neuroscience, 24*, 10223-10228.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology, 55*, 468-484.

Grimshaw, G. M., Bulman-Fleming, M. B., & Ngo, C. (2004). A signal-detection analysis of sex differences in the perception of emotional faces. *Brain and Cognition, 54*, 248-250.

Guillaume, F., & Tiberghien, G. (2001). An event-related potential study of contextual modifications in a face recognition task. *Neuroreport, 12*, 1209-1216.

Gur, R. E., McGrath, C., Chan, R. M., Schroeder, L., Turner, T., Turetsky, B. I., Kohler, C., Alsop, D., Maldjian, J., Ragland, J. D., & Gur, R. C. (2002). An fMRI study of facial emotion processing in patients with schizophrenia. *American Journal of Psychiatry, 159*, 1992-1999.

Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology, 28*, 240-244.

Halgren, E., Baudena, P., Heit, G., Clarke, J. M., Marinkovic, K., & Clarke, M. (1994). Spatio-temporal stages in face and word processing. I. Depth-recorded potentials in the human occipital, temporal and parietal lobes. *Journal of Physiology, 88*, 1-50.

Halgren, E., Raij, T., Marinkovic, K., Jousmaki, V., & Hari, R. (2000). Cognitive response profile of the human fusiform face area as determined by MEG. *Cerebral Cortex, 10*, 69-81.

Hamann, S. B., Ely, T. D., Grafton, S. T., & Kilts, C. D. (1999). Amygdala activity related to enhanced memory for pleasant and aversive stimuli. *Nature Neuroscience, 2*, 289-293.

Hannula, D. E., Federmeier, K. D., & Cohen, N. J. (2006). Event-related potential signatures of relational memory. *Journal of Cognitive Neuroscience, 18*, 1863-1876.

Hansen, C. H., & Hansen, R. D. (1988). Finding the face in the crowd: an anger superiority effect. *Journal of Personality and Social Psychology, 54*, 917-924.

Hariri, A. R., Tessitore, A., Mattay, V. S., Fera, F., & Weinberger, D. R. (2002). The amygdala response to emotional stimuli: a comparison of faces and scenes. *Neuroimage, 17*, 317-323.

Hasson, U., Hendler, T., Ben Bashat, D., & Malach, R. (2001). Vase or face? A neural correlate of shape-selective grouping processes in the human brain. *Journal of Cognitive Neuroscience, 13*, 744-753.

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

Heisz, J. J., Watter, S., & Shedden, J. M. (2006). Progressive N170 habituation to unattended repeated faces. *Vision Research, 46*, 47-56.

Hempel, R. J., Tulen, J. H., van Beveren, N. J., van Steenis, H. G., Mulder, P. G., & Hengeveld, M. W. (2005). Physiological responsivity to emotional pictures in schizophrenia. *Journal of Psychiatry Research, 39*, 509-518.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology, 50*, 243-271.

Henson, R. N., Goshen-Gottstein, Y., Ganel, T., Otten, L. J., Quayle, A., & Rugg, M. D. (2003). Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cerebral Cortex, 13*, 793-805.

Herrmann, M. J., Ehlis, A. C., Ellgring, H., & Fallgatter, A. J. (2005). Early stages (P100) of face perception in humans as measured with event-related potentials (ERPs). *Journal of Neural Transmission, 112*, 1073-1081.

Herrmann, M. J., Ehlis, A. C., Muehlberger, A., & Fallgatter, A. J. (2005). Source localization of early stages of face processing. *Brain Topography, 18*, 77-85.

Herrmann, M. J., Ellgring, H., & Fallgatter, A. J. (2004). Early-stage face processing dysfunction in patients with schizophrenia. *American Journal of Psychiatry, 161*, 915-917.

Hess, U., Blairy, S., & Kleck, R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal behavior, 21*, 241-257.

Heuer, F., & Reisberg, D. (1990). Vivid memories of emotional events: the accuracy of remembered minutiae. *Memory & Cognition, 18*, 496-506.

Hill, H., Schyns, P. G., & Akamatsu, S. (1997). Information and viewpoint dependence in face recognition. *Cognition, 62*, 201-222.

Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceeding of the National Academy of Sciences of the USA, 95*, 781-787.

Hollingworth, A., & Henderson, J. M. (1998). Does consistent scene context facilitate object perception? *Journal of Experimental Psychology: General, 127*, 398-415.

Holmes, A., Vuilleumier, P., & Eimer, M. (2003). The processing of emotional facial expression is gated by spatial attention: evidence from event-related brain potentials. *Cognitive Brain Research, 16*, 174-184.

Howell, D. C. (2002). *Statistical methods for psychology.*: Duxbury Thomson Learning. 5th Ed.

Iidaka, T., Matsumoto, A., Haneda, K., Okada, T., & Sadato, N. (2006). Hemodynamic and electrophysiological relationship involved in human face processing: evidence from a combined fMRI-ERP study. *Brain and Cognition, 60*, 176-186.

Itier, R. J., Latinus, M., & Taylor, M. J. (2006). Face, eye and object early processing: what is the face specificity? *Neuroimage, 29*, 667-676.

Itier, R. J., & Taylor, M. J. (2004a). Effects of repetition learning on upright, inverted and contrast-reversed face processing using ERPs. *Neuroimage, 21*, 1518-1532.

Itier, R. J., & Taylor, M. J. (2004b). N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cerebral Cortex, 14*, 132-142.

Itier, R. J., & Taylor, M. J. (2004c). Source analysis of the N170 to faces and objects. *Neuroreport, 15*, 1261-1265.

Jeffreys, D. A. (1989). A face-responsive potential recorded from the human scalp. *Experimental Brain Research, 78*, 193-202.

Jemel, B., Schuller, A. M., Cheref-Khan, Y., Goffaux, V., Crommelinck, M., & Bruyer, R. (2003). Stepwise emergence of the face-sensitive N170 event-related potential component. *Neuroreport, 14*, 2035-2039.

Jenkins, R., Lavie, N., & Driver, J. (2005). Recognition memory for distractor faces depends on attentional load at exposure. *Psychonomic Bulletin Review, 12*, 314-320.

Johansson, M., Mecklinger, A., & Treese, A. C. (2004). Recognition memory for emotional and neutral faces: an event-related potential study. *Journal of Cognitive Neuroscience, 16*, 1840-1853.

Jokisch, D., Daum, I., Suchan, B., & Troje, N. F. (2005). Structural encoding and recognition of biological motion: evidence from event-related potentials and source analysis. *Behavioural Brain Research, 157*, 195-204.

Joyce, C., & Rossion, B. (2005). The face-sensitive N170 and VPP components manifest the same brain processes: the effect of reference electrode site. *Clinical Neurophysiology, 116*, 2613-2631.

Joyce, C. A., & Kutas, M. (2005). Event-related potential correlates of long-term memory for briefly presented faces. *Journal of Cognitive Neuroscience, 17*, 757-767.

Kawasaki, H., Kaufman, O., Damasio, H., Damasio, A. R., Granner, M., Bakken, H., Hori, T., Howard, M. A., 3rd, & Adolphs, R. (2001). Single-neuron responses to emotional visual stimuli recorded in human ventral prefrontal cortex. *Nature Neuroscience, 4*, 15-16.

Keltner, D., & Kring, A. M. (1998). Emotion, social function and psychopathology. *Review in General Psychology, 2*, 320-342.

Kerr, S. L., & Neale, J. M. (1993). Emotion perception in schizophrenia: specific deficit or further evidence of generalized poor performance? *Journal of Abnormal Psychology, 102*, 312-318.

Kim, H., Somerville, L. H., Johnstone, T., Polis, S., Alexander, A. L., Shin, L. M., & Whalen, P. J. (2004). Contextual modulation of amygdala responsivity to surprised faces. *Journal of Cognitive Neuroscience, 16*, 1730-1745.

Kirita, T., & Endo, M. (1995). Happy-face advantage in recognizing facial expressions. *Acta Psychologica, 89*, 149-163.

Kirouac, G., & Dore, F. Y. (1983). Accuracy and latency of judgment of facial expressions of emotions. *Perceptual Motor Skills, 57*, 683-686.

Kleinsmith, L. J., & Kaplan, S. (1963). Paired-associate learning as a function of arousal and interpolated interval. *Journal Experimental Psychology, 65*, 190-193.

Klucharev, V., & Sams, M. (2004). Interaction of gaze direction and facial expressions processing: ERP study. *Neuroreport, 15*, 621-625.

Kohler, C. G., Bilker, W., Hagendoorn, M., Gur, R. E., & Gur, R. C. (2000). Emotion recognition deficit in schizophrenia: association with symptomatology and cognition. *Biological Psychiatry, 48*, 127-136.

Kohler, C. G., Turner, T. H., Bilker, W. B., Brensinger, C. M., Siegel, S. J., Kanes, S. J., Gur, R. E., & Gur, R. C. (2003). Facial emotion recognition in schizophrenia: intensity effects and error pattern. *American Journal Psychiatry, 160*, 1768-1774.

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

Kowatari, Y., Yamamoto, M., Takahashi, T., Kansaku, K., Kitazawa, S., Ueno, S., & Yamane, S. (2004). Dominance of the left oblique view in activating the cortical network for face recognition. *Neuroscience Research, 50*, 475-480.

Krolak-Salmon, P., Fischer, C., Vighetto, A., & Mauguiere, F. (2001). Processing of facial emotional expression: spatio-temporal data as assessed by scalp event-related potentials. *European Journal of Neuroscience, 13*, 987-994.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*, 161-163.

Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences, 23*, 571-579.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1999). International Affective Picture System (IAPS): Technical Manual and Affective Ratings. *NIMH Center for the Study of Emotion and Attention, University of Florida, Gainesville.*

Lang, P. J., Bradley, M. M., Fitzsimmons, J. R., Cuthbert, B. N., Scott, J. D., Moulder, B., & Nangia, V. (1998). Emotional arousal and activation of the visual cortex: an fMRI analysis. *Psychophysiology, 35*, 199-210.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 451-468.

Lavie, N. (2005). Distracted and confused?: selective attention under load. *Trends in Cognitive Sciences, 9*, 75-82.

LeDoux, J. E. (1996). *The emotional brain*. New York: Simon & schuster.

Leppänen, J. M., Tenhunen, M., & Hietanen, J. K. (2003). Faster choice-reaction times to positive than to negative facial expressions. *Journal of Psychophysiology, 17*, 113-123.

Lewis, M. B., & Edmonds, A. J. (2003). Face detection: mapping human performance. *Perception, 32*, 903-920.

Loftus, G. R., Nelson, W. W., & Kallman, H. J. (1983). Differential acquisition rates for different types of information from pictures. *Quarterly Journal of Experimental Psychology, 35*, 187-198.

Luce, R. D. (1986). *Response times. Their role in inferring elementary mental organization.*: Oxford University Press, New York.

Luck, S. J. (2006). *An introduction to the event-related potential technique*. Cambridge: MIT Press.

Lundqvist, D., Flykt, A., & Ohman, A. (1998). *The Karolinska Directed Emotional Faces-KDEF*: CD-ROM from the Department of Clinical Neuroscience, Psychology Section, Karolinska Institute Stockholm. ISBN 91-630-71649.

Maddock, R. (2000). Reply. *Trends in Neurosciences, 23*, 196-197.

Maddock, R. J. (1999). The retrosplenial cortex and emotion: new insights from functional neuroimaging of the human brain. *Trends in Neurosciences, 22*, 310-316.

Mandler, G. (1980). Recognizing: the judgment of previous occurrence. *Psychological Review, 87*, 252-271.

Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology, 32*, 4-18.

Maratos, E. J., Allan, K., & Rugg, M. D. (2000). Recognition memory for emotionally negative and neutral words: an ERP study. *Neuropsychologia, 38*, 1452-1465.

Mathis, K. M. (2002). Semantic interference from objects both in and out of a scene context. *Journal of Experimental Psychology: Learning Memory & Cognition, 28*, 171-182.

McGaugh, J. L. (2003). *Memory and Emotion. The making of lasting memories.*: Columbia University Press, New York.

McKelvie, S. J. (1995). Emotional expression in upside-down faces: evidence for configurational and componential processing. *British Journal of Social Psychology, 34*, 325-334.

McNeely, H. E., Dywan, J., & Segalowitz, S. J. (2004). ERP indices of emotionality and semantic cohesiveness during recognition judgments. *Psychophysiology, 41*, 117-129.

Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences of the USA, 102*, 16518-16523.

Memon, A., & Bruce, V. (1983). The effects of encoding strategy and context change on face recognition. *Human Learning, 2*, 313-326.

Mesquita, B., & Frijda, N.H. (1992). Cultural variations in emotions: a review. *Psychological Bulletin, 112*, 179-204.

Miyoshi, M., Katayama, J., & Morotomi, T. (2004). Face-specific N170 component is modulated by facial expressional change. *Neuroreport, 15*, 911-914.

Mobbs, D., Weiskopf, N., Lau, H. C., Featherstone, E., Dolan, R. J., & Frith, C. D. (2006). The Kuleshov effect: the influence of contextual framing on emotional attributions. *Social Cognitive and Affective Neuroscience, 1*, 95-106.

Morris, J. S., Friston, K. J., Buchel, C., Frith, C. D., Young, A. W., Calder, A. J., & Dolan, R. J. (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain, 121*, 47-57.

Mueser, K. T., Doonan, R., Penn, D. L., Blanchard, J. J., Bellack, A. S., Nishith, P., & DeLeon, J. (1996). Emotion recognition and social competence in chronic schizophrenia. *Journal of Abnormal Psychology, 105*, 271-275.

Murphy, G. L., & Wisniewski, E. J. (1989). Categorizing objects in isolation and in scenes: what a superordinate is good for. *Journal of Experimental Psychology: Learning Memory & Cognition, 15*, 572-586.

Nunez, P. L., & Srinivasan, R. (2006). *Electric fields of the brain. The neurophysics of EEG.* New York: Oxford University Press.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: detecting the snake in the grass. *Journal of Experimental Psychology: General, 130*, 466-478.

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: a threat advantage with schematic stimuli. *Journal of Personality and Social Psychology, 80*, 381-396.

Onitsuka, T., Niznikiewicz, M. A., Spencer, K. M., Frumin, M., Kuroki, N., Lucia, L. C., Shenton, M. E., & McCarley, R. W. (2006). Functional and structural deficits in brain regions subserving face perception in schizophrenia. *American Journal of Psychiatry, 163*, 455-462.

Palermo, R., & Coltheart, M. (2004). Photographs of facial expression: accuracy, response times, and ratings of intensity. *Behavioral Research Methods, Instruments & Computers, 36*, 634-638.

Paller, K. A., Gonsalves, B., Grabowecky, M., Bozic, V. S., & Yamada, S. (2000). Electrophysiological correlates of recollecting faces of known and unknown individuals. *Neuroimage, 11*, 98-110.

Palmer, S. E. (1975). The effect of contextual scenes on the identification of objects. *Memory and Cognition, 3*, 519-526.

Patterson, K. E., & Baddeley, A. D. (1977). When face recognition fails. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 406-417.

Penn, D. L., Combs, D. R., Ritchie, M., Francis, J., Cassisi, J., Morris, S., & Townsend, M. (2000). Emotion recognition in schizophrenia: further investigation of generalized versus specific deficit models. *Journal of Abnormal Psychology, 109*, 512-516.

Phelps, E. A., O'Connor, K. J., Gatenby, J. C., Gore, J. C., Grillon, C., & Davis, M. (2001). Activation of the left amygdala to a cognitive representation of fear. *Nature Neuroscience, 4*, 437-441.

Phillips, M. L., Drevets, W. C., Rauch, S. L., & Lane, R. (2003). Neurobiology of emotion perception II: Implications for major psychiatric disorders. *Biological Psychiatry, 54*, 515-528.

Phillips, M. L., Young, A. W., Senior, C., Brammer, M., Andrew, C., Calder, A. J., Bullmore, E. T., Perrett, D. I., Rowland, D., Williams, S. C., Gray, J. A., & David, A. S. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature, 389*, 495-498.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Jr., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D., & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology, 37*, 127-152.

Pizzagalli, D. A., Lehmann, D., Hendrick, A. M., Regard, M., Pascual-Marqui, R. D., & Davidson, R. J. (2002). Affective judgments of faces modulate early activity (~160 ms) within the fusiform gyri. *Neuroimage, 16*, 663-677.

Potter, M. C. (1975). Meaning in visual search. *Science, 187*, 965-966.

Puce, A., Allison, T., & McCarthy, G. (1999). Electrophysiological studies of human face perception. III: Effects of top-down processing on face-specific potentials. *Cerebral Cortex, 9*, 445-458.

Puce, A., Smith, A., & Allison, T. (2000). ERPs evoked by viewing facial movements. *Cognitive Neuropsychology, 17*, 221-239.

Puce, A., Syngeniotis, A., Thompson, J. C., Abbott, D. F., Wheaton, K. J., & Castiello, U. (2003). The human temporal lobe integrates facial form and motion: evidence from fMRI and ERP studies. *Neuroimage, 19*, 861-869.

Rainis, N. (2001). Semantic contexts and face recognition. *Applied Cognitive Psychology, 15*, 173-186.

Reinitz, M. T., Morrissey, J., & Demb, J. (1994). Role of attention in face encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition., 20*, 161-168.

Righart, R., & de Gelder, B. (2006). Context influences early perceptual analysis of faces--an electrophysiological study. *Cerebral Cortex, 16*, 1249-1257.

Righart, R., & de Gelder, B. (2007). Recognition of facial expressions is influenced by emotional scene gist. *Manuscript submitted for publication.*

Rockstroh, B., Junghofer, M., Elbert, T., Buodo, G., & Miller, G. A. (2006). Electromagnetic brain activity evoked by affective stimuli in schizophrenia. *Psychophysiology, 43*, 431-439.

Rolls, E. T., Aggelopoulos, N. C., & Zheng, F. (2003). The receptive fields of inferior temporal cortex neurons in natural scenes. *Journal of Neuroscience, 23*, 339-348.

Rousselet, G. A., Mace, M. J., & Fabre-Thorpe, M. (2004a). Spatiotemporal analyses of the N170 for human faces, animal faces and objects in natural scenes. *Neuroreport, 15*, 2607-2611.

Rousselet, G. A., Mace, M. J. M., & Fabre-Thorpe, M. (2004b). Animal and human faces in natural scenes: How specific to human faces is the N170 ERP component? *Journal of Vision, 4*, 13-21.

Rozin, P., & Fallon, A. E. (1987). A perspective on disgust. *Psychological Review, 94*, 23-41.

Rugg, M. D., & Coles, M. G. H. (1995). *Electrophysiology of mind*: Oxford University Press.

Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature, 392*, 595-598.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin, 115*, 102-141.

Russell, J. A. (1997). Reading emotion from and into faces: Resurrecting a dimensional-contextual perspective. In J. A. Russell & J. M. Fernández-Dols (Eds.), *The psychology of facial expression*: Cambridge University Press.

Salem, J. E., Kring, A. M., & Kerr, S. L. (1996). More evidence for generalized poor performance in facial emotion perception in schizophrenia. *Journal of Abnormal Psychology, 105*, 480-483.

Schienle, A., Stark, R., Walter, B., Blecker, C., Ott, U., Kirsch, P., Sammer, G., & Vaitl, D. (2002). The insula is not specifically involved in disgust processing: an fMRI study. *Neuroreport, 13*, 2023-2026.

Schneider, F., Gur, R. C., Koch, K., Backes, V., Amunts, K., Shah, N. J., Bilker, W., Gur, R. E., & Habel, U. (2006). Impairment in the specificity of emotion processing in schizophrenia. *American Journal of Psychiatry, 163*, 442-447.

Schupp, H. T., Ohman, A., Junghofer, M., Weike, A. I., Stockburger, J., & Hamm, A. O. (2004). The facilitated processing of threatening faces: an ERP analysis. *Emotion, 4*, 189-200.

Schweinberger, S. R. (1995). Repetition priming and associative priming on face recognition: evidence from event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition., 21*, 722-736.

Schweinberger, S. R., Pickering, E. C., Jentzsch, I., Burton, A. M., & Kaufmann, J. M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Cognitive Brain Research, 14*, 398-409.

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: evidence for time- and spatial scale-dependent scene recognition. *Psychological Science, 5*, 195-200.

Smith, A. P., Dolan, R. J., & Rugg, M. D. (2004). Event-related potential correlates of the retrieval of emotional and nonemotional context. *Journal of Cognitive Neuroscience, 16*, 760-775.

Smith, A. P., Henson, R. N., Dolan, R. J., & Rugg, M. D. (2004). fMRI correlates of the episodic retrieval of emotional contexts. *Neuroimage, 22*, 868-878.

Smith, A. P., Stephan, K. E., Rugg, M. D., & Dolan, R. J. (2006). Task and content modulate amygdala-hippocampal connectivity in emotional retrieval. *Neuron, 49*, 631-638.

Smith, N. K., Cacioppo, J. T., Larsen, J. T., & Chartrand, T. L. (2003). May I have your attention, please: Electrocortical responses to positive and negative stimuli. *Neuropsychologia, 41*, 171-183.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*, 34-50.

Spence, I., Wong, P., Rusan, M., & Rastegar, N. (2006). How color enhances visual memory for natural scenes. *Psychological Science, 17*, 1-6.

Sprengelmeyer, R., Young, A. W., Calder, A. J., Karnat, A., Lange, H., Homberg, V., Perrett, D. I., & Rowland, D. (1996). Loss of disgust. Perception of faces and emotions in Huntington's disease. *Brain, 119*, 1647-1665.

Stekelenburg, J. J., & de Gelder, B. (2004). The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect. *Neuroreport, 15*, 777-780.

Sternberg, S. (1969). Memory-scanning: mental processes revealed by reaction-time experiments. *American Scientist, 57*, 421-457.

Sterpenich, V., D'Argembeau, A., Desseilles, M., Balteau, E., Albouy, G., Vandewalle, G., Degueldre, C., Luxen, A., Collette, F., & Maquet, P. (2006). The locus ceruleus is involved in the successful retrieval of emotional memories in humans. *Journal of Neuroscience, 26*, 7416-7423.

Sugase, Y., Yamane, S., Ueno, S., & Kawano, K. (1999). Global and fine information coded by single neurons in the temporal visual cortex. *Nature, 400*, 869-873.

Surguladze, S. A., Brammer, M. J., Young, A. W., Andrew, C., Travis, M. J., Williams, S. C., & Phillips, M. L. (2003). A preferential increase in the extrastriate response to signals of danger. *Neuroimage, 19*, 1317-1328.

Takahashi, H., Koeda, M., Oda, K., Matsuda, T., Matsushima, E., Matsuura, M., Asai, K., & Okubo, Y. (2004). An fMRI study of differential neural response to affective pictures in schizophrenia. *Neuroimage, 22*, 1247-1254.

Tanaka-Matsumi, J., Attivissimo, D., Nelson, S., & D'Urso, T. (1995). Context effects on the judgment of basic emotions in the face. *Motivation and Emotion, 19*, 139-155.

Taylor, M. J., Edmonds, G. E., McCarthy, G., & Allison, T. (2001). Eyes first! Eye processing develops before face processing in children. *Neuroreport, 12*, 1671-1676.

Taylor, S. F., Phan, K. L., Britton, J. C., & Liberzon, I. (2005). Neural response to emotional salience in schizophrenia. *Neuropsychopharmacology, 30*, 984-995.

Taylor, S. F., Phan, K. L., Decker, L. R., & Liberzon, I. (2003). Subjective rating of emotionally salient stimuli modulates neural activity. *Neuroimage, 18*, 650-659.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*, 520-522.

Tipples, J., Young, A. W., Quinlan, P., Broks, P., & Ellis, A. W. (2002). Searching for threat. *Quarterly Journal of Experimental Psychology, 55*, 1007-1026.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. *Psychological Review, 113*, 766-786.

Tsivilis, D., Otten, L. J., & Rugg, M. D. (2001). Context effects on the neural correlates of recognition memory: an electrophysiological study. *Neuron, 31*, 497-505.

Tsivilis, D., Otten, L. J., & Rugg, M. D. (2003). Repetition effects elicited by objects and their contexts: an fMRI study. *Human Brain Mapping, 19*, 145-154.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352-373.

Van den Stock, J., Righart, R., & de Gelder, B. (2006). Emotions expressed in the whole body influence recognition of facial expressions and emotional prosody. *Emotion, In press.*

Vogel, E. K., & Luck, S. J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology, 37*, 190-203.

Vogt, B. A., Absher, J. R., & Bush, G. (2000). Human retrosplenial cortex: where is it and is it involved in emotion? *Trends in Neurosciences, 23*, 195-197.

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience, 6*, 624-631.

Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J., & Dolan, R. J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nature Neuroscience, 7*, 1271-1278.

Vuilleumier, P., & Schwartz, S. (2001). Emotional facial expressions capture attention. *Neurology, 56*, 153-158.

Walker, E., Marwit, S. J., & Emory, E. (1980). A cross-sectional study of emotion recognition in schizophrenics. *Journal of Abnormal Psychology, 89*, 428-436.

Wallbott, H. G. (1988). Faces in context: the relative importance of facial expressions and context information in determining emotion attributions. In K. R. Scherer (Ed.), *Facets of emotion (pp. 139-166).*

Watanabe, S., Miki, K., & Kakigi, R. (2002). Gaze direction affects face perception in humans. *Neuroscience Letters, 325*, 163-166.

White, M. (1995). Preattentive analysis of facial expressions of emotion. *Cognition and emotion, 9*, 439-460.

Whittaker, J. F., Deakin, J. F., & Tomenson, B. (2001). Face processing in schizophrenia: defining the deficit. *Psychological Medicine, 31*, 499-507.

Wilding, E. L. (2000). In what way does the parietal ERP old/new effect index recollection? *International Journal of Psychophysiology, 35*, 81-87.

Williams, L. M., Palmer, D., Liddell, B. J., Song, L., & Gordon, E. (2006). The 'when' and 'where' of perceiving signals of threat versus non-threat. *Neuroimage, 31*, 458-467.

Windmann, S., & Kutas, M. (2001). Electrophysiological correlates of emotion-induced recognition bias. *Journal of Cognitive Neuroscience, 13*, 577-592.

Winograd, E., & Rivers-Bulkeley, N. T. (1977). Effect of changing context on remembering faces. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 397-405.

Yi, D. J., Woodman, G. F., Widders, D., Marois, R., & Chun, M. M. (2004). Neural fate of ignored stimuli: dissociable effects of perceptual and working memory load. *Nature Neuroscience, 7*, 992-996.

Yip, A. W., & Sinha, P. (2002). Contribution of color to face recognition. *Perception, 31*, 995-1003.

Yonelinas, A. P. (2002). The nature of recollection and familiarity. A review of 30 year of research. *Journal of Memory and Language, 46*, 441-517.

Young, A. W., Hay, D. C., & Ellis, A. W. (1985). The faces that launched a thousand slips: everyday difficulties and errors in recognizing people. *British Journal of Psychology, 76*, 495-523.

Yovel, G., & Paller, K. A. (2004). The neural basis of the butcher-on-the-bus phenomenon: when a face seems familiar but is not remembered. *Neuroimage, 21*, 789-800.

## Summary

Numerous studies in face recognition have centered on the question how identity and expression are processed by the brain. In the past, many studies have overlooked the importance of the context in which we perceive a person and his/her facial expression.

The context in which we perceive others often is meaningful. For example, if we are used to see the butcher in his butchery, it can take considerably more time to recognize the butcher when he is suddenly seen in a different context, for example a bus. We are able to interpret a facial expression of emotion readily, oftentimes because it is part of a reaction to an emotional event. For example, if we see someone with dilated eyes, the reason for this reaction may at first hand be unclear. However, if we perceive this expression in the context of a burning house, then the emotional salience is immediately clear to us.

The experimental research that has been reported in this dissertation is concerned with the question what role contexts play in face recognition. The research questions were focussed on the influence of emotional meaningful contexts on the recognition of facial identity and expression. These questions were studied by using pictures of faces that were positioned centrally on the background of an emotional scene. Behavioral and EEG measures were used to investigate performance and brain activity that is related to identity and expression recognition for faces in contexts.

Behavioral performance (=the number of correctly recognized faces) and response times (=the time in milliseconds it takes to recognize faces correctly) were acquired in an emotion categorization task in which participants choose among a number of expressions (three alternatives: fear, happiness, disgust) that were positioned in emotional scenes (fear, happiness, disgust). We predicted that facial expressions would be better and faster categorized if the emotional value of the context matched with the facial expression (for example, a facial expression of happiness that is shown in front of a landscape of flowers).

By using EEG measures (=Electroencephalography) it was investigated how the time course of face processing is affected by emotional contexts, which can be envisioned by analyzing the ERP components (=event-related potential components). The ERPs represent brain activity that is related to the processing of the faces in context. The ERP components can be derived from the ongoing EEG by means of signal averaging. The ERPs that have been related to early stages of face processing are the P1 and N170 component, which occur at about 100 ms and 170 ms respectively after stimulus presentation. Previous studies have reported that the amplitudes of these ERP components are not only sensitive to the presence of a face, but also change by facial expressions of emotion. Given these findings, we were interested whether the amplitudes of these ERP components would also change if facial expressions were presented in emotional salient contexts, for example a fearful expression in

the context of a burning house. In addition to healthy participants, we also studied a group of schizophrenic patients because of their deficits on emotion recognition.

ERPs were also used to investigate whether recognition memory for faces would be influenced by the emotional context in which faces had been encoded. In addition to the P1 and N170 components that are related to the early stages of face processing, we have measured ERP components that are related to recognition memory. The LPC (=late positive complex) has been related to recognition memory and occurs at longer latencies (~400 ms after stimulus onset). The amplitude of the LPC is found to be more positive to faces that have been correctly recognized as being "old" (=seen in the previous session) as compared to faces that have been correctly recognized as being "new" (=not seen in the previous session). We predicted that the emotional context that accompanies faces during encoding may distract attention from face processing, which may result in worse memory performance and less positive amplitudes for the LPC.

The results show that emotional contexts influence face processing. It was found that facial expressions of emotion were recognized better and faster if the emotional content of the facial expression and context were similar. The effects occur with very short presentation times (=200 milliseconds), which suggests a fast processing of emotional scene gist. The ERPs show that the brain is capable to process emotional information very fast. The N170 amplitude was increased (=more negative amplitudes) for faces, most prominently for fearful faces, that were shown in context scenes that evoke fear, as compared to scenes that were neutral and happy. Schizophrenia patients showed like the healthy participants larger N170 amplitudes for fearful faces in fearful contexts as compared to happy contexts, but not if compared to neutral contexts. These results suggest that schizophrenia patients rely on contrasting valences (negative-positive) in emotion processing.

It was found that recognition memory performance for faces was not influenced by emotional scene contexts in which faces were studied. Differences were found on the LPC, as amplitudes were more positive for correctly recognized "old" faces than correctly recognized "new" faces, which corresponds with previous studies that did not use context stimuli. Faces that were studied in a fear-evoking contexts showed less positive amplitudes, which agrees with the formulated hypothesis.

Rapid processing to facial expressions and its surrounding context may be necessary to respond optimally and adaptively. The results indicate that contexts in which facial expressions are perceived play a key role in emotion perception. The results suggest that prior knowledge about emotional events may affect how facial expressions are perceived, and how easily a situation is understood. An explanation of top-down mechanisms by which the brain processes complex information is consistent with the time-courses that have been found in previous face and object recognition studies and will be discussed.

# Samenvatting

Veel wetenschappelijk onderzoek is uitgevoerd naar hoe we herkennen dat we een persoon eerder hebben gezien (identiteit) en hoe we emotionele gelaatsexpressies kunnen herkennen (expressies). Hierbij is echter het belang van de context waarin we een persoon en zijn/haar gelaatsexpressie waarnemen over het hoofd gezien.

De context waarin we anderen ontmoeten is vaak betekenisvol. Bijvoorbeeld, als we gewend zijn de slager in zijn slagerij te treffen, kan het moeite kosten dezelfde persoon te herkennen in een andere omgeving, bijvoorbeeld als we de slager onverwacht in de bus tegen komen. Een emotionele gelaatsexpressie kunnen we onmiddellijk interpreteren, vaak omdat deze een reactie is op een emotionele gebeurtenis in de omgeving. Als we bijvoorbeeld iemand met opengesperde ogen zien, dan kan onduidelijk zijn waarom iemand deze gelaatsexpressie toont. Als we deze gelaatsexpressie echter waarnemen in de context van een brandend huis, dan is de emotie gelijk duidelijk.

Het experimenteel onderzoek dat is gerapporteerd in dit proefschrift heeft zich gericht op de vraag welke betekenis de context in gezichtsherkenning heeft. De vragen zijn toegespitst op de invloed van emotioneel betekenisvolle gebeurtenissen op herkenning van identiteit en gelaatsexpressies. De invloed van contexten is onderzocht door foto's van gezichten centraal te positioneren op een achtergrond van een emotionele scène (foto van een gebeurtenis waarin iets emotioneels gebeurt). Gedragsmetingen en EEG metingen zijn verricht om de prestatie en de hersenactiviteit die gerelateerd zijn aan identiteit en expressieherkenning voor gezichten in contexten te meten.

De accuratesse (= het aantal emotionele gelaatsexpressies die worden herkend) en de responstijd zijn gemeten (= de tijd in milliseconden die nodig is om een expressie te herkennen) voor een categorisatie taak waarin proefpersonen kiezen welke emotionele gelaatsexpressie ze waarnemen (drie alternatieven, angst, blijdschap, walging), en deze expressies waren gepositioneerd op een scène (angst, blijdschap, walging). We voorspelden dat gezichtsexpressies beter en sneller worden herkend indien de emotie van de context overeenkomt met de gelaatsexpressie (b.v. een blije expressie voor een landschap van bloemen).

Door gebruik te maken van EEG metingen (=Elektroencefalografie) is onderzocht hoe het tijdsverloop van gezichtsverwerking beïnvloed wordt door emotionele contexten. Dit kan onderzocht worden door de bijbehorende ERP componenten (ERPs = event-related potentials) te analyseren. ERPs representeren hersenactiviteit die gerelateerd is aan de verwerking van een stimulus, in dit geval het gezicht in context. De ERP componenten kunnen worden geanalyseerd na middeling van het EEG signaal dat bij iedere gebeurtenis hoort. De ERPs die vaak worden gerelateerd aan vroege stadia van verwerking zijn de P1 en N170. De P1 component en de N170 component treden op na respectievelijk circa 100 ms (milliseconden) en 170 ms na de aanbieding van het plaatje van het gezicht. In eerder onderzoek is al gevonden dat deze componenten niet alleen worden opgeroepen bij de presentatie van een gezicht op een beeldscherm, maar dat de amplituden van deze componenten verandert indien de gezichten emotionele gelaatsexpressies hebben. Naar aanleiding van deze bevindingen waren we geïnteresseerd of de amplituden van deze ERP componenten ook veranderen indien gelaatsexpressies worden aangeboden in een emotionele context, bijvoorbeeld een angstige expressie in de context van een brandend huis. Naast gezonde proefpersonen hebben we onderzocht hoe schizofrenie patiënten gezichten in emotionele contexten verwerken, gegeven hun problemen met emotieherkenningstaken.

ERPs zijn ook gebruikt om te onderzoeken of herkenningsgeheugen voor gezichten beïnvloed wordt door de emotionele context waarin deze verwerkt waren. Naast de P1 component en de N170 component die gerelateerd zijn aan de vroege stadia van gezichtsverwerking, zijn ook ERPs gemeten die gerelateerd zijn aan geheugen. De LPC (= Late Positive Complex) is gerelateerd aan herkenningsgeheugen en treedt op na ongeveer 400 ms nadat een gezicht is gepresenteerd. De amplitude is positiever voor gezichten die correct als "oud" zijn herkend (=proefpersonen hebben deze gezien in een voorgaande sessie) in vergelijking met de gezichten die correct als "nieuw" zijn herkend (=proefpersonen hebben deze niet gezien in voorgaande sessie). We voorspelden dat de emotionele context de aandacht afleidt van gezichtsverwerking, wat resulteert in een slechtere prestatie op de geheugentaak en minder positieve amplituden op de LPC.

Uit de resultaten komt naar voren dat emotioneel betekenisvolle contexten invloed hebben op hoe gezichten verwerkt worden. Er werd gevonden dat emotionele gelaatsexpressies beter en sneller worden herkend indien ze in een emotioneel passende context worden waargenomen. Dit effect treedt op bij zeer korte aanbiedingstijden (=200 milliseconden), wat er mogelijk op wijst dat er snelle globale verwerking van de scène plaatsvindt naast de verwerking van de gelaatsexpressie. De ERP resultaten laten zien dat de hersenen de emotionele informatie zeer snel verwerken. De N170 piek amplitude was groter (=meer negatief) voor gezichten die werden getoond in een angstopwekkende context dan gezichten die werden getoond in een neutrale of een vreugdevolle context. Schizofreniepatiënten tonen net als de gezonde personen grotere N170 amplitudes voor angstige gezichten in een angst opwekkende context vergeleken met een blijde context, maar niet vergeleken met een neutrale context. De ERP resultaten suggereren dat schizofreniepatiënten de contrasterende waarden van emotie (negatief-positief) verwerken, maar moeilijk neutrale contexten kunnen onderscheiden.

In het onderzoek naar herkenningsgeheugen voor gezichten is gevonden dat de geheugenprestatie niet verandert onder invloed van de emotionele context waarin een gezicht bestudeerd was. Wel waren er verschillen op de LPC, waar de amplitudes meer positief waren voor correct als "oud" herkende gezichten dan correct als "nieuw" herkende gezichten, wat overeenkomt met de literatuur over gezichten zonder contexten. Gezichten die waren verwerkt in een angstige context lieten minder positieve amplituden zien, wat overeenkomt met de opgestelde hypothesen.

De snelle verwerking van gezichtsexpressies en de context zijn mogelijk noodzakelijk om optimaal en adaptief op de omgeving te kunnen reageren. De resultaten geven aan dat de context waarin gezichtsexpressies zijn waargenomen een hoofdrol spelen bij de waarneming van emoties. Voorkennis over emotionele lading van gebeurtenissen kan beïnvloeden hoe gezichtsexpressies worden waargenomen, en dat een situatie sneller wordt begrepen. De "top-down mechanismen" die de hersenen hanteren om complexe informatie snel te verwerken zal worden bediscussieerd.

## Acknowledgements

I am grateful to the following persons who have supported me with this thesis or otherwise during this period: first I would like to thank my promotor Prof. dr. Beatrice de Gelder for offering me a position in her lab. It was very exciting to set up and work out this novel area of research!

I thank the members of the dissertation committee (Prof dr. P. Bertelson, Dr. G.J.M. van Boxtel, Prof. dr. E. de Haan, Prof. dr. N.H. Frijda, Prof. dr. K.R. Ridderinkhof) for accepting the invitation for this graduation.

Ton Heinen, Geert van Boxtel, Jean Vroomen, Ton van Boxtel, Sjakko de Jong (GGZ-west brabant), Jeroen Stekelenburg, Hanneke Meeren, Karen Mansfield, Ilja Frissen (Now Tübingen), Viona Wijnen, Wim van de Riet, Jan Van den Stock, Sabine van Linden (Now TNO), Mirjam Keetels, Corné van Heijnsbergen (Now Neurofeedback), Thessa Caus (Now GGZ), Bernard Stienen, Charlotte Sinke, Mariska Kret, Maurice Magnee (Utrecht University), and the colleagues of psychology & health for creating the right atmosphere to work across different domains of expertise (behavioral, neuropsychology, EEG, fMRI).

Many thanks to those who made it all possible: the participants in research for prosopagnosia, schizophrenia, alzheimer's disease and students of the Tilburg University. Jeske Hendriks, Annemieke Vink and Thessa Caus who have realised a master's thesis. SPITS, a.o., Charles Rambelje for support in the experimental timing.

I've learnt that visiting other labs is an important experience! I would like to thank Bea and the researchers of MGH (Nouchine Hadjikhani, Ksenija Marinkovic, Wim van Duffel and John Arsenault), MIT (Sinha) and the Bar Lab in Boston for their time and enthusiasm to show their work.

Finally, I would like to thank my family and friends for continuous support, in special my mother, my brother David, and my little (but great) sister Marie-Colien.

Tilburg / Rotterdam,
April 2007

Ruthger

# Curriculum Vitae

Ruthger Righart was born on the 27th of July 1977 in Brouwershaven. In 1996 he started studying construction engineering in Rotterdam. After succeeding his first-year examination (Propaedeutic year), in 1997 he decided to switch to psychology. From 1997 until 2003 he studied psychology at Leiden University and obtained his master's degree in clinical psychology with a specialization in neuropsychology. From 2003 until 2007 he conducted research at Tilburg University for his PhD dissertation using behavioral measures and event-related potentials. During his education, he has worked with different neurological and psychiatric populations (depression, schizophrenia, Huntington's disease, prosopagnosia).

www.ruthger.righart.net

# Uitnodiging

raag nodig ik u uit voor
het bijwonen van de
penbare verdediging
van mijn proefschrift:

*"The Role of Contexts
in Face Processing.
Behavioral and
ERP studies"*

p vrijdag 15 juni 2007
16.00 in de aula van de
niversiteit van Tilburg,
/arandelaan 2, Tilburg.

Aansluitend is er een
receptie ter plaatse.

Ruthger Righart
3023 SM Rotterdam
06-18083564
R.G.R.Righart@uvt.nl

SBN: 978 90 5335 123 9