THE COSTS AND BENEFITS OF EMOTIONAL PROCESSING

BEHAVIORAL PERFORMANCE AND ELECTROCORTICAL DYNAMICS

Dissertation zur Erlangung des Doktorgrades

Eingereicht an der Naturwissenschaftlichen Fakultät der
Universität Konstanz

Von Niklas Ihssen

Tag der mündlichen Prüfung: 18. Juli 2007

1. Referent: PD Dr. Andreas Keil
2. Referent: Prof. Dr. Harald Schupp

Contents

DANK

Dank gebührt an erster Stelle meiner betreuenden Doppelspitze, Sabine Heim und Andreas Keil. Sie haben diese Arbeit ermöglicht, mich in allen Belangen unterstützt und gehören zweifelsohne zum Besten, was man sich als Doktorand wünschen kann. Wissenschaft ist ein Gemeinschaftswerk. Und daher möchte ich allen danken, die an der Vorbereitung, Durchführung und Auswertung dieser Studien mitgeholfen haben: Nadine Stumpf, Julia Hansmann, Leonie Koban, Hanna Gruell, Alexander Dienst, Christian Pietreck, Annette Gomolla und Thekla Witte. Dank auch an Nathan und Stephan für manch kleineren und größeren Ratschlag, an die große Zahl von Studierenden, die an unseren Experimenten teilgenommen haben, an meine Familie und Freunde für ihre Geduld. Und natürlich Danke, Shanna: Большое спасибо, Солнышко.

1. INTRODUCTION

In a complex and rapidly changing environment, adaptive behavior requires extraction and efficient processing of sensory information. Specialized mechanisms are implemented in the architecture of our mental processing systems that structure the incoming flow of information according to current needs, long-term goals, and situational demands. Among other functions, these mechanisms act to distinguish hostile from benign events and potential threats from benefits. In scientific contexts, these capabilities have been linked to evolutionary old emotional systems regulating perception and optimizing behavioral outcome in motivationally relevant situations (Lang, Bradley, & Cuthbert, 1997). Empirical evidence from various disciplines converges to reveal emotion as a potent modulator of both perceptual and attentional processes. Contemporary lines of this research have a strong focus on examining the biological underpinnings of these phenomena. Subsumed under the research framework of affective stimulus processing, these approaches have been stimulated by the emergence of new physiological recording and data analysis techniques ranging from advanced electroencephalographic to hemodynamic and optical imaging methods but also including improved measures of peripheral physiology.

The growing physiological data base has lead to a comparatively coherent picture denoting an organized pattern of bodily and central-nervous changes during affective stimulus perception. With respect to brain activity indices of perception, findings converge to show amplified cortical responses towards emotional information. Accordingly, a perspective has evolved stating that the human brain processes emotional information in a near-optimal manner. In experimental-laboratory settings, emotion-related perceptual enhancement has been typically documented for situations of passive viewing (or hearing) a stimulus. However, empirical findings are surprisingly inconclusive as to how affective stimulus perception translates into behavioral responses, especially when mental operations are carried out on the stimulus. The present thesis centers on the question whether behavioral performance is facilitated or impaired by the affective content of a stimulus. Within this broader research frame, specific conditions for behavioral facilitation and interference are investigated presenting emotional stimuli in different cognitive and perceptual tasks. Four series of studies are reported capitalizing on several performance behaviors modulated by emotional stimulus perception. Dependent variables include both behavioral (e.g., reaction times, identification accuracy) and electrocortical (event-related potentials) measures. Emphasis is thus given to the relation between neuronal and behavioral indices of affective stimulus processing. As illustrated in the following chapter, a suitable theoretical framework to approach the outlined issues is to view emotions as action dispositions that are implemented in neuronal networks (Lang & Davis, 2006).

1.1 Emotion as Action Dispositions: A Theoretical Framework

1.1.1 Appetitive and Defensive Motivation

Emotion theorists from various disciplines emphasize a motivational basis of emotion (e.g., Cacioppo & Berntson, 1999; Chen & Bargh, 1999; Frijda, 1986; Lang & Davis, 2006). Emotions are considered as being organized around two basic and evolutionary old motive systems, which have been labeled with varying terms as “appetitive and defensive” (Lang, 1979), “preservative and protective” (Konorski, 1967), or “attractive and aversive” (Dickinson & Dearing, 1979). From a functional perspective, these anatomically distinct systems have evolved to prepare and organize adaptive behavior in situations relevant for survival (Bradley, Codispoti, Cuthbert, & Lang, 2001). The defensive system becomes active in contexts that signal danger or directly threaten the organism’s physical survival. It coordinates appropriate withdrawal or attack behavior. Conversely, the appetitive system is engaged in situations promoting an organism’s well-being. It initiates approach behavior towards appetitive cues, such as food or mating partners. In influential work by Peter Lang (for a review see Lang & Davis, 2006) these ideas have been elaborated and tested under the framework of a two-dimensional model of emotion. The polymorphism of emotional reactions and subjective experiences is broken down to two basic dimensions of emotion: hedonic valence and emotional arousal. The model draws on multivariate analyses of emotional language using semantic differentials (e.g., Osgood, Suci, & Tannenbaum, 1957). In these studies, the factors of valence and arousal were found to be sufficient predictors of the principal variance in the data. According to Lang (e.g., Lang et al., 1997), the two dimensions are indices of motivational activation, with valence indicating which system is active (appetitive or defensive) and emotional arousal reflecting the intensity of motivational activation. To test the model’s assumptions, an extensive research program has been launched using as primary measures subjective-experiential, biological reflexes, and peripheral physiological changes in response to standardized sets of emotional stimuli. Findings confirm that the variance of affective responses can be considered as the output of two motivational systems varying in activation level. For instance, affective judgments of large samples of words or pictures do not yield a random distribution within the two-dimensional affective space but rather form two gradients reflecting the degree of appetitive and defensive motivation (Lang, Bradley, & Cuthbert, 2005). Further, during picture viewing the startle blink reflex is differentially affected for aversive (potentiation) versus appetitive (inhibition) items suggesting congruence-dependent modulation driven by pre-activated/primed motive systems (Lang, Bradley, & Cuthbert, 1990). Other measures dissociating appetitive versus defensive motivation include heart rate deceleration and electromyographic activity of facial muscles. Instead, parameters such as skin conductance covary systematically with the arousal dimension, which is seen as an instance of common somatic output of the two motivational systems (Bradley et al., 2001).

In summary, emotions are phenomena having multiple facets and thus call for multiple description levels (physiology, behavior, experience). Yet, the multivariate data outlined so far offer plausible reason to conceive emotion as the output of two motive systems which i) have distinct neuronal substrates and ii) evoke a systematic and temporally sequenced pattern of mental and bodily changes in response to discrete events. Functionally, these changes prepare for action. Of course, the notions reviewed above are only one possible theoretical avenue to emotion. A multitude of other approaches have been made to define emotion. The approaches vary according to the major psychological concepts, e.g., reinforcement (Rolls, 1990) or evaluation (Cacioppo & Berntson, 1999), and the data being used, e.g., animal data or human neuropsychological data (Damasio, 1994). In addition, some authors have emphasized the existence of qualitatively distinct "basic emotions" (Panksepp, 1992). Others have stressed the role of cognitive appraisal processes and consider emotion as a decoupling of stimulus-response relations that increases behavioral flexibility (Scherer, 1984). In the current framework, biological approaches as offered for instance by Lang are especially useful with respect to three aspects: i) the models specify the neuronal underpinnings of hypothetical motive systems, ii) predictions are available as to the modulation of perceptual processing of affective stimuli, iii) emotions are viewed as networks. These three aspects are detailed as follows.

1.1.2 Neuronal Circuits Involved in Appetitive and Defensive Motivation

According to Lang, the two motive systems are implemented in specific, largely subcortical neuronal circuits having widespread connections to other cortical and subcortical brain regions including perceptual, visceral, and motor areas (for a review see Lang & Davis, 2006). Emotional arousal is not viewed as having a separate substrate, but rather qualifies the neuronal/metabolic intensity of system activation. Drawing from animal research, a key role in defensive motivation is assigned to the amygdaloid complex in the anterior medial temporal lobe. Sensory afferents reach the amygdala via the basolateral nuclei (lateral, basal, and basomedial nucleus), which project to the central nucleus and “extended amygdala” (bed nucleus of the stria terminalis), among other target structures. Efferents from here include projections to the lateral hypothalamus and diverse areas of the brain stem (e.g., central gray), which mediate the vegetative and somatic reflexes (e.g., heart rate deceleration or freezing) indexing action readiness. It is likely that these amygdaloid circuits are also engaged in appetitive motivation (Zald, 2003). For systems initiating approach behavior, additional projections have been emphasized connecting the amygdala with the mesolimbic reward/reinforcement system which comprises several dopaminergic structures including the ventral tegmental area and the nucleus accumbens in the ventral striatum (Everitt, Cador, & Robbins, 1989). The outlined proposals are also consistent with other suggestions as to the functional anatomy of emotion. For instance, Gray (1983) has linked the amygdala and its connections to a hypothetical fight-and-flight-system (FFS). Instead, the so-called behavioral approach system is assumed to be primarily supported by mesolimbic circuits.

1.1.3 Motive Systems and Visual Processing

Relevant for the current issues, the amygdaloid systems are also linked to a variety of cortical areas involved in perceptual, memory, and executive processes. Fibers from the basal nucleus innervate the hippocampus, the cingulate cortex, and the orbito-frontal cortex. Further, as revealed by anatomical tracing (Amaral, Behniea, & Kelly, 2003), the amygdala projects to all processing levels of the ventral visual stream beginning at early sensory stages in striate areas of the occipital lobe and continuing to the fusiform gyrus and higher-order encoding areas in inferior temporal cortices. Functional brain imaging work by Vuilleumier and colleagues (Vuilleumier, Richardson, Armony, Driver, & Dolan, 2004) provides compelling evidence that the amygdale modulates visual responses to emotional stimuli. In this work, patients with lesions of the amygdala were dissociated from patients with hippocampus damage and healthy controls in that the first group did not show increased activity of the fusiform cortex in response to fearful versus neutral faces. According to the dual-pathway hypothesis (LeDoux, 2000; Vuilleumier, 2005), this modulation may rely on a feedback loop providing the amygdala with sensory information via two different routes i) A slow route that conveys the results of high-level visual analysis in inferior temporal areas to the amygdala, and ii) a fast route that transmits coarse visual information from thalamic nuclei (pulvinar) and possibly the superior colliculi to the amygdala. Based on this fast sensory input, the amygdala is capable to modulate perception before ventral stream processing is completed. Alternatively to ii), it has been proposed that rapid amygdala activation (and reentrant modulation of early ventral stream processing) is due to an initial feedforward-sweep within the visual systems consisting of crude magnocellular information about the stimulus (Vuilleumier, 2005). Supportive evidence for the existence of a fast route has been gathered by studies showing that differential (right) amygdala responses to emotional versus neutral stimuli persist although the stimuli are masked/unseen (Morris, Öhman, & Dolan, 1998; Whalen et al., 1998) or the subjects are cortically blind (Morris, de Gelder, Weiskrantz, & Dolan, 2001; Pegna, Khateb, Lazeyras, & Seghier, 2005). Consistent with the proposed thalamic-collicular pathway, unseen fear-conditioned stimuli also activated the pulvinar and superior colliculi (Morris, Öhman, & Dolan, 1999). Among other phenomena, the outlined mechanisms may account for early (< 200 ms) modulation of event-related potentials during encoding of new emotional stimuli. Further, they may underlie the short-term reorganization of primary visual cortex observed across experimental trials during fear conditioning (Stolarova, Keil, & Moratti, 2006). Reentrant modulation of perceptual processing may also be relevant for a modulation of perceptual/cognitive performance as examined here.

1.1.4 Emotions as Networks and Knowledge Structures

Memory or knowledge structures can be described as propositional networks, where complex relations between pieces of meaning are formed by associative learning (Anderson & Bower, 1973). Formally, propositional networks comprise arguments (nodes in the network, e.g., “crocodile”, “Sabine”, “swim”) that are linked to attributes (“dangerous”,”green”) or to other arguments by predicates (“is”, “has”, “can”). When sensory input matches a sufficient number of network units, network activity is propagated to other network parts by means of spreading activation.

As noted by Lang (Lang & Davis, 2006), most affective cues are not hard-wired, i.e., they do not represent a survival-relevant event that directly activates the motive systems. Rather, they form learned associations to appetitive or defensive motivation. In the network view, symbolic emotional objects (words, pictures) as used in the studies here may be similarly described: They do not directly engage the motive systems but represent stimuli that match elements of an emotional network. According to Lang (Lang, 1993; Lang & Davis, 2006), these emotional networks are unique in that they include connections to the amygdaloid circuits. They thus integrate cortical information about perceptual or cognitive stimulus aspects (e.g., color, shape, category) with behavioral and physiological response patterns. Given that sufficient network nodes are activated, emotional memories or symbolic affective cues will thus be accompanied by emotion-specific autonomic and somatic responses. Due to social constraints, network activation will not necessarily result in execution of the associated behavioral responses (e.g., attack). Accordingly, Lang (1995) suggested that emotional experiences typically occur when the motive systems are activated but motivational behaviors are inhibited or delayed.

Crucially, cortico-amygdaloid connections are reciprocal (Amaral et al., 2003). In the current context, it may thus be expected that activated motivational network parts will alter activity of cognitive network parts and other knowledge structures, e.g., when mental operations are carried out on the stimulus, such as retrieving category membership or lexicality.

1.2 Brain Responses to Emotional Stimuli

A core tenet of the motivational model pertains to perceptual and attentional effects arising from appetitive/defensive motivation. Emotional events prompt reflexive or “motivated attention” (Lang et al., 1997), i.e., the processing systems are adjusted to increase sensory input, enhance perceptual processing, and gather stimulus and contextual information (Bradley et al., 2001). With the advent of brain imaging techniques, evidence has been accrued supporting the notion of emotion-driven perceptual enhancement. In the following sections, major hemodynamic and electrocortical findings of this research are summarized.

1.2.1 Hemodynamic Responses

Early work using positron emission tomography (PET, Lane et al., 1997) and functional magnetic resonance imaging (fMRI, Lang, Bradley, Fitzsimmons et al., 1998) reported a widespread posterior increase of functional brain activation to emotionally evocative versus neutral pictures, including areas of the occipital, inferior/superior parietal, and inferior temporal cortex. Similar results were obtained in studies using phobic stimuli (Fredrikson et al., 1993), emotional faces (e.g., Sprengelmeyer, Rausch, Eysel, & Przuntek, 1998) or fear-conditioned (Knight, Smith, Stein, & Helmstetter, 1999) stimuli. Subsequent fMRI studies refined these results showing blood flow increases in secondary visual areas as a function of pictures rated emotional arousal (Bradley et al., 2003; Sabatinelli, Flaisch, Bradley, Fitzsimmons, & Lang, 2004). Emotional enhancement of perceptual processing, even at early sensory stages, has been linked to recurrent influence of the amygdala, which may receive visual information via the thalamus quickly and prior to awareness (see chapter 1.1.3). Accordingly, hemodynamic measures have reliably observed increased amygdala activation in response to emotionally arousing relative to neutral stimuli including face, word, picture, and auditory stimuli (for a review see Vuilleumier, 2003). Further, amygdala activity covaries with visual cortex activity in response to stimuli varying in emotional arousal, which indicates a functional connectivity of those structures (Sabatinelli, Bradley, Fitzsimmons, & Lang, 2005).

1.2.2 Electrocortical Responses

Insights into the temporal dynamics of emotion-driven perceptual enhancement on a subsecond scale are offered by studies using electrophysiological measures. A multitude of event-related potential (ERP) studies reported diverging electrocortical patterns for emotional and neutral stimuli (including pictures, faces, words) beginning early after stimulus onset (for a review see Schupp, Flaisch, Stockburger, & Junghöfer, 2006). In studies of affective picture processing, two key findings have been replicated consistently:

(i) Viewing emotional pictures was related to enhancement of a late positive component (termed late positive potential LPP) reaching maximum between 400 - 600 ms over centroparietal sensors (Cacioppo, Crites Jr., Berntson, & Coles, 1993; Cuthbert, Schupp, Bradley, Birbaumer, & Lang, 2000; Keil et al., 2002). The LPP is a sensitive measure of rated picture arousal being most pronounced for evolutionary significant stimuli such as mutilation or erotica scenes (Schupp, Cuthbert et al., 2004). Source space projections using the minimum norm estimate (Keil et al., 2002) or discrete regional source modeling (Sabatinelli, Lang, Keil, & Bradley, 2007) indicated that the LPP originates from the joint activation of extrastriate occipital, inferiortemporal, and parietal regions. During picture viewing, the LPP correlates with other measures of appetitive/defensive motivation, including autonomic responses (Cuthbert et al., 2000) and BOLD (blood oxygen level-dependent) intensity in visual cortices (Sabatinelli et al., 2007). Functionally, enhanced LPPs have been proposed to reflect increased attentional allocation and continued analysis of motivational relevant stimuli at higher-order and post-sensory stages of stimulus processing (Cuthbert et al., 2000).

(ii) For arousing stimuli, a negative amplitude shift has been reported that occurs between 150 and 300 ms over occipitotemporal scalp regions (Junghöfer, Bradley, Elbert, & Lang, 2001; Schupp, Junghöfer, Weike, & Hamm, 2003). This early ERP modulation, termed Early Posterior Negativity (EPN), is thought to reflect a fast attention mechanism, which selects motivationally relevant information for further processing (Schupp et al., 2003). This view is supported by work showing that attention to specific semantic pictures categories (e.g., animal pictures) is associated with increased negativities having a similar time course and topography (Codispoti, Ferrari, De Cesarei, & Cardinale, 2006). Interestingly, explicit and reflexive/motivated forms of attention appear to affect the early ERP modulations in an additive manner, while later affective modulations (in the LPP range) are disproportional larger when emotional pictures serve the roles of attended targets, relative to non-targets (Schupp, Stockburger, Codispoti et al., 2007).

Bradley and colleagues (Bradley, Hamby, Low, & Lang, 2007) recently added to these findings that perceptual picture composition may also mediate early posterior negativities. Pictures containing simple figure ground scenes elicited less positivity (more negativity) between 150 ms and 250 ms, relative to more complex picture content. Emotional picture category modulated LPP responses but had only little effect on the early negativity.

While LPP and early ERP modulations have been reliably shown for both emotional pictures and facial expressions of emotion, electrocortical findings related to emotionally charged verbal material are less conclusive showing consistent effects in the LPP range only. (for a review see Kissler, Assadollahi, & Herbert, 2006). Relevant for the current experimental setup, perceptual enhancement of emotional pictures as indexed by the LPP and posterior negativity does not depend on long stimulus durations (Schupp, Junghöfer, Weike, & Hamm, 2004). Further, as opposed to the autonomic responses and startle reflex modulations, these electrocortical differences do not habituate with repeated stimulus exposure suggesting that novelty plays a minor role for enhanced processing of affective cues (Codispoti, Ferrari, & Bradley, 2007; Schupp, Stockburger et al., 2006). However, early affective modulations are subject to interference when a competing cognitive task is performed. Posterior negativities to arousing pictures were reduced when subjects had to count the occurrence of line patterns overlaid on the pictures (Schupp, Stockburger, Bublatzky et al., 2007). Worth mentioning, in this study emotion-cognition interactions were approached from a perspective that represents the complementary counterpart to the approach used here, namely how a cognitive task modulates the capability of affective processing.

As an alternative to the stimulus-locked ERP approaches detailed above, several authors examined electrocortical oscillations in the frequency domain as a sensitive measure of affective stimulus perception, for instance the evoked or induced gamma band response or the steady-state response elicited by flickering stimuli (for a review see Keil, 2006). Similar to time-domain ERP measures, GBA (e.g., Keil et al., 2001) and SSR responses (e.g., Keil, Moratti, Sabatinelli, Bradley, & Lang, 2005) reliably vary with emotional stimulus aspects.

1.3 Task Performance[1] and Emotional Stimuli

To engage appropriate actions, the motive systems hinge on accurate processing of the perceptual objects indicative of emotional significance. The above findings confirm that processing resources are automatically allocated to visual elements indexed as motivationally significant. Commensurate their adaptive value, emotional stimuli appear to profit from a perceptual and attentional primacy. These benefits may include optimized encoding at early stages (Schupp et al., 2003) and preferential access to higher-order processing resources (Cuthbert et al., 2000), thus “facilitating encoding and recognition in sensory systems” (Bradley et al., 2003, p. 370) . However, one might argue that ascribing quality to a mental process (i.e., using terms like “facilitated” or “optimized”) requires the registration of performance measures, such as response speed or response accuracy. Most brain imaging work capitalized on affective stimulus processing under passive viewing conditions and did not encompass dependent variables at the level of behavioral outcome. Thus, there is no unequivocal answer to the question, whether and how amplified cortical responses to emotional stimuli translate into behavior. In a narrow interpretation of the motivational model, appetitive/defensive motivation facilitates only appetitive/defensive behaviors. Except for certain reflexive behaviors (e.g., the startle blink), few predictions are available how the motive systems modulate simultaneous but independently evoked behaviors, such as responses in cognitive tasks. Drawing from the so-called defense cascade model (Lang et al., 1997), one might hypothesize detrimental effects of aversive content on manual responses due to motor inhibition. However, these suggestions are speculative. Thus, in the following sections, findings of several behavioral studies involving emotional stimuli are reviewed. These studies have employed a broad range of experimental paradigms differing in many aspects. Often, results appear contradictory showing either enhanced (Anderson, 2005) or impaired (McKenna & Sharma, 1995) performance on emotional stimuli, compared to neutral material. As shown below, the behavioral outcome may result from an interplay among a variety of dimensions (see also Keil, 2006) including the role of emotional stimuli as targets or non-targets, the broader task context (e.g., spatial detection versus perceptual identification tasks), and specific task instructions (e.g., speeded versus non-speeded processing).

---

[1] Task performance in the current context is conceptualized as either the speed (reaction times) or the accuracy (percent correct responses) of behavioral responses towards a defined class of experimental stimuli. Benefits/costs of affective processing are inferred from shorter RTs/higher accuracy or longer RTs/lower accuracy, respectively, observed for emotionally charged stimuli relative to neutral “baseline” stimuli.

1.3.1 Tasks Showing Benefits of Emotional Processing

Facilitated Responses in Spatial Attention Tasks

In an elegant application of the ***visual search paradigm***, Hansen and Hansen (1988) examined the speed of detecting an emotionally deviant target inside a spatial array of faces. As a key finding, participants were faster to detect photographs of angry faces inside crowds of happy faces than vice versa, which was called an anger-superiority-effect. More than 10 years later, Öhman. Lundqvist, and Esteves (2001) elaborated the initial paradigm of Hansen using i) larger search arrays, ii) physically controlled schematic faces, and iii) different emotional expressions (angry, happy, sad, and neutral). Results confirmed faster detection of angry faces, compared to happy and sad expressions, irrespective of whether the background crowd was composed of neutral or emotional items. Similar processing benefits were obtained for target pictures showing fear-relevant evolutionary threats (snakes and spiders compared to flowers and mushrooms, Öhman, Flykt, & Esteves, 2001), "modern" threats (e.g., Blanchette, 2006), and eye-regions of threatening faces (Fox & Damjanovic, 2006). Crucially, when display sizes are increased progressively, search times for angry faces grow but the search slopes are comparatively shallow (Eastwood, Smilek, & Merikle, 2001). Thus, threatening pictorial stimuli may not "pop out" in the traditional sense (which would be indicated by near zero search slopes), but may be detected by highly efficient parallel and preattentive search mechanism that scans for threatening events outside the spotlight of focused attention (Öhman, Flykt et al., 2001). An on-going debate centers on the boundary conditions of these effects. For instance, Harris, Pashler, and Coburn (2004) showed that emotionally charged words did not evoke faster detection in visual search arrays.

Results from visual search paradigms indicate that pictorial emotional stimuli can guide spatial attention. This interpretation is consistent with the ability of emotional stimuli to facilitate responses in ***spatial cueing*** or ***dot-probe paradigms***. Response latencies to hemifield targets were reduced when they are validly cued by preceding negative words (Stormark, Nordby, & Hugdahl, 1995) or fear-conditioned stimuli (Armony & Dolan, 2002; Stormark, 1999), relative to neutral cues. Similarly, when presented with vertically arranged word pairs (one neutral, one threatening), anxious participants responded faster to dot probes that replace the emotional word (Mogg, Bradley, de Bono, & Painter, 1997).

Facilitated Responses in Other Speeded Tasks

More sporadically, facilitatory effects of emotion on reaction times are found for responses not mediated by spatial attention, i.e., when stimuli are centrally presented as single exemplars. Such a genuine processing speed benefit might be inferred, for instance, from recent work by Gaillard and Tranel (2006) who showed that the identification of famous faces is faster for positive facial expressions than neutral expressions. Further, several studies reported that emotional content (positive and negative) expedited ***lexical decisions*** (Matthews, Pitcaithly, & Mann, 1995; Williamson, Harpur, & Hare, 1991). These findings are, however, at odds with other lexical decision studies reporting either null effects (MacKay et al., 2004) or slowing effects for negative material (Algom, Chajut, & Lev, 2004). A similar response time gradient with slower RTs to unpleasant relative neutral stimuli has been occasionally reported also in other ***non-evaluative choice reaction time tasks***, for example, when subjects discriminate newly presented from previously seen pictures (Bradley, Greenwald, Petry, & Lang, 1992) or pictures containing animals or humans (De Cesarei & Codispoti, 2006). Noteworthy, in the latter study a facilitation effect was found for pleasant pictures specifically.

Pertinent to the question of speed benefits associated with emotional stimuli, a widely held view posits that evaluative discrimination (i.e. the classification as either “good” or “bad”) is a quick and effortless process (Bargh, Chaiken, Govender, & Pratto, 1992; Dijksterhuis & Smith, 2002). This perspective has been derived to a great part from ***affective priming*** studies manipulating valence congruency between prime-target pairs. In the prototypical case, the evaluation of a probe stimulus (usually a word) is faster when it is preceded by an affectively congruent versus incongruent prime. Congruency effects occur for both supraliminal and subliminal (masked) primes (Greenwald, Klinger, & Liu, 1989) making a strong case for the automatic activation of evaluations (Fazio, Sanbonmatsu, Powell, & Kardes, 1986). As reviewed by Klauer and Musch (2003), reliable facilitatory effects, however, are restricted to prime-target stimulus-onset asynchronies (SOAs) below 300 ms and to evaluative tasks or tasks requiring simple affirmative responses (e.g., “Is it a word?”). In another typically evaluative task setup, the speed of ***evaluative categorization*** is compared in response to facial expression of positive versus negative affect. Interestingly, findings converge to show faster responses to positive or happy expressions of emotion, relative to negative expression including anger (Leppänen, Tenhunen, & Hietanen, 2003), sadness (Feyereisen, Malet, & Martin, 1986), and disgust (Stalans & Wedding, 1985).

A major drawback of the outlined priming and evaluation studies refers to the lack of a referential neutral baseline condition. They do thus not allow conclusions about behavioral effects as conceptualized in the current context, i.e., as processing benefits for emotional relative to non-emotional stimuli. Summarizing the above results, emotional stimuli appear capable to speed processes related to spatial attention, while reaction time studies using centrally presented stimuli have yielded inconsistent results. As illustrated below, emotional content facilitates responses more reliably under non-speeded task conditions.

Facilitated Responses in Accuracy[2] Tasks

In recent years, a growing number of studies have assessed effects of stimulus' emotional content in the context of ***Attentional Blink*** (AB) paradigms. The AB refers to a decrease in accuracy for reporting the second (T2) of two targets that are presented in close temporal vicinity during Rapid Serial Visual Presentation (RSVP). A variety of studies reported facilitated identification of emotional T2 stimuli during the AB period, relative to neutral stimuli (see also chapter 2.3). Enhanced report rates were consistently shown for emotional words (Anderson, 2005; Anderson & Phelps, 2001; Keil & Ihssen, 2004; Ogawa & Suzuki, 2004), for emotional faces (Mack, Pappas, Silverman, & Gay, 2002; Milders, Sahraie, Logan, & Donnellon, 2006), for phobia-related stimuli, and for positive and negative natural scene pictures (Trippe, Hewig, Heydel, Hecht, & Miltner, 2007). Several authors have emphasized that the AB reflects a late, post-perceptual, bottleneck related to working memory consolidation (e.g., Vogel, Luck, & Shapiro, 1998). In this view, perceptual enhancement related to emotional content may act to support the formation of working memory representations.

A facilitatory impact of emotional content on accuracy measures is also indicated by studies examining ***long-term memory***. Enhanced declarative memory for emotionally arousing material is a robust phenomenon showing improved performance for emotional pictures and words, incidentally or explicitly learned material, and at immediate or delayed (up to 1 year) retrieval conditions (e.g., Bradley et al., 1992). Work with brain-lesioned patients (Adolphs, Cahill, Schul, & Babinsky, 1997; Buchanan, Denburg, Tranel, & Adolphs, 2001) converge to show that the amygdala plays a key role in emotional memory enhancement. Several authors have examined the accuracy of responses when emotional and neutral stimuli are presented ***subliminally*** or ***near the threshold of conscious access***. For instance, Dijksterhuis and Aaarts (2003) showed that affective categorization of words presented for 13 ms with pre- and post-masks is more accurate for negative compared to positive words. In a related finding, negative backward-masked words elicited more accurate naming responses than neutral backward-masked words (Gaillard et al., 2006). Extending these findings, in a masking study by Zeelenberg and co-workers (Zeelenberg, Wagenmakers,

---

[2] Note that in the present work the term "accuracy" is used to denote performance in non-speeded tasks and does not refer to response accuracy in reaction time experiments.

& Rotteveel, 2006) enhanced recognition accuracy was found for both negative and positive words relative to neutral items. Similarly, recognition thresholds for briefly flashed words were lower (i.e., accuracy was higher) for affective versus neutral words, given that the stimuli were frequent words and their occurrence had been expected by the participants (Kitayama, 1990).

In accordance with the notion of enhanced accuracy under non-speeded task condition is a finding of Phelps and colleagues (Phelps, Ling, & Carrasco, 2006). Using centrally presented stimuli displayed at a SOA of 125 ms, they found that emotional events can facilitate subsequent visual processing. In an orientation discrimination task, ***contrast thresholds*** were lower (i.e., accuracy was higher) for grating patterns presented after fearful faces compared to gratings following neutral faces. Phelps and colleagues concluded that "emotion facilitates early vision: People actually see better in the presence of emotional stimuli" (p. 298).

1.3.2. Tasks Showing Costs of Emotional Processing

Impaired Responses in the Emotional Stroop Task

A well-replicated impairment effect related to affective content is the slowing of stimulus color naming in the emotional Stroop task. Relative to neutral words, reaction times are delayed when subjects are required to name the font color of for phobia-related (Watts, McKenna, Sharrock, & Trezise, 1986), negatively valenced (Pratto & John, 1991), personally relevant (Wingenfeld et al., 2006), and taboo words (MacKay & Ahmetzanov, 2005). Response impairment was also found for color naming of angry versus neutral schematic faces (Putman, Hermans, & van Honk, 2004). Two major theoretical accounts for the emotional Stroop effect have been proposed. In initial studies, parallels were drawn to the traditional Stroop effect assuming a conflict of selective attention that arises from automatic activation of word meaning. Increased attention to the meaning of emotional words leads to "intrusive cognitions" and a failure to attend to word color (McKenna & Sharma, 1995; Williams, Mathews, & MacLeod, 1996). The attention account, however, has been subject to several criticisms. Thus, McKenna and Sharma (2004) showed that slowing effects are observed reliably only in blocked conditions (consisting of either neutral or negative words) and not in randomly intermixed trials suggesting a "slow component" underlying the emotional Stroop. Doubts about a relation between the emotional and traditional Stroop have also been raised by Algom, Chajut, and Lev (2004). By showing that negative content also slows naming and lexical decision responses, he argued for a generic defense or inhibition mechanism that disrupts on-going cognitive activity in presence of possible threats. This account shares features to other notions assuming that humans are biologically prepared to respond to threat, e.g., “automatic vigilance” (Pratto & John, 1991) or the threat-detection mechanism proposed by Öhman (2001).

Impaired Responses in Spatial Attention Tasks

Emotional stimuli can impair the processing of other stimuli presented simultaneously but at different spatial locations in the perceptual field. For instance, judging the parity of two digits required more time when the digits flank a centrally presented emotional versus neutral word (Harris & Pashler, 2004). However, these disruptive effects were subject to rapid habituation and decreased when multiple (neutral and emotional) words were presented.

Evidence for emotional interference with spatial attention is also evident in affective variants of the visual search or spatial cueing task. Thus, participants responded slower to the absence of a discrepant face in crowds of either angry or happy faces (Eastwood, Smilek, & Merikle, 2003). Similarly, counting of features in schematic faces was slower for negative compared to neutral faces suggesting that attention was captured to the global face level particularly during processing of negative affect (Eastwood et al., 2003). Further, Fox and colleagues demonstrated that response latencies to hemifield targets were more strongly delayed when their location was (incorrectly) cued by an emotional versus neutral word that had appeared in the opposite hemifield (Fox, Russo, Bowles, & Dutton, 2001). Thus, affective cues do not only capture, i.e., facilitate the engage or shift component of spatial attention (Posner, 1980), but may also hold attention (i.e., modulate disengagement from emotional signals and increase attentional dwell time).

Impaired Responses in Temporal Attention Tasks

The assumption of increased dwelling on emotional signals is consistent with impaired processing of a target stimulus that is presented with a temporal gap but at the same spatial location as a briefly displayed emotional stimulus. Given that they were foveally presented, task-irrelevant threat words slowed lexical decisions to neutral probe words at a SOA of 300 ms (Calvo & Castillo, 2005). At similar SOAs, Schimmack and Derryberry (2005) documented slowing of reaction times contingent on picture arousal using a math problem and line localization task. Emotional costs for subsequent processing were also found in studies using non-speeded accuracy measures. Thus, detection accuracy for rotated landscape/architecture pictures amidst a RSVP stream was lower when the targets were preceded by a negative compared to a neutral distractor within a time window of 200 ms (Most, Chun, Widders, & Zald, 2005).

1.4 Synopsis and Overview of Studies

In daily life, emotionally laden visual objects appear endowed with special properties making them salient or “eye-catchers” in advertisements, mass media, and many other environmental contexts. Neuroscientific studies have accrued evidence that these stimuli indeed capture natural selective attention and enhance perceptual processes. As opposed to the coherent picture delineated by brain activity measures, behavioral task responses to emotional stimuli are less predictable. Intuitively, one might assume that attention capture of emotional stimuli leads to behavioral benefits when emotional stimuli serve the role of targets and to behavioral costs when they are presented as concurrent or competing non-target stimuli (or distractors[3]). Consistent with this, emotional targets are preferentially selected from temporal streams and spatial arrays of information, while emotional non-targets that are presented in spatial or temporal distance to a target stimulus can render its processing worse. However, often the borders between target versus non-target/distractor tasks are blurred indicating that this view may be too simplistic. Two examples are illustrated showing obvious and subtle problems inherent in a prediction based on target status solely. In a broader sense, these notions also guide the issues and hypotheses addressed in the empirical part of this thesis (chapter 2).

i) Except for visual search tasks, only few studies could verify a behavioral benefit for emotional targets at tasks entailing speeded responses, such as choice-reaction time experiments. Interestingly, pleasant but *not* unpleasant stimuli appear to be processed faster than neutral stimuli. Further, given that emotional content facilitates access to target lexicality or target identity, it is not clear why emotional content should not facilitate access to target color. Accordingly, based on the target status hypothesis, in the emotional Stroop task one might expect faster responses to negative words, relative to neutral items. Empirical evidence, however, clearly shows a reversed pattern, with slower color naming for negative stimuli. As indicated in the above chapters, for paradigms presenting emotional stimuli as targets, it might thus be useful to distinguish tasks according the dependent behavioral measures employed,

---

[3] Compared to the term “distractor”, the term “non-target” may be more appropriate as it does not imply any (detrimental) behavioral consequences.

namely reaction times versus non-speeded accuracy: Converging evidence suggests a facilitatory impact of emotion (irrespective of valence) on the accuracy of responses under non-speeded task conditions, such as working memory encoding, long-term retrieval, or identification of masked stimuli. In contrast, assessment of speed parameters (reaction times) during affective stimulus processing has yielded mixed results, and often results vary with hedonic valence. Reliable facilitation effects for aversive stimuli have been limited to tasks manipulating the distribution of spatial attention, while more complex processes, such as font color naming, are consistently delayed in presence of negative stimuli. Thus, the purpose of the first experimental series was to scrutinize affective influences on the speed of comparatively simple behavioral responses, such as sensory detection and visual categorization. To this end, behavioral and electrocortical parameters during different simple and choice reaction times tasks were recorded using emotional material as target stimuli.

ii) In the study by Phelps et al. (2006), emotional stimuli (faces) were presented as task-irrelevant items prior to the target stimuli (Gabor patches). As reviewed in the previous sections, centrally presented, emotional non-targets typically impair processing of subsequently presented stimuli. However, the opposite was true in the study by Phelps et al.: Relative to neutral faces, fearful faces facilitated subsequent performance. Following these notions, it remains unclear whether and how emotional stimuli modify processing of information in temporal proximity to the affective cue. Brain imaging work has shown that cortical responses to affective stimuli are amplified at various levels of stimulus analysis (see chapter 1.2). An outstanding issue, particularly on a neuronal level, refers to whether and how perceptual enhancement for arousing signals translates into modified processing of subsequent information. Experimental series B aimed to examine facilitation and interference effects of task-irrelevant emotional pictures on subsequent verbal processing. In the context of forced-choice lexical decision tasks, pictures varying in hedonic valence and emotional arousal preceded word/pseudoword targets. Dependent measures included lexical decision latencies and electrophysiological measures derived from event-related potentials and regional source modeling.

In the behavioral studies of series C, emotion-performance interactions were examined in a Rapid Serial Visual Presentation (RSVP) paradigm. Identification accuracy was measured for target words (T1 and T2) that varied in subjectively rated pleasure and arousal and were embedded in a rapid stream of neutral filler words. T2 report rates were analyzed in relation to emotional content of T1 or emotional content of a task-irrelevant picture distractor. By assessing T2 identification at varying temporal distances (lags) to T1 (or the distractor), the experimental setup allowed to track temporal unfolding of emotion-related attention capture.

Experimental series D explored affective response modulation in the context of automatic multisensory binding. Using the well-known McGurk illusion, two behavioral experiments are reported examining whether audiovisual fusion, i.e., integration of auditory and visual channels, can be altered by affective content.

2. EXPERIMENTAL SERIES

2.1 Emotional Processing and the Speed of Visual Categorization (Series A)

From a functional-evolutionary perspective, processing of emotional information must be both accurate and swift. As reviewed in the previous chapter, there is compelling evidence that emotional target stimuli are related to enhanced accuracy at behavioral tasks requiring delayed responses after each experimental trial (Anderson, 2005; Zeelenberg et al., 2006). In contrast, there is less conclusive data available addressing the impact of emotion on performance under speeded task conditions. Though, the time component is vital in situations of motivational relevance, particularly in defensive contexts such as the encounter with a predator (Öhman, Flykt, & Lundqvist, 2000).

Overview of the Studies of Series A

This series of experiments[4] set out to examine whether emotional content precipitates a reaction time benefit in visual information processing. In particular, we raised the issue whether emotional content alters the speed of comparatively simple behavioral responses related to i) sensory detection (i.e., responses to stimulus onset) and ii) two-choice categorization (i.e., discrimination along a simple non-evaluative dimension). To scrutinize these issues, stimuli varying in emotional arousal and hedonic valence were centrally presented as single targets in the context of several SRT and CRT tasks. We also intended to close the gap that has emerged between studies examining neuronal reflections of affective processing and studies relying on behavioral measures. To this end, we recorded both behavioral and neural parameter of information processing, comparing response speed and error rates with electrocortical patterns obtained from high-density electroencephalography (EEG).

Since our intention was to model effects of emotion during encoding of real-word situations, we opted to focus on natural scene pictures as target stimuli. Each photograph was presented only once during a given task. Thereby, the current paradigm emphasized processes related to processing novel information versus retrieving well-learned stimulus

---

[4] This chapter contains adapted sections of the manuscript entitled “Acceleration and Deceleration of Choice Reactions: How Emotional Content Alters the Speed of Visual Information Processing” by Niklas Ihssen and Andreas Keil, which is being prepared for publication.

representations, as in the case of word encoding. A second reason for focusing on emotional pictures was that, compared to affective words and faces, pictorial material is capable to elicit a higher degree of emotional arousal, both physiologically and in terms of subjective ratings (Lang, Bradley, & Cuthbert, 1998). Interestingly enough, examination of facilitation/interference in cognitive tasks has mostly relied on verbal material or emotional faces. This neglects the potential that emotional intensity and stimulus channel have a prominent impact on such effects. In Experiment 1 we addressed the issue of differences between emotional words and pictures explicitly by including a verbal SRT and a verbal CRT task.

Within each experiment, SRT and CRT tasks comprised the same set of emotional pictures (and words, respectively). Applying the rationale of traditional processing stage approaches (Donders, 1868; Ritter, Simson, & Vaughan, 1983; Sternberg, 1969), SRT latencies reflect the sum of three processes: sensory detection, motor preparation, and motor execution. In contrast, CRTs additionally involve the operation of a discrimination (and/or response selection) mechanism that maps stimulus codes to response codes. Thus, any difference between SRTs and CRTs should reflect emotional effects on a cognitive stage of processing rather than modulation of motor processes. This hypothesis was scrutinized in Experiment 4, where we measured event-related potentials (ERPs). By this means, emotion effects could be tracked at various processing stages and with high temporal resolution.

General Methods

Selection of Standardized Emotional Pictures

Pictorial stimuli were selected from the International Affective Picture System IAPS (Lang et al., 2005). The IAPS comprises normative affective ratings of more than 1200 colored photographs and covers a wide range in terms of arousal and valence ratings allowing for systematic manipulation of these dimensions. For each experiment, different pictures were selected from the IAPS, thus testing the replicability of possible emotion effects across stimulus sets. In Experiment 1 to 3, we compared SRT/CRT responses towards pictures rated as either pleasant, neutral, or unpleasant. To elucidate possible interactive effects of valence and arousal on processing speed, Experiment 4 included systematic arousal variations within the pleasant and unpleasant categories. Here, low-arousing negative/positive pictures were contrasted with high-arousing negative/positive stimuli. In all experiments, each emotional category contained heterogeneous content, thus avoiding salient differences related to semantic classes (e.g., erotica). As best as possible, emotional categories were also matched for stimulus complexity (e.g., single versus multiple objects etc.). Further, two experiments (3 and 4) exclusively included pictures of humans, controlling for differences in person versus non-person object encoding. Similar to previous categorization studies (e.g., VanRullen & Thorpe, 2001), pictures were presented briefly (100 ms). A short exposure time helps to minimize the effects of exploratory or emotion-related (e.g., avoiding fixation of highly aversive content) eye movements.

The Picture Versus Non-Picture Task and the Number-of-Persons Task

We employed two different versions of CRT paradigms, thus ensuring the task independence of possible effects: a two-choice picture versus non-picture discrimination (Experiments 1 -3), and a two-choice number-of-persons discrimination task (Experiment 3). The picture versus non-picture task was developed as a pictorial analogue to the lexical decision paradigm that composes the prototypical choice reaction task in the verbal domain. For each picture, we constructed a (distorted) non-picture exemplar using a standardized image filtering procedure. Pictures were submitted to an algorithm implemented in the open-access GIMP software that interprets the cartesian coordinates of each pixel as polar coordinates and generates a planar image of this projection. Compared to pixel-wise randomization, coordinates transformation has the advantage that the distorted picture contains no artificial grid-patterns which would make the picture easily discriminable from non-distorted pictures. In the resulting distorted image, object information is no longer extractable (see figure 1 for two example pictures). In contrast, basic physical properties are preserved as the distorted picture comprises the same color and luminance range as the plain image. Task performance on non-pictures, grouped for affective categories underlying the source images, thus represents a control condition for testing effects related to physical differences between picture categories. The possibility to control for category differences associated with physical image parameters was particularly valuable during recording of event-related potential in Experiment 4. In the second version of the CRT task, which was applied in Experiment 3, participants had to decide whether the depicted scene comprised one or two persons. In addition to the CRT task, each participant performed a SRT task, requiring responses contingent on stimulus onset but involving the same material as in the CRT task.

EXPERIMENT 1 - SRTS AND CRTS TO EMOTIONAL PICTURES AND WORDS

Four RT tasks, completed by each participant, tested response speed and error rates as a function of target's emotional content: i) SRTs and CRTs (lexical decisions) to pleasant, neutral, and unpleasant words, and ii) SRTs and CRTs (picture versus non-picture discriminations) to pleasant, neutral, and unpleasant natural scene pictures.

Methods

Participants

22 university students (16 females) with a mean age of 22.7 years (Standard Deviation SD = 3.0) consented to participate in Experiment 1. They all were native speakers of German and reported normal or corrected-to-normal visual acuity. Participants received course credits or a financial bonus.

Stimuli

For the pictorial tasks, 20 pleasant (p; including cheery kids or adults, family scenes, erotica), 20 neutral (n; including neutral kids or adults, household scenery), and 20 unpleasant (u; including animals [spiders, snakes, sharks], human attack scenes, mutilated bodies) colored images were selected. According to the normative 9-point scale IAPS ratings (1 = unpleasant, 9 = pleasant), mean valence was 7.31 (SD = 0.57) for pleasant pictures, 4.87 (0.24) for neutral pictures, and 2.29 (0.84) for unpleasant pictures. Regarding judgments of emotional arousal (1 = calm, 9 = excited), both pleasant and unpleasant pictures were more arousing than neutral images (mean arousal p: 5.44 [SD = 1.08], u: 6.85 [0.31], n: 2.82 [0.59]). 60 non-pictures were generated for the CRT task by transforming each picture to polar coordinates as described above. Pictures size was 7.1 x 5.3 cm.

Based on pre-experimental ratings obtained from a different sample (N = 41, see Keil & Ihssen, 2004), 90 German verbs were selected for the verbal tasks. 30 pleasant (e.g., “to kiss”, mean valence: 7.87, [SD = 0.66], mean arousal: 7.06, [SD = 0.72]), 30 neutral (e.g., “to note”, 4.99 [0.33], 2.61 [0.54]), and 30 unpleasant (e.g., “to kill”, 1.59 [0.37], 7.66 [0.49]). Affective categories were carefully matched for phonological word length (number of syllables p: 2.67, n: 2.50, u: 2.60), mean number of letters (p: 8.17, n: 8.10, u: 8.33), and word (lemma) frequency (p: 163.9, n: 150.2, u: 143.1) according the CELEX (Baayen, Piepenbrock, & Gulikers, 1995) database. For the lexical decision task, 90 pronounceable, arbitrary pseudowords were generated by permuting syllables and letters of each word but preserving the word ending (“-en”), which defines the verb infinitive in German language. The 180 white-colored words and pseudowords were presented in 26-point Helvetica font.

Design and Procedure

Participants were seated in a sound-attenuated and dimly lit experimental chamber. They viewed the stimuli on a 19’’ computer screen. All stimuli were centrally presented against a black background. A chin rest located 70 cm in front of the monitor helped to standardize visual angle, which was 5.8° horizontally and 4.3° vertically for pictures, and 0.6° vertically for words, respectively. The order of the four tasks was pseudo-randomized. Between the tasks, participants had the possibility to rest. Prior to each task, participants were instructed and performed 10 practice trials (5 for the lexical decision) comprising stimuli from all conditions of the respective experiment but not used in the test trials

Stimuli were not repeated within a given task and presented in random order. During each trial, the target stimulus was displayed for 6 refresh cycles on a 60 Hz monitor, resulting in 100 ms stimulus duration. The next trial started after a randomly varying interval showing a black screen and a white fixation cross that was included to warrant centrally oriented gaze direction. Inter-Trial-Intervals (ITIs) had a minimum duration of 2000 ms plus an additional time determined by an exponentially decreasing random function that had a mean of 2500 ms and a maximum of 12.5 seconds.

Participants executed CRTs with their right index and middle finger using two adjacent buttons on the computer keyboard (ALTERNATE and CONTROL). They were instructed to indicate as quickly and accurately as possible whether the target stimulus was a normal picture (a real word in German) or a distorted non-picture (a nonword). Button assignments were counterbalanced across subjects. SRTs were performed by pressing the spacebar simultaneously with the left and right index finger.

Data Analysis

Participants’ speed of choice reactions was calculated for each condition as the mean RT of correct responses that occurred in a time segment between 200 ms and 1000 ms following target onset. We selected a specific cutoff value (1000 ms) for dealing with RT outliers instead of applying thresholds based on individual RT distributions (SD trimming, Windzorizing etc.). With regard to statistical power, fixed thresholds have been proven equally or even superior, compared to variable criteria (Ratcliff, 1993). To assess response accuracy, we measured error rate as the percentage of wrong responses, missed targets, or response occurring later than 1000 ms. For the SRT tasks, mean RTs were recorded for responses falling in a time window between 150 ms and 1000 ms after target onset. The dependent measures were submitted to repeated measurement Analysis of Variance (ANOVA), for each task separately. CRTs to pictures were evaluated by two-factorial ANOVA crossing the factors picture distortion (2; picture versus non-picture) and emotional content (3; pleasant, neutral, unpleasant). SRTs/CRTs to words and SRTs to pictures were analyzed using one-factorial ANOVA with emotional content (3; pleasant, neutral, unpleasant) as within-subject factor. Significant effects were followed by post-hoc ANOVAs and Newman-Keuls tests.

Results

Responses to Words

Words' emotional content did not affect simple or choice reaction times, Fs < 1.0 (see Table 1). ANOVA on CRT error rates showed a significant main effect of emotional content, F(2, 42) = 4.8, p < 0.5. Participants made fewer errors with pleasant, Newman-Keuls, p < 0.05, and unpleasant words, p < 0.05, compared with neutral items.

Table 1. Mean RTs and error rates (standard errors) of simple and choice responses (lexical decisions) to verbal stimuli in experiment 1.

	Pleasant Words	Neutral Words	Unpleasant Words
SRTs (ms)	307 (9)	310 (10)	310 (9)
CRTs (ms)	652 (15)	660 (16)	654 (14)
CRT error rates (%)	14.4 (2.7)	20.9 (3.7)	15.8 (3.1)

Table 2. Mean SRTs (ms) and standard errors to pictures as a function of emotional content.

	Pleasant	Neutral	Unpleasant
Experiment 1	325 (11)	320 (10)	324 (12)
Experiment 2a	308 (16)	305 (16)	302 (16)
Experiment 2b	278 (14)	284 (16)	279 (15)
Experiment 3	309 (9)	304 (9)	303 (8)
Experiment 4	PL: 319 (8), PH: 318 (10)	321 (10)	UL: 318 (11), UH: 317 (9)

Note. PL: pleasant low-arousing, PH: pleasant high-arousing, UL: unpleasant low-arousing, UH: unpleasant high-arousing.

Responses to Pictures

SRTs to pictorial stimuli did not vary with stimulus' emotional content, F(2, 42) = 1.1 (see Table 2). ANOVA on CRTs yielded a main effect of picture distortion, F(1, 21) = 17.6, $p < 0.001$, reflecting slower responses for non-pictures than for pictures (see Figure 1). Further, CRTs were modulated by emotional content, F(2, 42) = 15.2, $p < 0.0001$, and this modulation varied as a function of picture distortion, emotional content x picture distortion: F(2, 42) = 15.1, $p < 0.0001$. Simple effects ANOVA revealed that emotional content modified CRTs to pictures, F(2, 42) = 29.9, $p < 0.0001$, but not to non-pictures, $F < 1.0$. As illustrated in Figure 1, participants responded faster to pleasant compared to neutral pictures, Newman-Keuls $p < 0.05$. At the same time, responses were delayed for unpleasant compared to neutral material, $p < 0.001$.

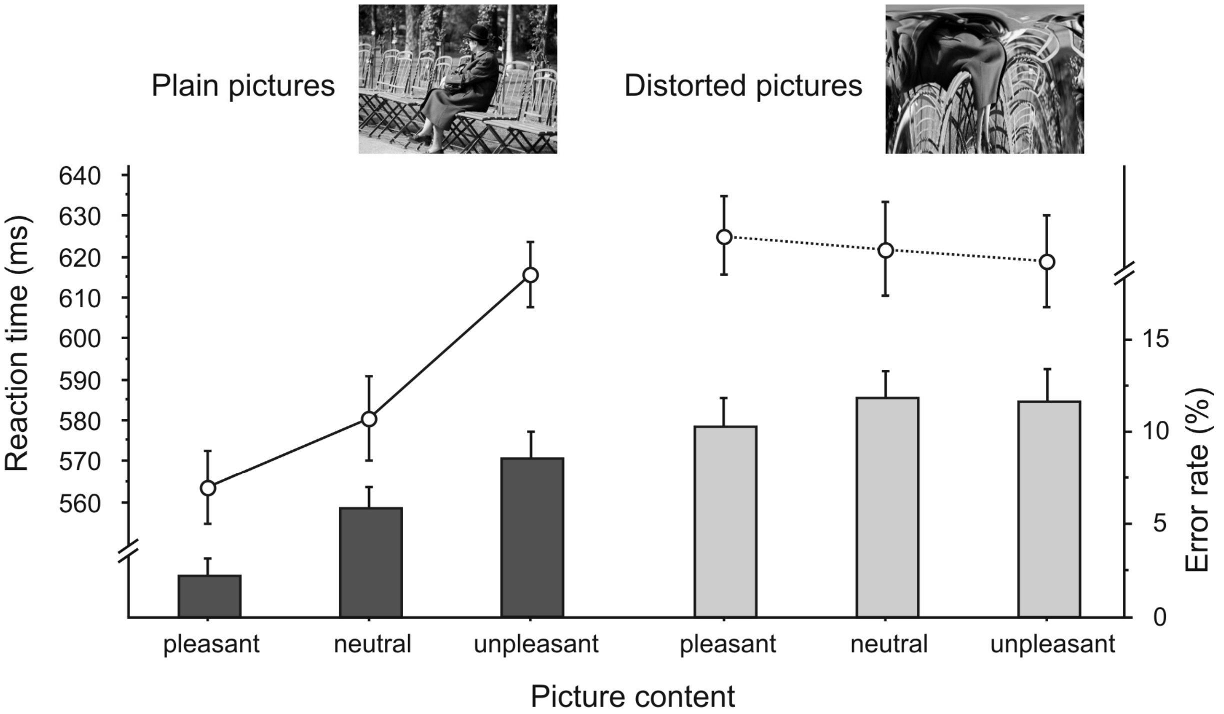

Figure 1. Mean CRTs and error rates as a function of picture distortion and emotional content in the picture versus non-picture discrimination task of Experiment 1. Error bars indicate standard errors. Top: Two example stimuli showing a neutral picture and its corresponding non-picture analogue created by cartesian-to-polar coordinates transformation. To maintain the research value of the IAPS, an example picture was chosen not contained in the IAPS collection.

Analysis of error rates showed that distorted pictures were more often categorized wrongly than plain stimuli, F(1, 21) = 20.8, $p < 0.001$. ANOVA also yielded a main effect of emotional content, F(2, 42) = 6.2, $p < 0.01$. Post-hoc testing demonstrated that pleasant stimuli elicited less errors compared to neutral, $p < 0.05$, and unpleasant items, $p < 0.01$. The interaction between emotional content and picture distortion did not reach significance, $F(2, 42) = 1.9$.

Discussion

CRTs to pictures were substantially delayed for aversive stimuli compared to the neutral baseline condition. In contrast, appetitive picture content accelerated choice responses. This pattern was mirrored by error rates showing an increase of incorrect responses from pleasant to neutral and unpleasant stimuli. Defensive activation appears to draw resources away from the discrimination task leading to pronounced response delays. Showing opposite effects, pleasant pictorial content seems capable to speed discrimination and/or response selection processes. Similar effects of pleasant content have been reported in research examining the efficiency of face evaluation. In most cases, however, a happy face advantage was shown in relation to RTs towards negative and not to neutral faces (Leppänen et al., 2003), confounding possible delays for unpleasant stimuli with facilitation effects for pleasant stimuli (see chapter 1.3.1).

Crucially, emotional content did not alter non-picture CRTs. Thus, affective modulation was not confounded with physical stimulus properties (e.g., luminance differences). This assumption is also supported by the absence of emotional content on SRTs, which are sensitive to variations of stimulus intensity only (Ulrich, Rinkenauer, & Miller, 1998). In general, it is assumed that visual SRTs are accomplished by direct transmittance of sensory activation in primary visual cortex to motor output stages (Miller & Ulrich, 2003). Any SRT modulation by emotional content would thus most likely index effects on response initiation or execution. Such effects of emotional signals on motor processes have been reported, for instance, by Derryberry (1991), who found that negative information can delay response execution. This inhibition was held to indicate the operation of a behavioral inhibition system (BIS), which suppresses already selected motor programs. In the present study, absent SRT modulation thus makes a good case that the CRT effects reside at the level of cognitive (discrimination and response selection) rather than motor processing.

As opposed to picture CRTs, choice responses to words were not reliably modulated by affective content showing moderate effects only for error rates. Two aspects are likely to contribute to this divergence of results (see Keil, 2006): Different temporal activation patterns in verbal versus pictorial processing and different degrees of emotional arousal. Apparently, words elicited similar arousal ratings than pictures (see methods section). However, it should be noted that the relationship between subjective rating and other measures of emotional arousal may not be linear. In addition, the normative picture scores as well as the pre-experimental word ratings reflect relative judgments for a given item within its stimulus class and thus are not directly comparable. An emotional picture (showing, e.g., a dead body) will most likely elicit a higher degree of arousal as a corresponding visually presented word (e.g., “death”). This is corroborated by work showing a pronounced physiological impact (e.g., heart rate skin conductance changes) for pictures rated as arousing (Lang, Bradley, & Cuthbert, 1998), while such a relationship is less pronounced for linguistic material. Further, complex colored photographs are more closely related to real-world situations and allow a more immediate access to object representations, even at short viewing times (Keil, 2006). Several groups (Codispoti et al., 2006; Potter, 1975; Thorpe, Fize, & Marlot, 1996) have shown that pictures scenes can be categorized remarkably quickly (e.g., contains an animal, contains no animal). ERP recordings during natural scene perception show that the brain differentiates between semantic categories as early as 80 ms (VanRullen & Thorpe, 2001) and between target and nontarget scenes within 150 ms (Thorpe et al., 1996), pointing to “ultrarapid visual categorization” of natural scenes (Fabre-Thorpe, Delorme, Marlot, & Thorpe, 2001), which may include fast activation of emotional features.

One possible account for the modulation of verbal error rates is that low levels of emotional arousal, conveyed by both pleasant and unpleasant words, have an alerting function resulting in small benefits during the lexical decision task. This is in agreement with the relative weak facilitation effects during other speeded tasks with emotional words (Matthews et al., 1995). Future work may address the question whether, for instance, phobia-related, taboo (curses, sexual colloquials), or personally relevant words, which may evoke a higher degree of arousal, are capable to modulate CRTs more strongly. This would be predicted from the emotional Stroop task, where RTs are most robustly delayed for such stimuli (MacKay & Ahmetzanov, 2005; Wingenfeld et al., 2006).

The emotional Stroop modulation shows similarity to CRT slowing for negative pictures. Both effects may arise because negative stimuli capture attention and interfere with concurrent processes (Pratto & John, 1991). In the Stroop paradigm, negative distractors are typically presented until a response occurs. This setup thus readily allows that selective attention is directed away from the target task. In turn, facilitatory effects on discriminative processes may be more likely elicited when conscious stimulus analysis is reduced (Dijksterhuis & Aarts, 2003). Following these conjectures, a second experiment was carried out, where we reduced conscious stimulus processing by using shorter stimulus durations and a pattern mask. Hereby, we aimed to reduce the effect of attention-capture mechanisms, which may have contributed to CRT slowing for negative pictures.

EXPERIMENT 2 – SRTS AND CRTS TO MASKED EMOTIONAL PICTURES

Tapping perceptual processes more rigorously, Experiment 2 examined SRTs and CRTs to emotional pictures for which visibility was reduced by backward masking. Methodologically, we thus came closer to the masking and RSVP studies detailed above, which reported facilitation for emotional stimuli under perceptually degraded viewing conditions (Dijksterhuis & Aarts, 2003; Gaillard et al., 2006; Zeelenberg et al., 2006). Specifically, we asked whether the findings of Experiment 1, i.e., CRT delay for aversive and CRT acceleration for pleasant pictures, would be replicable for masked stimuli presented as short as 83 (Experiment 2a) or 50 ms (Experiment 2b).

Methods

Participants

Data for Experiment 2a were obtained from 16 university students (11 females, mean age 23.6 years, SD = 4.9). Additional 8 students (6 females, mean age 22.4 years, SD = 1.8) participated in Experiment 2b. All respondents reported normal or corrected-to-normal vision, and none of them had been participating in Experiment 1. Participants received class credits or were paid.

Stimuli

60 pictures (20 pleasant, 20 neutral, 20 unpleasant) were drawn from the IAPS that had not been used in Experiment 1. Pleasant items (mean valence 7.30, SD = 0.53, mean arousal 5.61, SD = 0.83) depicted, for instance, babies, playing kids, happy adults, or erotic scenes. Unpleasant pictures (valence: 2.51, SD = 0.84; arousal: 6.29, SD = 0.59) showed dangerous animals, attack scenes, starving childs, medical interventions, and mutilated bodies (N = 10 items in each subset). Neutral pictures (valence 4.79, SD = 0.67; arousal: 3.36, SD = 0.52) comprised, among others, neutral kids and adults, animals, urban and household scenes. For the choice RT task, 60 non-pictures were generated using cartesian-to-polar coordinates transformation as described above.

Design and Procedure

The procedure and experimental design were the same as in Experiment 1, with the following exceptions: i) We assessed only responses to pictures (SRTs and CRTs), with task order counterbalanced across participants. ii) Pictures were presented for 83 ms (Experiment 2a), and 50 ms (Experiment 2b), respectively, followed by a backward pattern mask displayed for 200 ms. The colored pattern mask was generated by superimposing image parameters derived from principal component analysis (PCA) of each of the 120 pictures.

Data Analysis

Response speed and error rates were evaluated for Experiment 2a and 2b separately, using the factorial design described in Experiment 1. Effects of emotional content were further analyzed by planned comparisons (linear contrasts). Based upon the findings of Experiment 1, lower RTs and error rates were predicted for pleasant compared to neutral pictures. Further, we hypothesized higher RTs and error rates for unpleasant compared to neutral pictures.

Results

Experiment 2a – 83 ms Stimulus Duration

Results replicated the basic findings of Experiment 1. SRTs (see Table 2) were not sensitive to emotional content, $F(2, 30) = 1.3$, while ANOVA on CRTs showed a main effect of emotional content, $F(2, 30) = 5.2$, $p < 0.05$, and an interaction between picture distortion and emotional content, $F(2, 30) = 8.1$, $p < 0.01$. The main effect of picture distortion was not significant, $F(1, 15) = 1.2$. Planned comparisons confirmed that choice RTs were faster for pleasant pictures, $F(1, 15) = 9.6$, $p < 0.01$, and slower for unpleasant pictures, $F(1, 15) = 4.6$, $p < 0.05$, compared to the neutral category (see Figure 2a).

Paralleling CRT results, ANOVA on error rates yielded significant effects of emotional content, $F(2, 30) = 4.7$, $p < 0.05$, and emotional content x picture distortion, $F(2, 30) = 13.0$, $p < 0.0001$, but no effect of picture distortion, $F(1, 15) = 3.1$, $p = 0.1$. Linear contrast analysis showed that response accuracy decreased from neutral to unpleasant pictures, $F(1, 15) = 14.8$, $p < 0.01$. Error rates were no different for pleasant versus neutral pictures, $F(1, 15) < 1.0$.

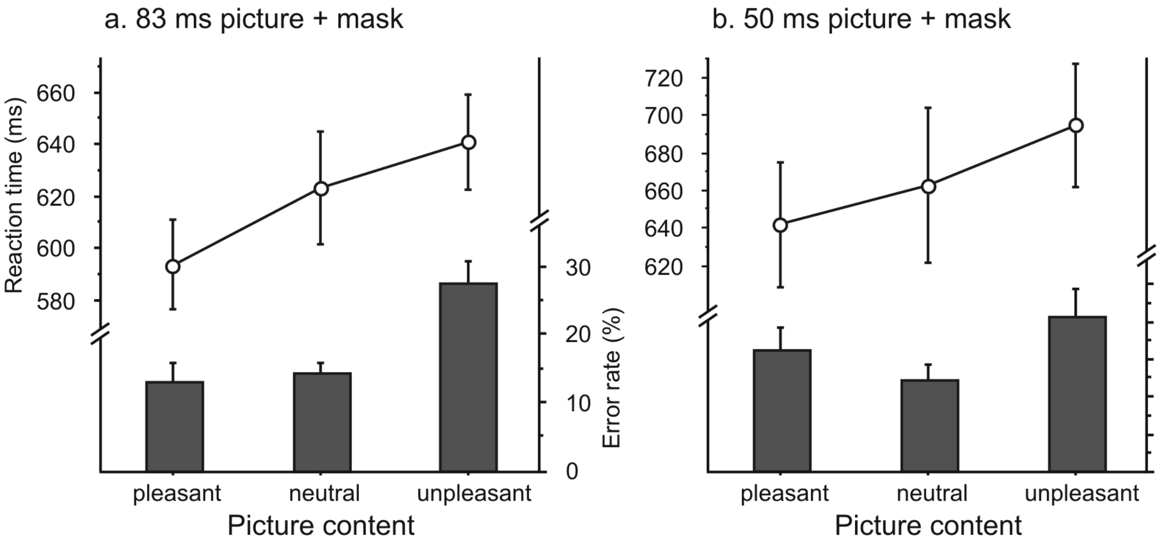

Figure 2. Mean CRTs and error rates for pleasant, neutral, and unpleasant pictures that were backward-masked and presented 83 ms (left panel, Experiment 2a) or 50 ms (right panel, Experiment 2b) in the picture versus non-picture paradigm. Error bars indicate standard errors.

Experiment 2b – 50 ms Stimulus Duration

Again, emotional content altered CRTs, F(2, 14) = 3.5, p = 0.06, emotional content x picture distortion: F(2, 14) = 4.6, p < 0.05, but not SRTs, emotional content: F(2, 14) = 1.1. ANOVA also showed slower CRTs to distorted than plain pictures, picture distortion: F(1, 7) = 14.2, p < 0.01. Although the basic direction of emotional effects on picture CRTs was preserved (i.e., an RT increase from pleasant to neutral and unpleasant pictures, see Figure 2b, right panel), effects were weaker in Experiment 2b, yielding only marginally significant contrasts, unpleasant versus neutral: F(1, 7) = 4.8, p = 0.06; pleasant versus neutral: F(1, 7) = 2.7, p = 0.15.

Regarding error rates, ANOVA showed no main effects of emotional content, F(2, 14) = 2.1, or picture distortion, F < 1.0, but a significant interaction between both factors, F(2, 14) = 6.1, p < 0.05. Contrast analysis showed that response accuracy was lower for unpleasant compared to neutral pictures, F(1, 7) = 9.1, p < 0.05.

Discussion

Reducing stimulus duration and masking did not abolish the effects observed in the previous study. Comparison of the overall RT level between Experiment 1 and 2 suggests that the difficulty of discrimination increased with decreasing exposure time, while SRTs were immune to this manipulation (see Table 2 and Figure 2). Crucially, again SRTs did not vary with emotional content, whereas choice responses were facilitated for pleasant pictures and impeded for unpleasant stimuli, although this pattern was weaker in the 50-ms condition. As a main conclusion, RT acceleration for pleasant and RT delay for unpleasant does not require fine-grained stimulus analysis. Rather, a relative small amount of stimulus information appears sufficient to evoke such speed effects. Accordingly, emotional facilitation/interference as observed here is fast and automatic and may be triggered before stimulus representations are fully activated.

Consistent with this view, Evans and Treisman (2005) have suggested that natural scene categorization involves a first stage of rapid and parallel detection of disjunctive target features having intermediate complexity. These features are sufficient to discriminate the scene along the target dimension, even though a comprehensive stimulus representation, where features have been bound by focused attention, is not available yet. Thus, based on the detection of few but distinctive appetitive features (e.g., a smiling mouth), positive affect may act to optimize visual categorization along any (evaluative and non-evaluative) stimulus dimension, yielding accelerated responses. In contrast, detection of salient aversive or threatening content may immediately disrupt all concurrent processing resulting in a state of “cognitive freezing”. Such a disruptive role of negative affect that operates on a pre-attentive level has been recently proposed by Algom, Chajut, and Lev (2004). Arguing in favor of a generic slowdown in presence of negative stimuli, the idea was challenged that the emotional Stroop reflects a conflict of selective attention, akin to the mechanisms that underlie the traditional Stroop effect. Experiment 2 indicates that CRT threat delay may share more similarities to a disruptive or inhibitory effect than to attention capture in the traditional sense.

One limitation of the previous experiments is the possible task dependence of the effects. Thus, it may be argued that some of the aversive pictures used here (mutilation and injury) intrinsically contain a larger amount of distortion information making the picture versus non-picture discrimination more difficult. Although his argument may also apply to pictures showing explicit sexual content, which did not delay response latencies but rather facilitated CRTs, a third experiment was conducted to examine these potential confounds.

EXPERIMENT 3 – CRTS TO EMOTIONAL PICTURES SHOWING ONE VERSUS TWO PERSONS

To test the replicability of emotional response modulation across different task contexts, we measured CRTs to pictures showing either one or two human agents. Compared to other content-based categorization paradigms (e.g., man-made vs. nature-made, living vs.

non-living), this target dimension can be easily crossed with an emotional dimension. Again, participants also performed SRTs using the same pictures as in the CRT task.

Methods

Participants

Informed consent for participation was obtained from 16 university students (8 females, mean age 23.9 years, SD = 3.0). All participants reported normal or corrected-to-normal visual acuity. For participation, they received course credits or a financial bonus.

Stimuli

Participants viewed 60 pictures showing one person and 60 pictures displaying two persons. Each target class comprised 20 pleasant, 20 neutral, and 20 unpleasant items showing in equal proportion persons oriented to observer (e.g., portraits of either nudes, neutral persons, or wounded persons) and persons in action (e.g., leisure versus office versus attack scenes). To meet these criteria, pictures were not only drawn from the IAPS (Lang et al., 2005) but also from an in-house data base containing pleasure and arousal ratings (N = 45) of 216 pictures that were found on various web sites and digital image collections. Mean valence scores were 6.96 (SD = 0.76) for pleasant pictures, 5.08 (0.54) for neutral pictures, and 2.38 (0.64) for unpleasant pictures. Mean arousal was 4.97 (0.77) for pleasant pictures, 5.92 (0.90) for unpleasant pictures and 3.01 (0.50) for neutral pictures. Pictures size was 10.6 x 7.9 cm (8.6° x 6.4° visual angle).

Design and Procedure

Participants were instructed and familiarized with the SRT and CRT task similarly to the previous experiments. Task order was balanced across respondents. During the SRT and CRT task, each of the 120 pictures (60 one-person-items, 60 two-person-items) was presented once in randomized order. Stimulus timing and ITI were identical to the picture versus non-picture paradigm in Experiment 1. In the CRT task, participants were required to indicate by button press whether the picture contained one or two persons. The same response modalities as in the previous experiments were used.

Data Analysis

RTs were analyzed for responses registered between 200 ms (150 ms for simple RTs) and 2000 ms after stimulus onset. Compared to the previous studies, a less restrictive CRT cutoff was used to accommodate possibly higher processing demands in the present task. Error rates were computed as the percentages of wrong responses, misses and responses above 2000 ms. SRTs, CRTs and error rates entered separate two-factorial ANOVAs crossing the within-subject factors target category (2; one- vs. two-persons) and emotional content (3; p, n, u). Similar to Experiment 2, CRTs were further analyzed by linear contrasts. Across target categories, we predicted an increase of RTs and error rates from pleasant to neutral and from neutral to unpleasant pictures.

Results

Neither target category (one-person vs. two-persons) nor emotional content affected SRTs, all Fs < 1.0 (see Table 2). ANOVA on CRTs showed a main effect of emotional content, F(2, 30) = 11.8, p < 0.001, but no significant effect involving target category, Fs < 1.2. Linear contrasts confirmed that across target categories, RTs increased from pleasant to neutral pictures, F(1, 15) = 6.1, p < 0.05, and from neutral to unpleasant pictures, F(1, 15) = 10.0, p < 0.01 (see Figure 3).

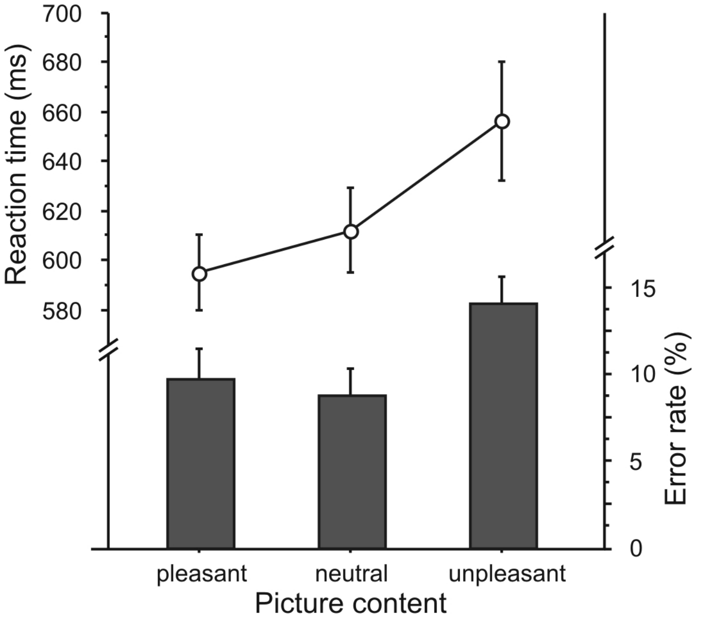

Figure 3. Experiment 3: Mean CRTs and error rates for pleasant, neutral, and unpleasant pictures obtained in the number-of-persons categorization task. Error bars indicate standard errors.

Paralleling RTs, error rates in the discrimination task were modulated by emotional content, F(2, 30) = 5.4, p < 0.05, whereas there were no effects of target category, F(1, 15) = 2.2, p = 0.16, or target category x emotional content, F < 1.0. Contrast analysis confirmed lower response accuracy for unpleasant compared to neutral pictures, $F(1, 15) = 8.5$, $p < 0.05$. There were no error rate differences between pleasant and neutral pictures, $F < 1.0$.

Discussion

Results of Experiment 3 validate that response facilitation for pleasant, and response impairment for unpleasant stimuli, respectively, is not specific to a task requiring speeded discrimination of plain versus distorted pictures. Instead, both the direction and magnitude of emotional modulation was replicated for choice reactions between pictures showing either one versus two human agents. Thus, accelerated CRTs to pleasant and delayed CRTs to unpleasant pictures appear to reflect a robust phenomenon of emotion-performance interactions, no matter whether holistic picture versus non-picture or content-based discrimination responses are considered.

In sum, the findings of Experiments 1 to 3 indicate that emotional content exerts a prominent impact on the speed of information processing. We have demonstrated that the requirement to categorize a target rapidly yields benefits for appetitive stimuli but detriments for aversive material. These diverging effects for pleasant and unpleasant stimuli are not predictable from electrophysiological findings. Typically, ERP studies show similar modulation of electrocortical responses for both arousing pleasant and arousing unpleasant pictures, relative to neutral stimuli. Between 200 and 250 ms these stimulus categories elicit a negative amplitude shift over occipitotemporal recording sites (Junghöfer et al., 2001; Schupp et al., 2003; Schupp, Stockburger, Codispoti et al., 2007), followed by enhancement of late positive potentials. It is assumed that these modulations reflect increased allocation of processing resources to motivationally significant stimuli at different levels of analysis. However, posterior negativities and enhanced late positive potentials have been typically reported for non-speeded task conditions. Thus, the question arises whether speeded categorization may be related to other (or additional) electrocortical modulations mirroring the diverging behavioral effects of pleasant versus unpleasant content in the present studies.

EXPERIMENT 4 – CHOICE REACTIONS AND ELECTROCORTICAL RESPONSES TO EMOTIONAL PICTURES

Seeking to reveal electrocortical correlates of the observed acceleration/deceleration effects, we measured event-related potentials (ERPs) during the picture SRT and picture versus non-picture CRT task. ERPs afford spatio-temporal information about macroscopic brain dynamics. They are thus an expedient approach to track information processing at high temporal resolution. Here, emotional effects on electrocortical responses were assessed by two means: First, we obtained amplitude measures of ERP deflections that index information processing at different stages of analysis (e.g., the P1 component). Secondly, we measured latencies of a late parietal positive potential (LPP). In Experiment 4, we also aimed to disentangle the role of emotional arousal and hedonic valence during speeded discrimination. To this end, emotional arousal was varied within the aversive and appetitive categories, i.e., responses to low-arousing pictures were contrasted with responses to high-arousing items having similar valence scores.

Methods

Participants

Behavioral data and ERPs were obtained from 18 university students (12 females, mean age 22.6 years, SD = 3.1), who gave written informed consent to participate in the study. According to scores on the Edinburgh Handedness Inventory (Oldfield, 1971), 17 participants were right-handed, one subject was left-handed. Participants reported no history of neurological or psychiatric disease. They received class credits or a financial bonus for participation.

Stimuli

Five categories of pictures were drawn from the IAPS (N = 35 items in each category): (i) High-arousing pleasant (PH; mean valence: 6.62 [SD = 0.58], mean arousal: 6.39 [0.36]), (ii) low-arousing pleasant (PH; 7.12 [0.56], 4.09 [0.53], (iii) low-arousing neutral (NL; 5.00 [0.35], 3.31 [0.47]), (iv) low-arousing unpleasant (UL; 2.45 [0.37], 5.21 [0.56]), and (v) high-arousing unpleasant pictures (UH; 1.94 [0.40], 6.83 [0.26]). Again, affective categories did not represent homogenous and salient semantic classes. Figure 4 (left panel) illustrates the distribution of the stimuli within an affective space, spanned by the dimensions of hedonic valence and emotional arousal. As can be seen, the five affective categories covered distinctive, i.e., non-overlapping areas. For each of the 175 pictures, a non-picture analogue was created using polar coordinates transformation (see methods section Experiment 1). Again, picture size was 10.6 x 7.9 cm (8.6° x 6.4° visual angle).

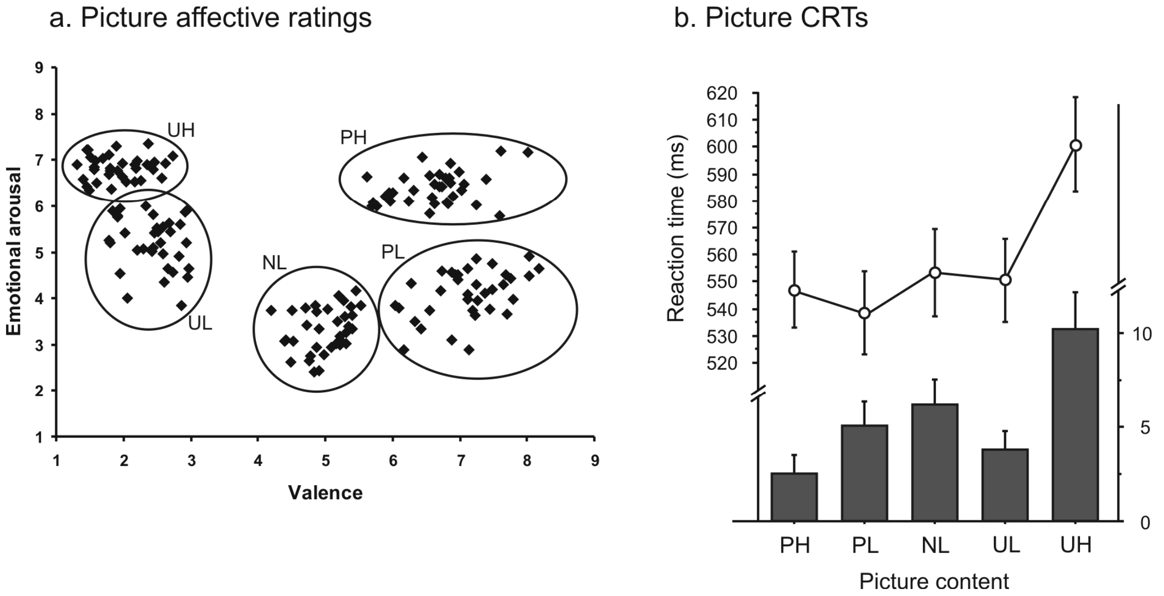

Figure 4. Experiment 4: a. Distribution of the selected picture stimuli in a two-dimensional affective space. Values correspond to normative valence and arousal ratings (PH = pleasant high-arousing, PL = pleasant low-arousing, NL = neutral low-arousing, UL = unpleasant low-arousing, UH = unpleasant high-arousing). b. Mean CRTs and error rates for PH, PL, NL, UL, and UH pictures. Error bars indicate standard errors.

Design and Procedure

After attachment of the electrode net, participants entered an electrically shielded recording chamber, where they were instructed and familiarized with the tasks. From each participant, ERPs were collected during the SRT and CRT task, with task order balanced across participants. During the CRT task, each of the 350 stimuli (175 pictures, 175 non-pictures) was presented once in randomized order. During the SRT task, we displayed all 175 pictures and 35 non-pictures in random order. Here, non-pictures were drawn randomly from the complete non-picture set (N = 175). Participants performed 10 practice trials prior to each task using non-experimental stimuli. Breaks were included in both tasks (two in the CRT task, one in the SRT task). The same stimulus duration and ITI were used as in Experiment 1. Participants delivered CRTs with their right index and middle finger as described above. In contrast to the previous studies, SRTs were performed with the right hand only, using the index finger and ALT button on a computer keyboard. Thereby, we intended to avoid ERP differences between the SRT and CRT task due to motor activity.

Analysis of Behavioral Data

Calculation of response speed and accuracy paralleled the procedure described in Experiment 3. Behavioral data were submitted to repeated measurement ANOVA using emotional content (5; PH, PL, NL, UL, UH) as within-subject factor. In addition, planned contrasts examined facilitation effects for PH and PL pictures and impairment effects for UL and UH pictures, compared to the neutral category.

EEG Recording and Preprocessing

The electroencephalogram was recorded continuously from 129 sensors using an Electrical Geodesics^{TM} (EGI) system. As recommended for the EGI high input-impedance amplifier, impedances were kept below 50 k. Data were digitized at a rate of 500 Hz and online band-pass filtered between 0.01 and 200 Hz. The vertex (electrode Cz) was used as recording reference. Channels with bad signal throughout the whole recording session (mean: 13 %) were interpolated offline using a spline interpolation method (Perrin, Pernier, Bertrand, & Echallier, 1989). An eye artifact correction was performed on the raw data applying the algorithm implemented in Besa^{TM} (Berg & Scherg, 1994b). For both tasks, stimulus-locked epochs were then extracted from 500 ms before until 1000 ms after picture/non-picture onset. Single trial epochs were inspected on artifact contamination using the semi-automatic artifact module of Besa. This tool allows for defining individual rejection criteria based on the distribution of epochs in terms of mean amplitude and gradient (first temporal derivate) values. For artifact rejection and all subsequent analyses, data were arithmetically re-referenced to average reference. Epochs being afflicted with artifacts were rejected. Artifact-free epochs were then averaged separately for each condition, considering only those trials with correct responses. ERPs were filtered between 1 and 30 Hz and baseline-corrected by subtracting the mean of a 200-ms segment prior to stimulus onset.

ERP Analysis: Amplitude Measures

As a first step, averaged data were visually inspected on whether the selected ERP components (see results section) were present in all subjects. We then determined sensor clusters covering the spatial foci of the components in the grand mean topographical maps (see Figure 7). For each participant, averaged waveforms were derived from the respective electrode cluster, which entered statistical analysis. Given that a high-density electrode net is employed, using such regional means instead of single sensor information enhances the reliability of the estimate used as dependent variable (Keil, Stolarova, Heim, Gruber, & Müller, 2003). Figure 5 shows the sensor clusters as used here, which are overlaid on an illustration of the employed electrode array.

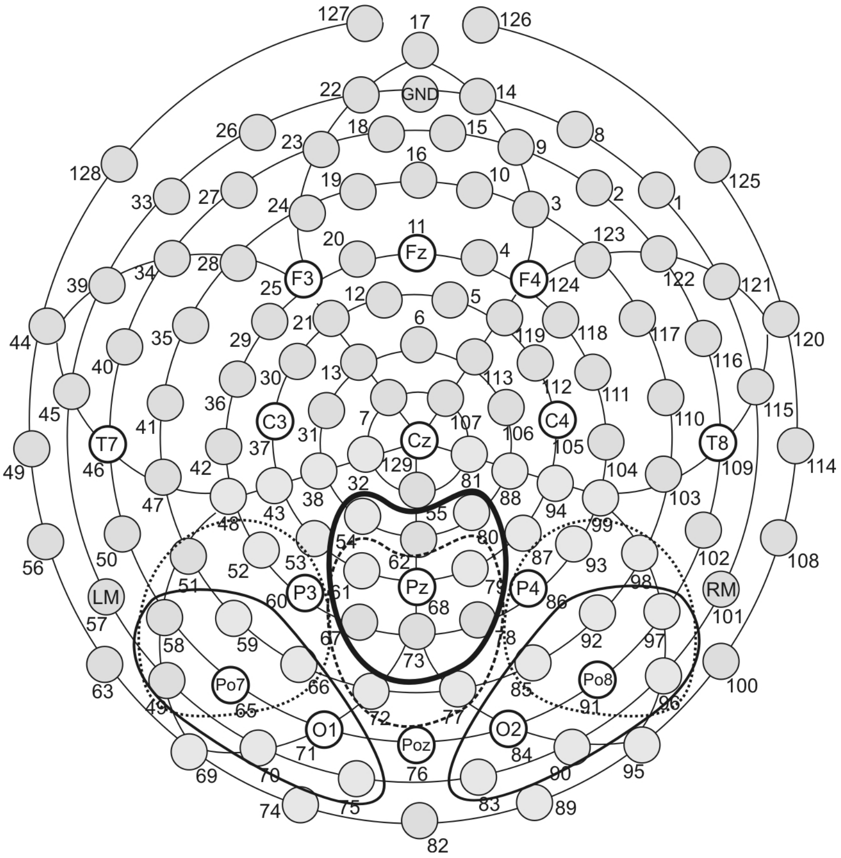

Figure 5. Layout of the electrode array employed in the event-related potential study (Experiment 4). Electrodes corresponding to sites of the international 10-20 system are labeled. The marked sensor clusters were selected to form regional means for each component: Dotted line = left and right sensor groups used to quantify the P1 component, thin solid line = left and right clusters used to measure the P2 component, dashed line = electrode group used to quantify the P3 component, thick solid line = sensor group used to assess LPP latencies. Note that the posteriors clusters did not include electrodes at the net margins, which are more prone to neck muscle artifacts or low signals than non-peripheral sensors.

As a second step, sensor cluster waveforms of single subjects were scored for the selected components. To this end, we determined the point of maximal deflection inside pre-specified time windows (P1: 90 -130 ms, P2: 180 - 300 ms, P3: 280 - 360 ms). The search windows were carefully examined such that they were not sensitive to the rise or fall of deflections at the window margins. Mean amplitudes were then calculated using a segment from +/- 30 ms (+/- 20 ms for P1) around the component's peak. Amplitude values of picture and non-picture responses were evaluated by ANOVA using emotional content (5; PH, PL, NL, UL, UH) and, in case of the bilateral P1 and P2, hemisphere (2; left, right), as within-subject factors. Paralleling behavioral analyses, differences between emotional and neutral picture categories were also tested by planned comparisons.

ERP Analysis: LPP Latency

Using the parietal sensor group shown in Figure 5, LPP latencies for different stimulus categories were computed and statistically analyzed with a jackknife-based approach proposed by Miller, Patterson, and Ulrich (1998). This method was originally developed to compare onset latencies of the lateralized readiness-potential but is also applicable to other ERP parameters, such as N400 peak latencies (Hohlfeld, Mierke, & Sommer, 2004). In essence, by utilizing grand average data jackknifing yields a more accurate estimate of true differences between conditions than measures derived from noisy single-subject latencies. Jackknifing involves two major steps. First, the latency for a particular condition is computed as the time point when the grand average waveform exceeds a specific criterion value. Similar to a traditional t-test, the grand average difference between two conditions is then tested against an estimated standard error of the difference derived from subsample grand averages. These so-called jackknife samples are obtained from calculating as many grand averages as there are participants omitting a different participant for each grand average. In the present experiment we selected a criterion of 2 microcvolt [µV] which would be exceeded in the late rise of the LPP, taking into account that the component displayed a plateau-like shape and a single peak could not be unambiguously identified (see Figure 9). Further, as recommended by Miller (Miller et al., 1998), an absolute instead of a relative (e.g., half of the peak amplitude) criterion is preferable when peak amplitudes may differ between conditions.

Results

Behavioral Responses

Paralleling the previous results, emotional content modulated CRTs to pictures, $F(4, 68) = 22.1$, $p < 0.0001$, but neither CRT responses to non-pictures, $F(4, 68) = 1.8$, nor SRT responses, $F(4, 56) < 1.0$. ANOVA on CRT error rates showed a significant main effect of emotional content for plain pictures, $F(4, 68) = 8.6$, $p < 0.0001$, but not for distorted items, F(4, 68) = 1.3. Most interestingly, paired comparisons with the neutral category revealed that CRTs were faster for PL pictures, F(1, 17) = 8.8, $p < 0.01$, but not for PH pictures, F(1, 17) = 1.4 (see Figure 4, p. 39). Vice versa, only UH pictures delayed RTs, F(1, 17) = 23.6, $p < 0.001$, but not UL pictures, $F < 1.0$. Error rates contrasts showed higher response accuracy for PH pictures, F(1, 17) = 8.7, $p < 0.01$, and UL pictures, F(1, 17) = 7.0, $p < 0.05$, but no significant differences for PL, F(1, 17) < 1, and UH pictures, F(1, 17) = 3.7, $p = 0.07$.

Electrocortical Responses: P1, P2, and P3

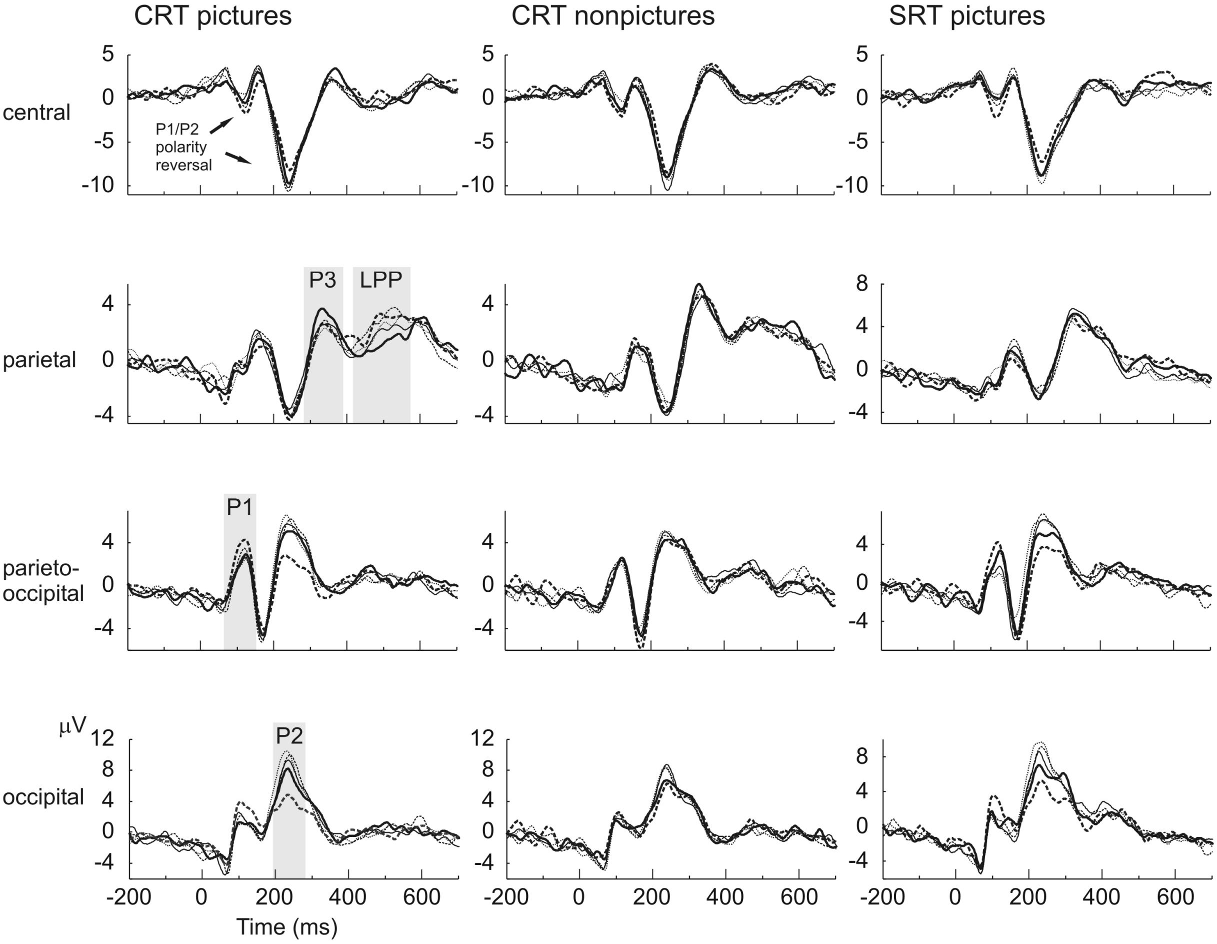

Figure 6. Grand average ERPs at representative centromedial (electrode # 129 [CZ]), parietomedial (# 62), right parietooccipital (# 97), and right occipital (# 92) sensors. Time zero represents stimulus onset. Left and middle column: ERPs to pictures and non-pictures in the CRT task, right column: ERPs to pictures in the SRT task. Dashed thin line = pleasant low-arousing content, dashed thick line = pleasant high-arousing content, dotted line = low-arousing neutral content, solid thin line = low-arousing unpleasant content, solid thick line = high-arousing unpleasant content.

Figure 6 illustrates grand average ERPs elicited by different picture categories during the SRT (right column) and CRT (left and middle columns) tasks at a subset of representative sensors. Two major early components over posterior areas are evident having distinct spatial topographies (see Figure 7): i) A P1 distributed over left and right lateral-occipital sensors and peaking at 110 ms, and ii) a P2 showing a more inferior, bilateral occipitotemporal topography and maximal at 220 ms.

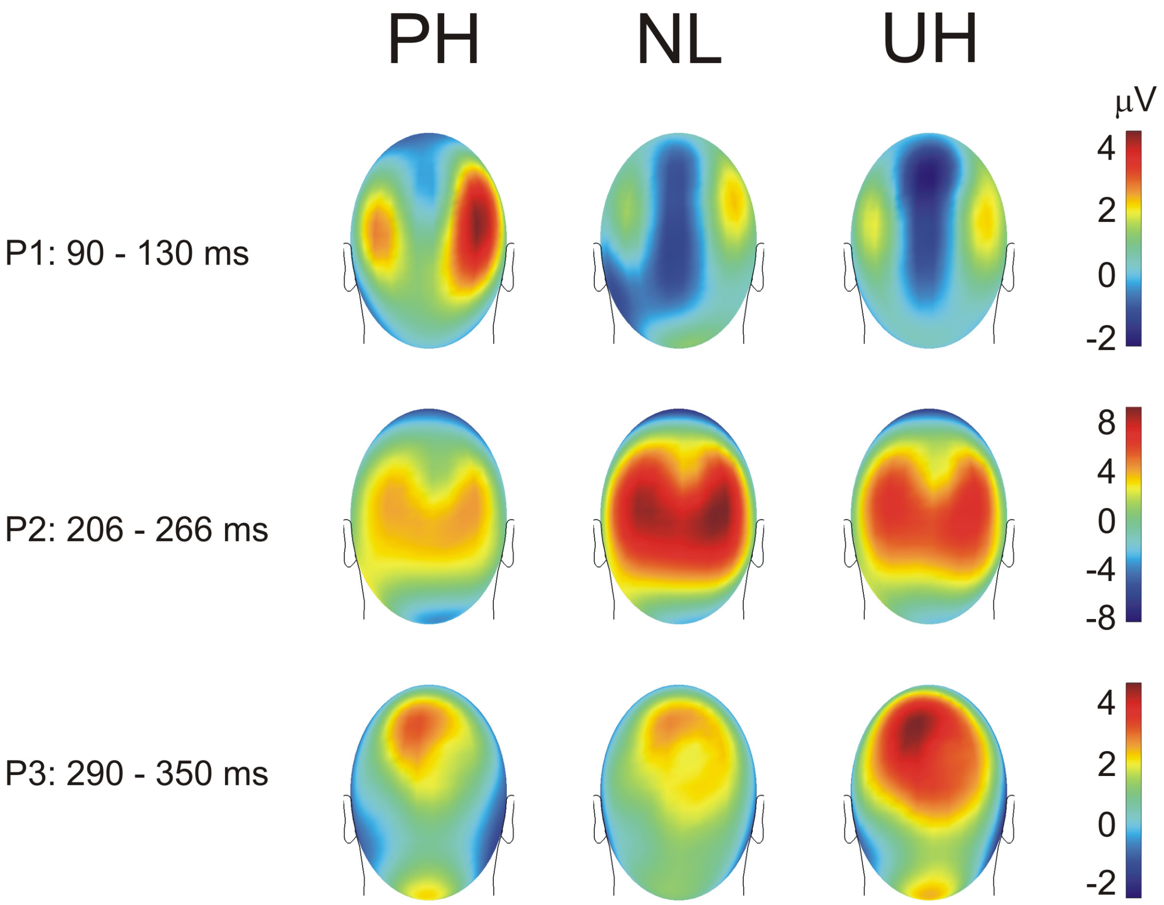

Figure 7. Spatial topographies of the scalp potentials represented as the mean activity across the P1, P2, and P3 time windows. Topographical maps were created by spherical spline-interpolation of sensor voltages (Perrin et al., 1989). Scalp topographies illustrate grand mean amplitude differences between high-arousing pleasant (PH), low-arousing neutral (NL), and high-arousing unpleasant (UH) pictures.

Both for the CRT and SRT task, ANOVA on picture P2 amplitudes yielded main effects of emotional content, CRTs: $F(4, 68) = 22.0$, SRTs: $F(4, 68) = 12.0$, $ps < 0.0001$. Hemisphere did not modulate P2 responses, $Fs < 1.6$. As can be seen in Figure 8b, P2 amplitudes in the CRT task showed strong negative shifts for PH, $F(1, 17) = 73.7$, $p < 0.0001$, and UH pictures, F(1, 17) = 10.7, p < 0.01, compared to NL pictures. Modulation was smaller for UL pictures, F(1, 17) = 4.5, p < 0.05, and absent for PL pictures, F(1, 17) = 2.1, p = 0.17. A similar pattern was found for P2 amplitudes in the SRT task, PH: F(1, 17) = 40.9, p < 0.0001, UH: F(1, 17) = 6.5, p < 0.05, UL: F(1, 17) = 4.9, p < 0.05, PL: F < 1.0. Thus, these results closely replicate previous findings showing negative amplitude shifts around 200 - 250 ms for emotionally arousing stimuli (Schupp et al., 2003). A weak modulation of P2 amplitudes by emotional content was also evident for non-pictures CRT responses, F(4, 68) = 2.9, p < 0.05, reflecting more negative values for PH versus NL items, F(1, 16) = 6.6, p < 0.05. There were no effects of hemisphere, Fs < 1.0.

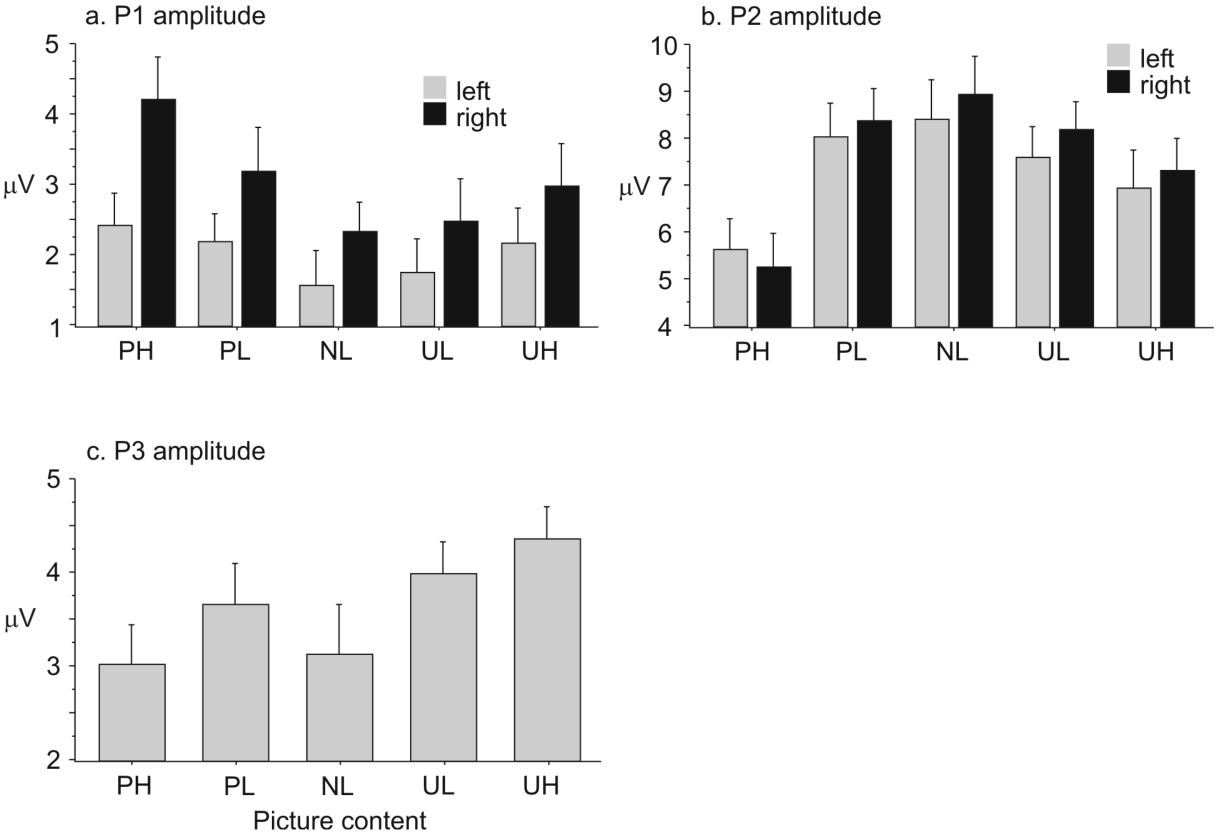

Figure 8. a. Mean amplitudes of the P1 component as a function of hemisphere and picture emotional content (PH = pleasant high-arousing, PL = pleasant low-arousing, NL = neutral low-arousing, UL = unpleasant low-arousing, UH = unpleasant high-arousing). b. Mean amplitudes of the P2 component grouped for hemisphere and picture content. c. Mean amplitudes of the P3 for emotional pictures categories.

As opposed to the P2 modulation, analyses of the earlier P1 indicated amplitudes differences contingent on stimulus valence. ANOVA on picture P1s in the CRT task showed larger amplitudes in the right compared to the left hemisphere, F(1, 17) = 4.0, p = 0.06.

Further, we obtained a main effect of emotional content, F(4, 68) = 3.6, $p < 0.01$. Planned contrasts with the neutral category revealed that this effect was due to enhanced amplitudes for both pleasant categories, PL: F(1, 17) = 4.9, $p < 0.05$, PH: F(1, 17) = 9.2, $p < 0.01$ (see Figure 8a). In contrast, P1 responses to aversive pictures (UL and UH) were not different to neutral picture responses, Fs < 2.4. ERPs to distorted pictures were not subject to emotional modulation, Fs < 2.1. Similar to the CRT task, P1s in the SRT task were larger in the right compared to the left hemisphere, F(1, 17) = 9.5, $p < 0.01$. We also found a main effect of emotional content, F(4, 68) = 4.4, $p < 0.01$, reflecting enhanced responses for PH compared to NL pictures, F(1, 17) = 15.1, $p < 0.01$.

In addition to the early posterior components (P1 and P2), CRT but not SRT ERPs yielded a biphasic pattern of positive-going potentials starting around 300 ms over midline parietal sensors (see Figure 6). A first positive peak (labelled P3) of this biphasic complex was evident around 320 ms, i.e., intermediate between the EPN and LPP windows typically examined in studies using emotional pictures. ANOVA showed that during choice reactions P3 amplitudes varied with emotional content, F(4, 68) = 5.0, $p < 0.01$. Paired contrasts revealed a significant P3 enhancement for UH pictures solely, F(1, 17) = 6.9, $p < 0.05$ (see Figure 8c). During non-picture trials, P3 amplitudes were not sensitive to emotional category, Fs < 2.1.

LPP Latency

The P3 component was followed by a second positive deflection (termed late positive potential LPP) having a similar topography but a longer duration (450 - 650 ms) as the earlier positive wave. As can be seen in Figure 9, LPP latencies varied strongly with emotional content. Using the electrode cluster and criterion value specified in the methods section, grand mean latency was 486 ms for neutral pictures. Paralleling RTs, latencies were shorter for pleasant pictures (PL: 462 ms, PH: 470 ms) and longer for arousing unpleasant pictures (UH, 574 ms). Only a small difference was evident for UL pictures (482 ms). As detailed above, each latency difference between the emotional categories (PH, PL, UL, UH) and the neutral category was statistically tested against an estimated standard error of the difference obtained from jackknife samples. Significant t-values were obtained for the difference between PL and NL, t(17) = 2.44, $p = 0.013$ (single-sided), and the difference between UH and NL pictures, t(17) = 7.13, $p < 0.0001$. The difference between PH and NL pictures did not reach significance, t(17) = 1.5. Analogous jackknife analyses on LPP latencies of non-picture responses showed no significant effects of emotional content, all ts < 1.1.

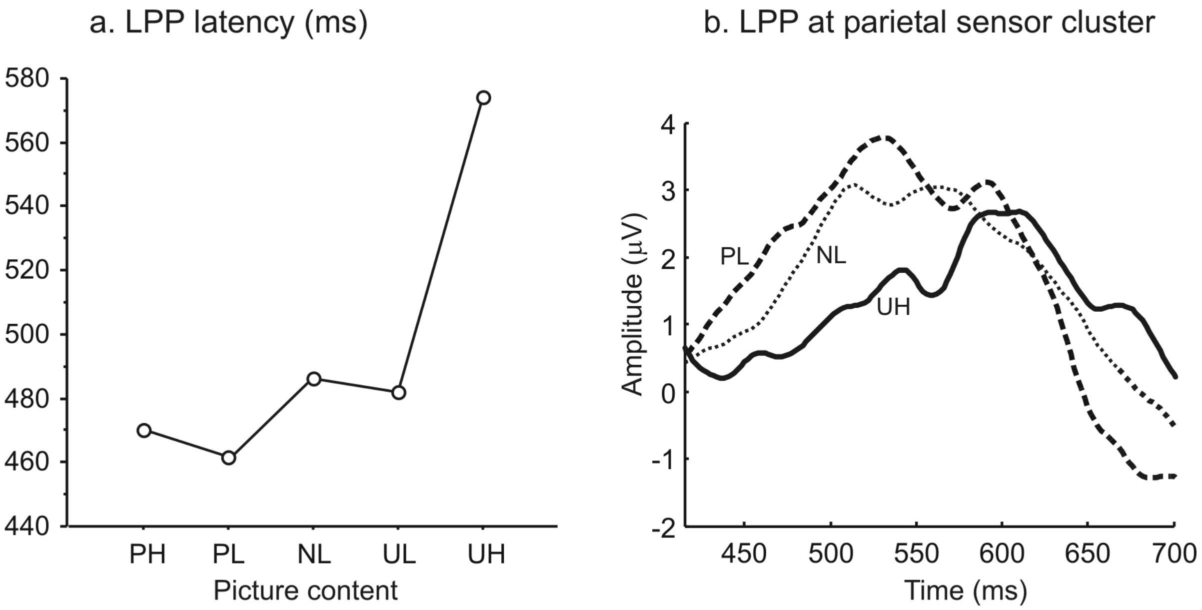

Figure 9. a. Grand average LPP latencies for high-arousing pleasant (PH), low-arousing pleasant (PL), low-arousing neutral (NL), low-arousing unpleasant (UL) and high-arousing unpleasant (UH) pictures. Latencies were quantified using a parietal sensor cluster (see Figure 5) and a criterion value of 2 μV. b. Grand average LPP waveforms for PL, NL, and UH pictures illustrating latency differences. Waveforms show sensor group means.

Discussion

Behavioral data in the present study extend the previous results in two important ways: Pleasant pictures speed visual categorization specifically for content that is judged as non- or mildly arousing, for instance pictures showing smiling persons or a mother that bathes her baby. In contrast, aversive content hampers CRTs only when the stimulus is associated with high degrees of emotional arousal, as in the case for pictures showing severe injuries or highly violent attack scenes. This pattern of behavioral results was mirrored by LPP latencies backing the assumption that CRT modulation is due to differences in cognitive processing time (stimulus evaluation and response selection) and not related to emotional effects on motor-preparation or -execution. Further, amplitude and latency parameters of non-pictures ERPs did not consistently vary with emotional content, suggesting that confounding physical category differences are marginal. As delineated below, electrocortical measures in this study also provide clues as to possible neuronal origins of behavioral effects.

P1 and P2 – Enhancing the Good But Separating the Important

ERP modulation specific to pleasant content was found for the lateral occipital P1 which is generated between 90 and 120 ms in extrastriate visual cortices. Amplitudes were amplified for high-arousing pleasant and low-arousing pleasant pictures. In spatial selective attention research, enhanced P1 responses have been reported for stimuli at attended versus non-attended stimulus locations. This effect is thought to reflect the operation of a control mechanism that acts to amplify sensory processing for attended stimuli (Hillyard, Vogel, & Luck, 1998). Similar "sensory gain" affecting P1 and N1 components of the ERP is triggered by non-instructional or reflexive forms of attention (e.g., orienting to abrupt changes of luminance, Hopfinger & Mangun, 1998). Thus, accelerated responses to pleasant pictures may reflect sensitivity to simple features associated with appetitive content (e.g., a smiling mouth) which enables fast processing of stimuli containing these features. Substantial effects of picture content were also evident for the second positive deflection (P2). Replicating previous studies (Schupp et al., 2003), in both tasks the P2 was more negative for high-arousing pictures (PH and UH), compared to the other categories. Further, emotional content affected non-pictures P2 responses to some extent, indicating that processing in this time range is highly sensitive to affective parameters, even to rudimentary features in scrambled images. In the domain of feature-based attention, increased negativity around 200 ms is linked to the selection of stimuli for further analysis (Näätänen, 1992). Posterior negativity have also been related to early semantic selection processes during natural scene categorization (Codispoti et al., 2006), showing amplitude shifts for target (e.g., items containing an animal) versus distractor (items containing no animal) pictures. Analogously, early posterior negativity (termed EPN) to arousing pictures is linked to a preparatory selection process that tags stimuli according to their motivational significance, i.e., irrespective of valence (Schupp, Stockburger, Codispoti et al., 2007). The relationship between early ERP modulations and later LPP/RT effects indicates that this selection mechanism may have inhibitory effects on the speed of behavioral responses. Though P1 responses were amplified for both PL and PH pictures, only the low-arousing pleasant category was related to a significant decrease of LPP/RT latencies. For PH pictures, the selection mechanisms thus counteracted their initially attained sensory gain, which in turn caused a lack of significant RT and LPP modulation. In the absence of P2 modulation (as it was the case for PL pictures), early perceptual facilitation was propagated unimpededly to discriminative processing and response-related stages.

Recent work has shown that occipital ERPs between 150 - 250 ms are also affected by picture composition showing more negativity for simple figure-ground pictures than for complex scenes (Bradley et al., 2007). Thus, posterior negativity was suggested to reflect at least partly a perceptual pop-out effect related to salient object features. In this view, simple figure-ground compositions are more easily mapped on mental stimulus representations than complex picture scenes. Such a mechanism may also account for findings showing that the EPN is most pronounced for erotica (e.g., De Cesarei & Codispoti, 2006; Schupp, Stockburger, Codispoti et al., 2007). Relative to other affective pictures, the perceptual structure of erotica pictures may be simple showing more conspicuous figure-ground segregation. In line with this, in the present study PH pictures (mainly comprising erotica) evoked a disproportional larger posterior negativity than UH pictures (see Figure 6) though they were related to similar arousal. Crucially, we found faster CRTs for PL pictures specifically, which included heterogeneous content but no salient erotica. Further, PL pictures did *not* elicit early negative amplitude shifts indicating that the figure-ground dimension is not sufficient to account for faster categorization responses in the present paradigm. In sum, more negative P2s to PH and UH pictures may reflect the joint effects of arousal-related selection and perceptual picture composition.

P3 and LPP – Attending to Threat and Slowing Cognition

Between 300 and 400 ms, i.e., intermediate between the EPN and LPP windows typically examined in affective processing studies, the present data reveals a specific enhancement for highly arousing aversive content. This enhancement was evident as an amplified positive deflection, which peaked around 320 ms over parietal sensors. Comparing the latency of the P3 to the average choice reaction time (558 ms) suggests that it did not represent a late positive component related to successful target categorization or the completion of stimulus evaluation (McCarthy & Donchin, 1981). Rather, it may relate to a reflexive attentional response which was prompted by high-arousing aversive stimuli specifically and, at the behavioral level, resulted in delayed RTs. Thus, upon encountering threatening stimuli the brain may disrupt cognitive analyses related to motivationally irrelevant, i.e., non-evaluative stimulus attributes (including the discrimination along the picture versus non-picture dimension). From an evolutionary perspective, this may have protective functions and serve to summon available resources for the preparation of appropriate actions.

During passive viewing, increased attention to motivationally significant stimuli is typically related to amplitude enhancement of late positive potentials (LPPs), which are maximal between 400 - 800 ms (Schupp, Flaisch et al., 2006). With long stimulus duration, the LPP merges in a positive slow wave, which can sustain up to several seconds (e.g. Cuthbert et al., 2000; Schupp, Cuthbert et al., 2004). In these studies, amplitude enhancement of the LPP was assumed to reflect continued analysis and heightened processing of arousing events. Speeded task conditions and comparatively short inter-trial intervals as used here may prevent an elaborate processing of arousing stimuli. In line with this notion, modulation of the LPP did not reveal as an amplitude enhancement. Rather, we found pronounced effects of emotional content on LPP latency, which mirrored the behavioral results. This finding is in accordance with the assumption that in reaction time tasks latencies of late positive potentials, such as the P300 (also referred to as P3b), index completion of stimulus evaluation (McCarthy & Donchin, 1981). Later work extended this view and emphasized the contribution of response selection processes to the timing of these components (Pritchard, Houlihan, & Robinson, 1999; Verleger, 1997). During speeded processing, latencies of late positive potentials thus provide a useful tool to estimate the time where a transition occurs between information processing (stimulus discrimination and/or response selection) and genuine motor processes (motor preparation and execution). In the current context, LPP latency differences between picture categories make a strong case for a cognitive (versus motor) origin of emotional response modulation.

GENERAL DISCUSSION OF SERIES A

Processing of emotionally arousing visual scenes activates evolutionary old motivational systems (appetitive and defensive) that are implemented as distinct neuronal networks in the brain (Lang & Davis, 2006). In situations requiring comparatively late behavioral access to mental computations, these systems jointly act to enhance the accuracy of responses to an emotionally arousing stimulus. Such accuracy benefits for arousing (pleasant and unpleasant) stimuli are found for instance in non-speeded attentional blink, memory or recognition tasks (Anderson, 2005; Bradley et al., 1992; Trippe et al., 2007; Zeelenberg et al., 2006). The present results show that the behavioral outcome may be fundamentally different as soon as the situation poses temporal constraints on information processing.

Three Affective Mechanisms during Speeded Information Processing

When a centrally presented visual scene must be categorized under speeded task instructions, behavioral decisions reflect neuronal calculations completed on average within

500 - 600 ms or earlier. During this period, emotional pictures trigger a rapid cascade of cortical processes having distinct temporal and topographical activation patterns. We show that at specific time ranges within the first 400 ms these processes display high sensitivity towards arousal and valence parameters of the stimulus. Initially (90 - 130 ms), a conspicuous amplification of electrocortical responses is found for pleasant content irrespective of arousal, presumably reflecting sensitivity to appetitive object features (e.g., a smiling mouth). Given that the stimuli do *not* engage motivational circuits (i.e., elicit no or low degrees of arousal), this initial perceptual benefit is available to discriminative analysis/response selection and thus propagated to the behavioral level. Conversely, initial perceptual benefits for high-arousing pleasant stimuli (e.g., erotica) are leveled off by a second mechanism operating around 200 - 250 ms. This mechanism may serve to label information related to high evolutionary significance (Schupp, Stockburger, Codispoti et al., 2007), such as sexuality or physical damage, and inhibit cognitive processes not relevant for reproductive or defensive behavior. Even more rigorously, irrelevant analyses are disrupted in a third phase (around 300 - 400 ms). Here, increased attention is commanded by to those arousing stimuli that signal threat for the organism. Attending to threat may be accompanied by a pervasive suppression of cognitive processes. Under speeded task conditions, the interplay among these early processes cumulates to diverging behavioral outcomes for different emotional categories. Across visibility and discrimination conditions we found facilitated responses for pleasant visual scenes. Experiment 4 indicated that these behavioral benefits are due to faster RTs to pleasant content that judged as mildly or non-arousing. This finding is consistent with the happy-face-advantage during evaluative categorization (Leppänen et al., 2003). Conversely, aversive pictures were related to impaired performance, particularly when they were judged as highly arousing.

The outlined interference effects in presence of arousing stimuli are presumably mediated by re-entrant input from distinct neuronal circuits in cortical and subcortical regions (e.g., the amygdala) that are rapidly activated in response to appetitive and aversive pictures and exert a regulatory influence on information processing. On a cellular level, connections of the appetitive/defensive system to categorization/response selection modules may be inhibitory in nature. In contrast, facilitated processing of non-arousing pleasant content may not originate from a regulation exerted by motivational circuits but reflect either i) perceptual salience of appetitive features (e.g., figure-ground segregation, see Bradley et al., 2007) or ii) a (learned or hard-wired) sensitivity of the brain to detect those features.

Emotional content did not modulate speed parameters when it was conveyed by a verbal information channel. Although many studies have employed words as instances of emotional events in general, it is doubtful whether linguistic material may elicit a comparable amount of arousal as photographs depicting real-life situations. Further, we have argued that the discrepancy between verbal and pictorial affective processing may result from differences in temporal activation patterns (Keil, 2006). Our electrocortical data support this view showing reliable effects of emotional category on picture ERPs within half of a second. Such content-based modulation is feasible because natural scenes stimuli trigger fast activation of object representations (Thorpe et al., 1996). In addition, as argued above early modulations may relate to the detection of few emotional features. In contrast, robust effects of emotional content on word ERPs typically start around 300 ms (Fischler & Bradley, 2006), probably reflecting the need to complete lexical analysis before emotional parameters can be accessed. This may explain why affective words reliably modulate performance in non-speeded accuracy tasks (Anderson, 2005), whereas speeded tasks fail to show a reliable influence of emotional content. Even for the well-studied emotional Stroop task, effects of are often weak and restricted to particular populations (Harris & Pashler, 2004) or word classes (MacKay & Ahmetzanov, 2005).

Threat Delay – A Functional Perspective

It is likely that the outlined processing modes – sensitivity to pleasant features and threat delay – have evolved as adaptive responses in a rapidly changing and complex sensory environment. The organism was challenged with the necessity to act successfully to a multitude of new events that could be either hostile or hospitable. In response to possible threats, slowing of certain cognitive analyses would serve to avoid erroneous decisions/categorizations that could be highly deleterious, for instance when the encountered object turns out as a predator. These notions corroborate the view that the default response to aversive material reflects a form of “adaptive conservativism” (Mineka, 1992). They also are pertinent with respect to the defense cascade model that specifies a characteristic physiological response pattern towards aversive stimuli (Lang et al., 1997). A core prediction of the defense cascade is an initial phase of motor inhibition. Motor freezing is accompanied by a variety of physiological reflexes, such as cardiac deceleration and electrodermal increase, but is also hold to reflect a phase of enhanced sensory intake. Our findings suggest that within the first several hundred milliseconds, freezing is even more pervasive disrupting also comparatively simple cognitive processes (e.g., visual discrimination and response selection).

Such cognitive freezing may reflect a rigid and secluded form of sensory intake, which serves to gather stimulus and context details prior to the initiation of action. By enhancing the accuracy of responses after approximately 500 - 1000 ms, the initial slowing of information processing thus turns out adaptively in the long run.

Recently, Vuilleumier (2005) and Phelps (2006) proposed a distinction between emotion effects elicited by either attention or perceptual processes. It is tempting to accommodate the present findings in this framework, i.e., relating accelerated responses to perceptual facilitation and delayed responses to attention capture effects. However, such a distinction may be arbitrary in case of early affective processing. Given that emotional stimuli elicit reflexive or implicit forms of attention (termed motivated attention, Lang et al., 1997), perceptual and attentional effects are hard to disentangle, particularly for early effects as observed here in the P1 time range. Indeed, it has been proposed that some attentional effects result from perceptual modulation (Carrasco, Ling, & Read, 2004). Further, we have shown that threat delay was also present for masked stimuli. Thus, negative stimuli do not only elicit comparatively slow attention effects, such as guiding spatial attention during visual search and capturing selective attention in the Stroop task, but might trigger a fast and comprehensive slowing of cognitive analyses (see also Algom et al., 2004).

The present findings underscore that it is essential to consider the time course of emotion-perception and emotion-cognition interactions, leading to different behavioral results for tasks tapping different time ranges of processing. A suitable approach to reveal such interactions are thus ERP measures having high temporal resolution. In the present study, ERP analysis indicated that amplitude modulations at different stages of affective processing can be exploited to predict information processing time. Future work may employ neuronal modeling approaches to examine such relationships more specifically. Additional data are also required to elucidate possible effects of emotional parameters on motor preparation and execution processes. The present variation of LPP latencies make a good case that affective modulation in the present task resides at cognitive rather than motor stages of processing. This assumption was also supported by supplementary analyses of the present EEG data using response-locked ERPs. However, caution is warranted in drawing conclusions from these results, as the time constant of the employed EEG system does not track slow oscillations, such as negative motor preparation potentials, reliably. Future work may apply appropriate devices to measure, for instance, lateralized- readiness potentials (LRPs), which afford an opportunity to examine hand-specific motor preparation and execution in motor cortices.

2.2 Emotional Processing and Subsequent Lexical Analysis (Series B)

The second empirical part[5] of this thesis addresses the question how emotional stimuli influence encoding of subsequent information. Specifically, three studies sought to explore electrophysiological and behavioral consequences for neutral word stimuli presented in temporal proximity to an emotionally arousing picture. As described below, the fate of stimuli following an emotional event can be predicted in two different ways.

Resource Capture by Emotional Events and Affective Interference

Findings of the emotional Stroop effect (see Introduction) suggest that emotional events capture attentional resources and interfere with competing processes. Pertinent to the present studies, these detrimental capture effects have also been shown for neutral information presented with a temporal gap to the emotional distractor. Thus, at short SOAs (300 ms) lexical decisions to neutral probe words are slower when they are preceded by task-irrelevant threat words (Calvo & Castillo, 2005). Similarly, within a time window of 200 ms, negative RSVP distractors impair the detection of subsequent target pictures more than neutral distractors (Most et al., 2005).

These findings indicate that behavioral facilitation effects, as shown for pleasant pictures in the previous studies, may be restricted to the emotional stimulus itself and entail impaired processing of subsequent information. Electrocortical studies of affective interference, however, are rare. One line of research analyzed the P300 to acoustic startle probes as an indicator of resource deployment by concurrently presented emotional background stimuli (e.g., Cuthbert, Schupp, Bradley, McManis, & Lang, 1998; Schupp, Cuthbert et al., 2004). These authors reported a reduction of probe P300 amplitudes when viewing arousing (pleasant and unpleasant) compared to non-arousing scenes. Yet, probe stimuli were typically presented as late as 2.5 - 4.5 s following picture onset (Schupp, Cuthbert et al., 2004), so that these studies did not shed light on the short-term dynamics of emotional interference.

---

[5] Parts of the present chapter are included in the article "The Costs of Emotional Attention: Affective Processing Inhibits Subsequent Lexico-Semantic Analysis" (in press, Journal of Cognitive Neuroscience) by Niklas Ihssen, Sabine Heim, and Andreas Keil.

An Emotion-Induced Mode of Facilitated Processing?

In addition to the described interference effects, a different prediction might be inferred from the work by Phelps et al. (2006). Using a SOA of 125 ms, they found that emotional faces can facilitate subsequent visual processing of grating patterns (see Introduction). Such transient but non-selective facilitation elicited by emotional stimuli might also be predicted from the studies using spatial attention paradigms. For instance, Pourtois, and colleagues (Pourtois, Grandjean, Sander, & Vuilleumier, 2004) presented pairs of faces in the left and right visual hemifield. One face showed a fearful and the other a neutral expression. With SOAs varying between 200 and 400 ms, the face pairs were followed by unilateral bars with either a horizontal or vertical orientation. Increased P1 amplitudes and faster discrimination responses to bars replacing fearful versus neutral faces indicated that fear cues facilitated processing of subsequent stimuli at their spatial location. In the context of event-related functional magnetic resonance imaging such facilitation effects were demonstrated for target stimuli following cues with acquired emotional relevance (Armony & Dolan, 2002). It is conceivable that the detection of any motivationally relevant event induces a transient attentional response, which temporarily tunes the brain for efficient encoding of subsequent sensory input. This implies that a quick glimpse on emotionally arousing scenes might shift the perceptual system to a facilitated processing mode, being unspecific with regard to all incoming information.

Overview of the Studies of Series B

The studies reviewed above have provided evidence for the capability of emotional stimuli to facilitate processing of temporally adjacent information, but also to impair subsequent processing. Thus the question arises, which variables mediate the divergent outcomes of facilitation versus interference as a consequence of affective processing. Importantly, differences related to emotional intensity/arousal and affective stimulus types (e.g., faces versus pictures) may contribute to the pattern of findings. Further, it is conceivable that the temporal distance or complexity of the primary task play a key role in eliciting behavioral costs versus benefits. Here, we used a lexical decision as the primary task, which was preceded by task-irrelevant emotional pictures. This setup aimed to be sensitive to the influence of arousing material on higher-order processing as is the case with lexico-semantic analysis (Binder et al., 2003). Linguistic as opposed to picture targets reduce variance related to physical properties. Compared to simple visual stimuli (e.g., geometric shapes), they provide an opportunity for using a high number of stimulus exemplars, thus tapping processes related to identification of new information more directly. Affective pictures were used as precedent stimuli as they allowed us to maximize differences in emotional intensity of the irrelevant stimuli, as compared to word or face stimuli.

In Experiment 1, we employed a temporal interval of 200 ms between picture and target, which has been recently shown to provoke reliable emotion effects on subsequent target processing (Most et al., 2005). Behavioral and electrocortical dynamics were recorded while participants responded to word/pseudoword targets preceded by high-arousing pleasant pictures (erotica or sports), high-arousing unpleasant pictures (attack or mutilation), or low-arousing neutral pictures (neutral persons). Thus, Experiment 1 was intended (i) to explore the occurrence of behavioral facilitation/interference elicited by briefly presented arousing pictures within the context of lexico-semantic processing and (ii) to track electrocortical correlates of such effects. Electrophysiological measures at high temporal resolution are particularly advantageous, because interference effects may differentially affect subsequent processing, depending on the time range and brain processes involved.

Using the same stimuli as in the first study, Experiment 2 tested the possibility that the direction of behavioral emotion effects (facilitation versus interference) depended on the temporal interval between picture and target. As suggested by the findings of Phelps et al. (2006), transient facilitation may emerge within a narrow time-window after offset of the emotional stimulus while later processes are suppressed. To test this hypothesis, picture-target pairs were presented at varying SOAs (80 ms, 200 ms, 440 ms).

Experiment 3, examined the role of emotional arousal versus hedonic valence in the present paradigm. In particular, we sought to rule out that response modulation was driven by salient differences related to semantic picture categories (e.g., erotica versus sports scenes) and not by variations in emotional arousal. To this end, responses to low-arousing pictures were contrasted with responses to high-arousing items having similar valence scores. In addition, post-experimental affective picture ratings were invited to predict RT modulation both on an individual subject and a single item level.

Methods

EXPERIMENT 1

Participants

Twenty-two healthy university students (11 females) with a mean age of 23.0 years (Standard Deviation, SD = 5.9) gave their informed consent to participate. As in the subsequent experiments, all participants were native speakers of German and reported normal or corrected-to-normal vision. Twenty-one participants were right-handed having a laterality quotient of ≥ 70 in the Edinburgh Handedness Inventory (Oldfield, 1971). One subject was ambidextrous, but was retained in the sample, as she did not show any signs of outlying ERP activity. In all three experiments, participants received course credits or a financial bonus of 7.50 Euros.

Stimuli

Verbal stimuli comprised neutral German verbs (e.g., ‘beantworten’ = to reply, ‘vorschlagen’ = to propose) and pronounceable, arbitrary pseudowords (e.g., ‘beglinfen’, ‘vorstelmen’). Selection of verbs was based on a pre-experimental rating study which used a paper-and-pencil version of the 9-point-scale Self-Assessment Manikin (SAM, Bradley & Lang, 1994). We obtained ratings for 160 verbs on the dimensions of emotional valence and arousal from 80 adults (mean age = 24.9 years, SD = 5.7, 38 females). We selected 90 verbs with low arousal (mean = 3.8, SD = 0.6) and neutral valence (mean = 5.4, SD = 0.4) scores. Using this set, 90 pseudowords were generated by permuting syllables and letters within or between verb stimuli. Pseudowords and words were matched for item length, initial letters, prefixes, and endings.

Pictures were taken from a pre-experimental study using a computerized version of the SAM. A group of 45 adults (mean age = 23.3 years, SD = 4.2, 26 females) judged 216 pictures from the IAPS (Lang et al., 2005), the Graphics Suite CD (Micrografx, 1997), and various web sites. Each stimulus displayed at least one human agent. Based on these ratings, 180 pictures were selected: (i) 60 high-arousing pleasant items including 30 erotica (mean valence = 6.6, SD = 0.5; mean arousal = 6.0, SD = 0.5) and 30 sports and adventure scenes (mean valence = 6.5, SD = 0.6; mean arousal = 5.5, SD = 0.4), (ii) 60 low-arousing neutral pictures depicting, for instance, office activities (mean valence = 5.2, SD = 0.4; mean arousal = 2.7, SD = 0.4), and (iii) 60 high-arousing unpleasant pictures showing 30 attack scenes (mean valence = 2.5, SD = 0.4; mean arousal = 6.3, SD = 0.5) and 30 exemplars of mutilated bodies or injured people (mean valence = 1.5, SD = 0.4; mean arousal = 7.0, SD = 0.6). Within each affective category, pairs of stimuli were selected displaying equivalent content (e.g., two pictures showing windsurfing, two pictures showing people in roller coasters). This resulted in two semantically congruent sets A and B of stimuli, consisting each of 30 pleasant, 30 neutral, and 30 unpleasant pictures.

Procedure

Subjects viewed the stimuli on a 19” computer screen at an average distance of 70 cm. A chin rest helped to reduce movement artifacts and inter-individual variability in terms of head position and visual angle. All stimuli appeared centrally against a gray background. Participants performed 5 practice trials involving stimuli not used in the test trials. Each trial started with a 14.5 x 20 cm picture appearing for 100 ms, followed by a random dot mask of equal size and duration. Insertion of a mask aimed to help separating electrocortical activation patterns evoked by picture versus word presentation. Verbal targets were shown in white color immediately after mask offset, resulting in a picture-target SOA of 200 ms. Targets were presented in 36-point Times New Roman font using lower-case letters. Participants were asked to indicate whether the target stimulus was a real word in German by pressing the left mouse button with their right index finger or a nonsense word by pressing the right button with their right middle finger. Button assignments were counterbalanced across subjects. Participants were asked to give their response as fast and accurate as possible within a time window of 2000 ms. Target presentation was terminated by the response. No feedback was provided. After an interval randomly varying between 2500 and 3500 ms and displaying a central fixation cross, the next trial started. The experimental session was divided into two runs separated by a short break. During each run, each of the 90 words and 90 pseudowords was presented once in random order, resulting in 360 trials for the whole session. In the first run, word targets were randomly combined with pictures from set A and pseudowords with those from set B (see above). The assignment of words/pseudowords to the picture sets was reversed in the second run. The order of set assignment was counterbalanced across participants.

Analysis of Behavioral Data

For each participant and picture x target combination, the mean reaction time (RT) and error rate was calculated. RTs were analyzed for correct responses occurring in a time segment of 200 to 2500 ms following target onset. A fixed cutoff value (2500 ms) was selected for dealing with RT outliers (Ratcliff, 1993). Error rate was computed as the percentage of trials with misses, wrong responses, and responses occurring earlier than 200 ms after target onset. The two dependent variables entered separate ANOVAs crossing the within-subjects factors picture category (erotica, sports, neutral, attack, mutilation) and target type (word, pseudoword). Significant RT effects including picture category were followed by trend analyses. We expected a quadratic trend exhibiting a parabolic pattern of mean RTs with disproportionately strong modulation for the erotica and mutilation subcategories (i.e., erotica > sports > neutral < attack < mutilation). This prediction is in line with recent notions that affective modulation of perception is most evident for stimuli with high evolutionary significance (Schupp, Flaisch et al., 2006).

EEG Recording and Preprocessing

EEG recording and preprocessing (eye blink correction, bad channel interpolation and rejection of artifact contaminated epochs) paralleled the procedure described in the ERP study of part 2. Epochs were extracted from 500 ms before until 900 ms after target onset. A mean of 4.5 (SD = 2.2) channels per subject were interpolated, and on average 3.7 % epochs per conditions were rejected. Using the average reference, artifact-free epochs were averaged separately for the 10 combinations of picture category (erotica, sports, neutral, attack, mutilation) and target type (word, pseudoword). In order to achieve comparable signal-to-noise-ratios across combinations, we randomly drew 50 % of the trials in the neutral picture – word/pseudoword conditions. Only those trials with correct responses to targets were considered. Artifact and error trial exclusion resulted in a mean of 27.5 (SD = 0.8) epochs per picture x target combination and subject. Averaged data were filtered between 1 and 30 Hz, and the mean of a 200-ms segment prior to picture onset was subtracted as baseline.

ERP Analysis

Statistical analysis of the three selected ERP deflections – picture P2, target N1, and target late positivity (LP, see Results) – was performed using regional means. Sensor clusters were selected according to the spatial focus of the respective component in the grand average topographical maps. For analysis of the picture P2 and target N1, a symmetrical pair of sensor clusters was chosen comprising 9 electrodes over the left and 9 electrodes over the right occipital cortex (left-hemisphere electrodes 59, 65, 66, 69, 70, 71, 72, 74, 75; right-hemisphere electrodes 77, 83, 84, 85, 89, 90, 91, 92, 95). For the target LP component, parietocentral regional means were formed each containing 9 electrodes over the left (32, 38, 43, 48, 52, 53, 54, 60, 61) and right hemisphere (79, 80, 81, 86, 87, 88, 93, 94, 99). Figure 10 (left panel) shows the selected sensor clusters with respect to sites of the international 10-20 system. For each regional mean, peak latencies and mean amplitudes were examined for each participant. Peak latency of the picture P2 component was determined within a time range from 120 to 200 ms after picture onset (i.e., -80 to 0 ms relative to target onset). Target N1 latency was scored between 184 and 284 ms, and LP latency between 412 and 712 ms after target onset. Amplitudes for the picture P2 and target N1 were averaged across a 60-ms segment around an individual's peak. A longer segment (300 ms) was used around individual peaks of the LP. Mean amplitudes and peak latencies were submitted to ANOVAs having factors of picture category (erotica, sports, neutral, attack, mutilation), target type (word, pseudoword), and hemisphere (right, left). Similar to behavioral analyses, amplitude differences between picture categories were further evaluated by quadratic trend tests.

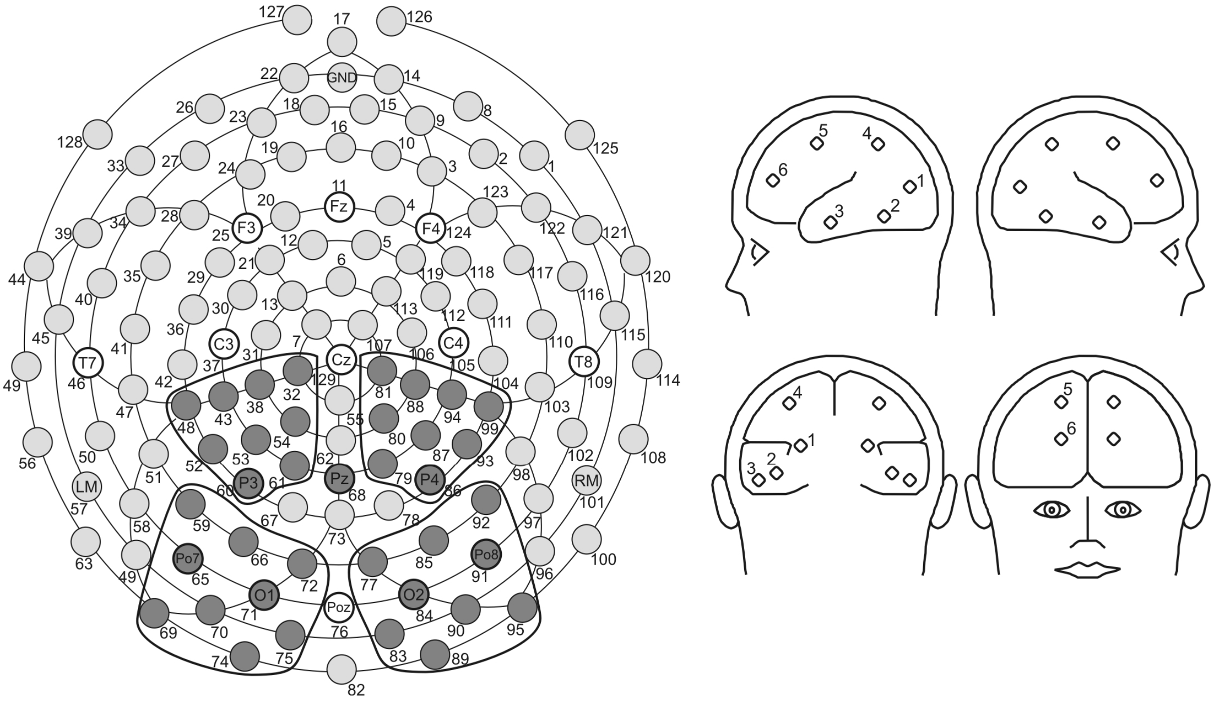

Figure 10. Left: Layout of the electrode array used in Experiment 1. Shaded electrode sites in the marked clusters were selected to form regional means entering statistical analyses. The posterior sensor clusters were used to quantify picture P2 and target N1 components, the central clusters were chosen to quantify the target late positivity (LP). Right: Regional source montage applied to single-subject ERPs. The model included 12 bilateral symmetric sources in occipital (labeled 1 in left hemisphere), posterior inferior temporal (pIT, 2), anterior inferior temporal (aIT, 3), parietal (4), central (5), and frontal (6) cortices. This procedure aimed to track word processing proceeding along the ventral stream in anterior direction with continued analysis.

Source Analysis

Given the goal of the present study to examine effects of emotional processing on subsequent lexico-semantic analysis, we aimed to increase the spatial and temporal specificity of our analysis by means of source modeling. To specifically highlight the temporal dynamics within and across different neuroanatomic structures, we employed discrete regional sources linked to landmark regions of language processing (Cohen & Dehaene, 2004; Hickok & Poeppel, 2004). Regional sources do not aim to provide precise neuroanatomic localization, but represent the multi-electrode ERP data in an anatomically meaningful low-dimensional space. This overcomes some of the problems associated with using voltage maps (Scherg, Ille, Bornfleth, & Berg, 2002). As a key advantage, the interpretation of source current amplitude enhancement as increase in brain electric activity is possible. Regional sources are sensitive to current changes irrespective of their orientation in space and therefore capture electrocortical activity originating from a wider range of areas. It is important to ensure that activity is not mis-represented in an area far from the true underlying source. Here, we used source sensitivity maps implemented in the Besa^{TM} software to measure whether the model dipoles were sensitive to local changes only. To calculate source sensitivity, the software models unit brain activity at different locations throughout the brain. Source sensitivity is defined as the fraction of power that is mapped on the selected source given the applied head model and regularization constant.

We seeded 12 symmetric bilateral sources in a 4-shell ellipsoidal head model (Berg & Scherg, 1994a). Positioning of regional dipoles drew on recent theories about the functional neuroanatomy of language (Cohen & Dehaene, 2004; Fernandez et al., 2001; Hickok & Poeppel, 2004), which posit that lexico-semantic analysis of verbal material is related to activation of the ventral processing system (Ungerleider & Mishkin, 1982). During visual word recognition, intracranial recordings in the inferior temporal cortex have revealed a directed processing stream proceeding in anterior direction with continued stimulus analysis (Fernandez et al., 2001). Left posterior inferior temporal areas also constitute a core element of theories which argue that visual word form activates specialized “word form areas” in the left fusiform gyrus (e.g., Cohen & Dehaene, 2004). The source montage as used here is illustrated in Figure 10 (right panel). It included symmetric bilateral dipoles in the occipital cortex as well as the posterior and anterior inferior temporal lobe (pIT and aIT). Left pIT was seeded according to the Talairach coordinates proposed “as the peak of the visual word form area” (McCandliss, Cohen, & Dehaene, 2003, p. 294) based on reviews of over 25 imaging experiments. Anterior IT was seeded to pick up ventral stream activity more rostrally. Further, the model included sources in parietal, central, and frontal cortices to control for activity outside the ventral system and to ensure sensitivity of each source for a particular brain region. Occipital, central, and frontal sources were equidistant to the midline; frontal and central sources were seeded according to equal z-locations. Talairach coordinates of the regional dipoles are listed in Table 3.

Table 3. Talairach coordinates of the regional sources used to model ERP data. Coordinates refer to sources in the left hemisphere. Right hemispheric sources were seeded symmetrically to the left sources, thus yielding deviant Coordinates only on the x-dimension.

	x	Y	z
Occipital	-25.0 (right 25)	-80.0	9.0
Posterior inferior temporal	-43.0 (right 43.0)	-54.0	-12.0
Anterior inferior temporal	-55.3 (right 55.3)	-2.9	-12.0
Parietal	-33.3 (right 33.3)	-54.8	46.1
Central	-19.2 (right 19.2)	0.9	52.0
Frontal	-19.2 (right 19.2)	45.0	27.1

The 12-source model was applied to individual ERPs. To reduce the number of comparisons, statistical evaluation of source intensity was restricted to three levels of picture category (pleasant, neutral, unpleasant). For each participant, mean source activity in the N1 window was calculated between 204 and 264 ms after target onset representing a 60-ms segment around N1 mean latency (234 ms). Mean source intensity in the LP window was obtained from 404 to 704 ms after target onset representing a 300-ms segment around LP mean latency (554 ms). For each source and target type (word, pseudoword), mean intensity values were submitted to ANOVAs using picture category as within-subject factor. Due to the augmented number of comparisons, the alpha-level was adjusted to 1%. Significant effects were analyzed by means of Newman-Keuls tests.

EXPERIMENT 2

Participants

12 university students consented to participate in Experiment 2 (11 females, mean age: 22.2 years, SD = 4.2).

Stimuli

Word/pseudoword targets were identical to Experiment 1. With regard to picture stimuli, three categories were used: erotica (high-arousing pleasant, 60 items), neutral persons (low-arousing, 60 items), and mutilation (high-arousing unpleasant, 60 items) pictures. The same pictures as in Experiment 1 were employed plus additional items required for the erotica and mutilation category. Again, the entire picture set comprised two semantically congruent subsets assigned to word versus pseudoword targets in a counterbalanced manner across participants.

Procedure

Experimental setting, physical stimulus properties, and response modalities were the same as in Experiment 1. Targets followed pictures at three different SOAs (80 ms, 200 ms, 440 ms). In the short SOA condition, targets were presented immediately after picture offset. In the 200 ms and 440 ms conditions, pictures were shown for 100 ms, followed by a blank screen with a duration of 100 ms and 340 ms, respectively. Each of the 180 pictures (90 assigned to words, 90 assigned to pseudowords, see above) was shown once at each SOA resulting in 540 trials, which were presented in random order and included two breaks. Again, picture-word pairs were separated by an inter-trial-interval randomly varying between 2500 ms and 3500 ms. After the lexical decision task, participants rated the pictures using the computerized SAM version described above. To reduce the length of the experiment, only those pictures were judged that had preceded word targets.

Data Analysis

RTs and error rates were calculated as specified in Experiment 1. The two measures were submitted to repeated measures ANOVA with the factors SOA (80 ms, 200 ms, 440 ms), picture category (pleasant, neutral, unpleasant), and target category (word, pseudoword). Mean picture ratings of each participant were analyzed by one-factorial ANOVA using picture category as within-subject factor. Significant effects were followed by Newman-Keuls tests.

EXPERIMENT 3

Participants

Data were collected from 16 university students (14 females) having a mean age of 21.3 (SD = 2.7) years. Rating data from one subject were excluded because of lack of compliance.

Stimuli

Verbal targets were the same as in the previous experiments. Pictures were 180 items drawn from the IAPS (Lang et al., 2005). Picture selection aimed at building 5 stimulus categories (low-arousing pleasant [PL], high-arousing pleasant [PH], low-arousing neutral [NL], low-arousing unpleasant [UL], high-arousing unpleasant [UH]; N = 36 items in each category) that were distinctive with regard to normative affective ratings but semantically heterogeneous. For instance, stimuli included both, low-arousing and high-arousing erotica, or low-arousing and high-arousing attack scenes. According to the normative IAPS ratings, valence scores ranged from 6.1 to 8.2 for PL, from 5.6 to 8.0 for PH, from 4.2 to 5.5 for NL, from 1.8 to 3.0 for UL, and from 1.3 to 3.8 for UH pictures. Arousal was between 2.9 and 4.9 for PL, between 5.3 and 7.2 for PH, between 2.4 and 4.2 for NL, between 3.9 and 5.8 for UL, and 5.8 and 7.3 for UH pictures. Figure 11 illustrates the distribution of the pictures in an affective space. Words/pseudowords were assigned to pictures by using the procedure described in Experiment 1. All pictures were rated after the experiment.

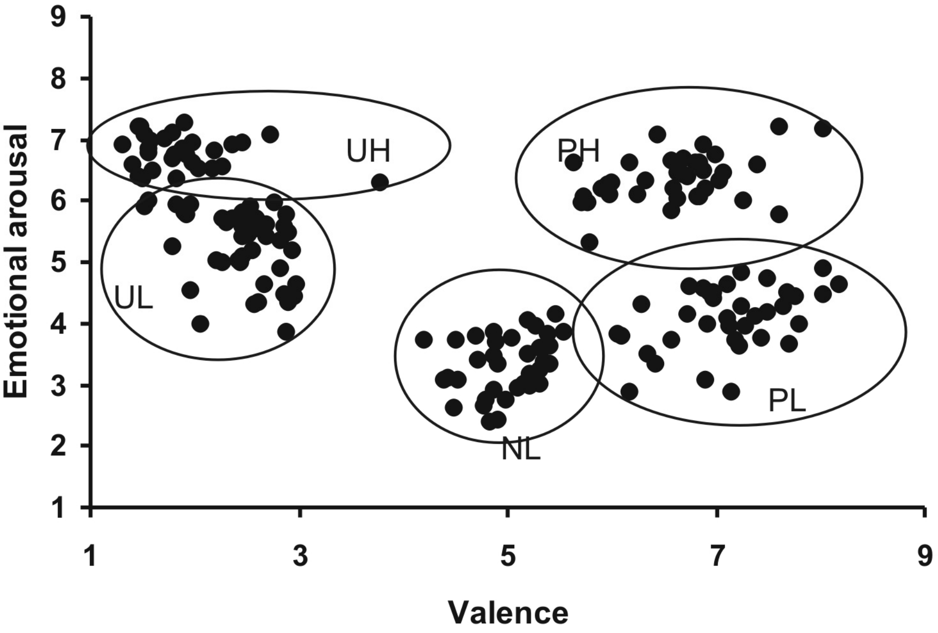

Figure 11. Affective space for the picture stimuli used in Experiment 3 defined by each picture’s normative valence and arousal rating (PH = high-arousing pleasant, PL = low-arousing pleasant, NL = low-arousing neutral, UL = low-arousing unpleasant, UH = high-arousing unpleasant).

Procedure

The procedure was identical to Experiment 1.

Data Analysis

Analysis of responses paralleled the approach employed in the previous experiments. RTs and error rates were analyzed by means of 5 (picture category; PH, PL, NL, UL, UH) x 2 (target category; words versus pseudowords) repeated measurement ANOVA. A quadratic trend was fitted to the data (see Experiment 1) expecting a moderate RT increase for PL and UL pictures and strong impairment for PH and UH items. Ratings were analyzed with ANOVA as specified in Experiment 2. In addition, a correlational analysis was performed. For each participant, mean differences were calculated between arousal ratings for emotional categories versus the neutral category, and these values were correlated with the respective RT differences. In addition to these by-subject analyses, we tested the predictive value of single pictures' arousal ratings with regard to modulatory effects on RTs. This by-item analysis was done by using single pictures as observations and correlating pictures' mean arousal ratings with mean RTs.

Results

EXPERIMENT 1

Behavioral Data

Consistent with previous work (Hauk, Davis, Ford, Pulvermüller, & Marslen-Wilson, 2006), subjects responded slower to pseudoword compared to word targets, $F(1, 21) = 29.0$, $p < 0.001$. As indicated by a significant main effect of picture category, $F(4, 84) = 3.9$, $p < 0.01$, and a significant interaction of picture category x target type, $F(4, 84) = 8.0$, $p < 0.001$, response latencies were also modulated by emotional content of the preceding stimulus. As illustrated in Figure 12, word RTs followed an U-shaped function with elevated values, i.e. slowed RTs, for pictures depicting erotic and mutilation scenes (quadratic trend, $F(1, 21) = 21.7$, $p < 0.001$). We did not find such a quadratic trend for responses to pseudowords, $F < 1.0$. Similar to RT analysis, ANOVA on error rates showed a significant interaction between picture category and target type, $F(4, 84) = 6.1$, $p < 0.001$. Post-hoc tests (Newman-Keuls) revealed higher error rates for word responses preceded by mutilation versus neutral pictures, $p < 0.05$. Error rates on pseudoword trials did not differentiate between neutral and emotional pictures, $ps > 0.53$.

Taken together, behavioral data indicate that lexico-semantic processing is impaired and *not* facilitated subsequent to information containing arousing appetitive or aversive content, such as erotic or mutilation scenes. Costs of processing these stimuli, however, are only evident for word but not for pseudoword responses. The question arises why arousing sports and attack pictures did not delay word responses (see Figure 12). We speculate that the small increase of subjectively rated arousal from sports to erotica, and from attack to mutilation, respectively (see Figure 12, right panel), underestimates the impact of pictures displaying content of high evolutionary significance (Schupp, Flaisch et al., 2006), such as sexuality or physical damage.

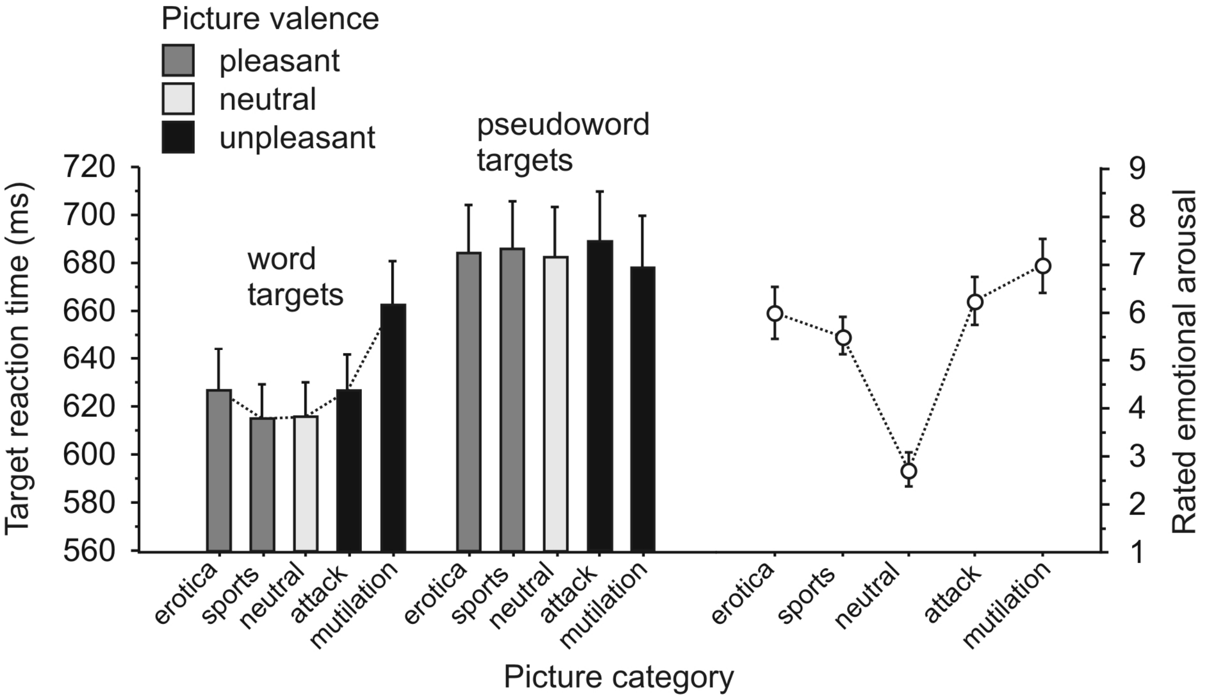

Figure 12. Experiment 1: Left panel: Mean reaction times (RTs) as a function of target type and category of the preceding picture. Right panel: Mean emotional arousal ratings of the presented pictures as obtained in a pre-experimental study (1 = low arousing, 9 = high arousing). The dotted line connecting mean RTs to words reflects a quadratic trend. Error bars indicate standard errors (standard deviations for arousal ratings).

Electrophysiological Data

Over posterior scalp, grand average ERPs were characterized by a series of deflections reflecting the rapid sequence of visual events (i.e., picture - mask - target) within each trial (see Figure 13). These posterior components demonstrated polarity reversal over anterior electrodes. Three deflections were reliably detected in all participants: (i) a positive component peaking 160 ms following picture onset, with a bilateral occipital maximum (picture P2); (ii) a bilateral occipitotemporal negativity reaching a maximum amplitude at about 230 ms after word/pseudoword onset (target N1); and (iii) a late positivity over centroparietal sensors extending from 400 to 700 ms post-target (target LP). Deflections between picture P2 and target N1 were not analyzed as they likely reflected sensory responses to the mask (see Figure 13). Regarding target ERPs, we replicated the electrocortical pattern recently reported by Hauk et al. (2006). Using a forced-choice lexical decision task with English nouns, these authors observed a prominent occipital negativity and a centroparietal positivity around 200 and 500 ms, respectively.

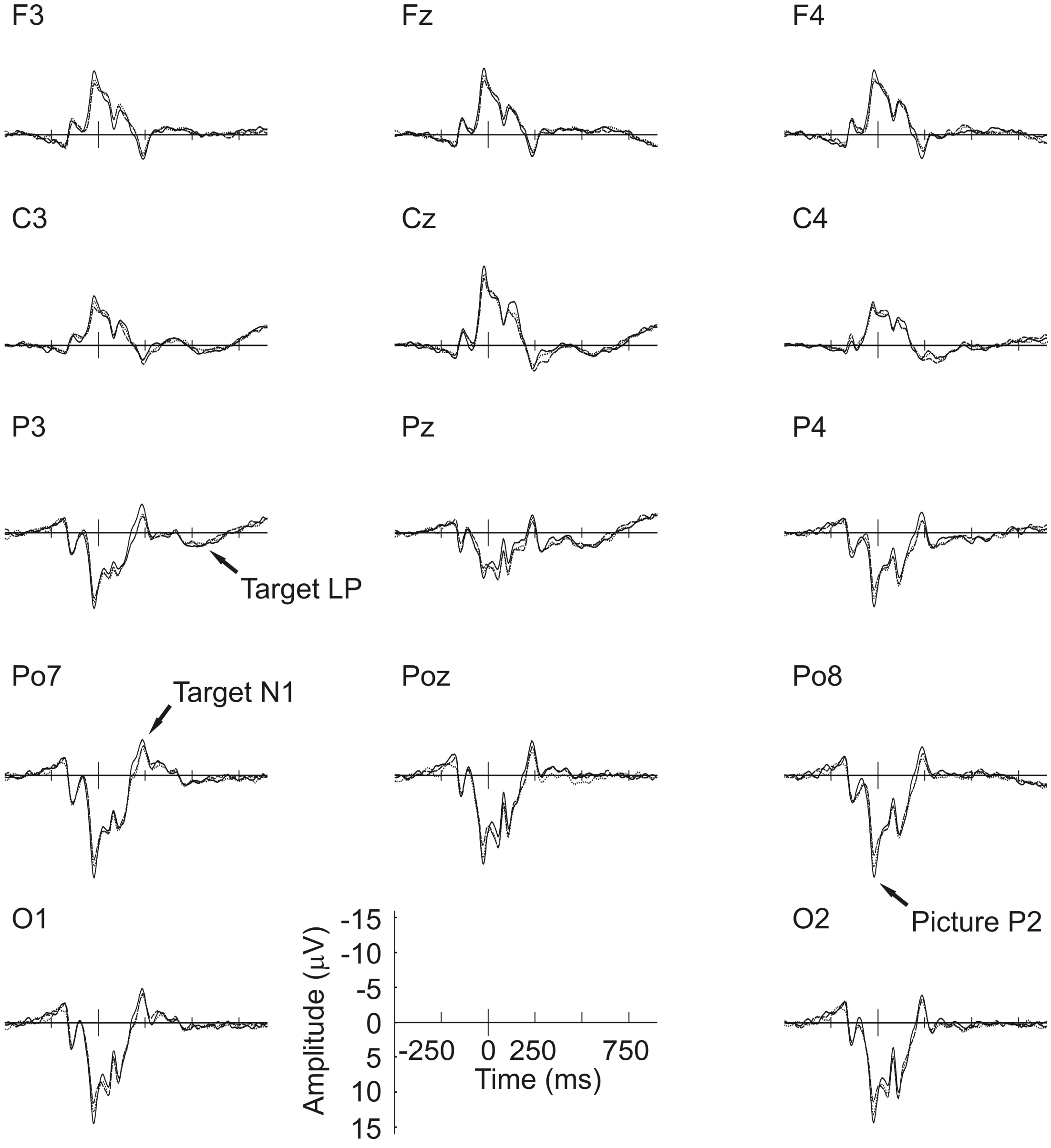

Figure 13. Grand mean ERPs to word targets following pleasant (dashed line), neutral (solid line), and unpleasant (dotted line) pictures at a subset of electrodes corresponding to sites of the international 10-20 system. Time zero represents word target onset.

ERP waveforms comprising the present picture P2 and target N1 at a representative left posterior electrode (PO7) are shown in Figure 14.

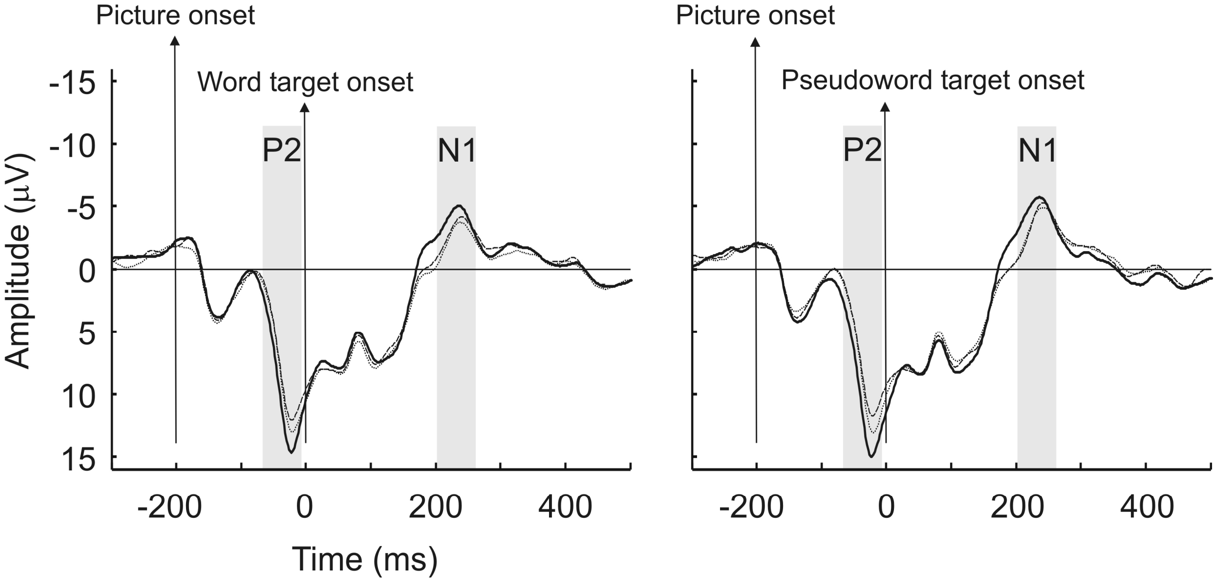

Figure 14. Picture P2 and target N1 at a representative left occipital sensor (# 65, corresponding to PO7). Waveforms show grand mean ERPs in response to words (left panel) and pseudowords (right panel) following pleasant (dashed line), neutral (solid line), and unpleasant (dotted line) pictures. Shaded bars illustrate that picture P2 and target N1 were statistically analyzed by extracting mean amplitudes across a 60-ms time segment around individual subjects' peak latencies.

Picture P2

Picture P2 amplitudes varied as a function of picture category, F(4, 84) = 36.2, p < 0.001, while there were no effects of target type or hemisphere, all Fs < 1.1. As can be seen in Figure 14, the P2 was more negative for emotional compared to neutral pictures. Trend analysis showed that across target types and hemispheres negativity increased slightly from neutral to sports/attack pictures and strongly from neutral to erotica/mutilation pictures, quadratic trend: F(1, 21) = 59.2, p < 0.001. This finding replicates previous work showing early ERP negativity as a function of emotional intensity (Schupp et al., 2003) and suggests that subjects discriminated early (and prior to target onset) between different affective picture categories. Peak latencies were not sensitive to picture category, F < 1.0, or target type, F < 1.1.

Target N1

A significant main effect of target type, F(1, 21) = 22.3, $p < 0.001$, indicated that N1 responses were smaller for words compared to pseudowords. N1 responses were also modulated by picture category, F(4, 84) = 8.0, $p < 0.001$, reflecting reduced amplitudes for targets following arousing pictures. Trend analysis confirmed that across hemispheres and target types, erotic and mutilation pictures suppressed N1 responses markedly, whereas sports and attack scenes elicited no or only a small N1 reduction (see Figure 15), quadratic trend: F(1, 21) = 13.6, $p < 0.01$. This result corroborates the notion that the observed effects were driven by affective-semantic factors and not by physical differences between the pleasant, neutral, and unpleasant stimulus categories.

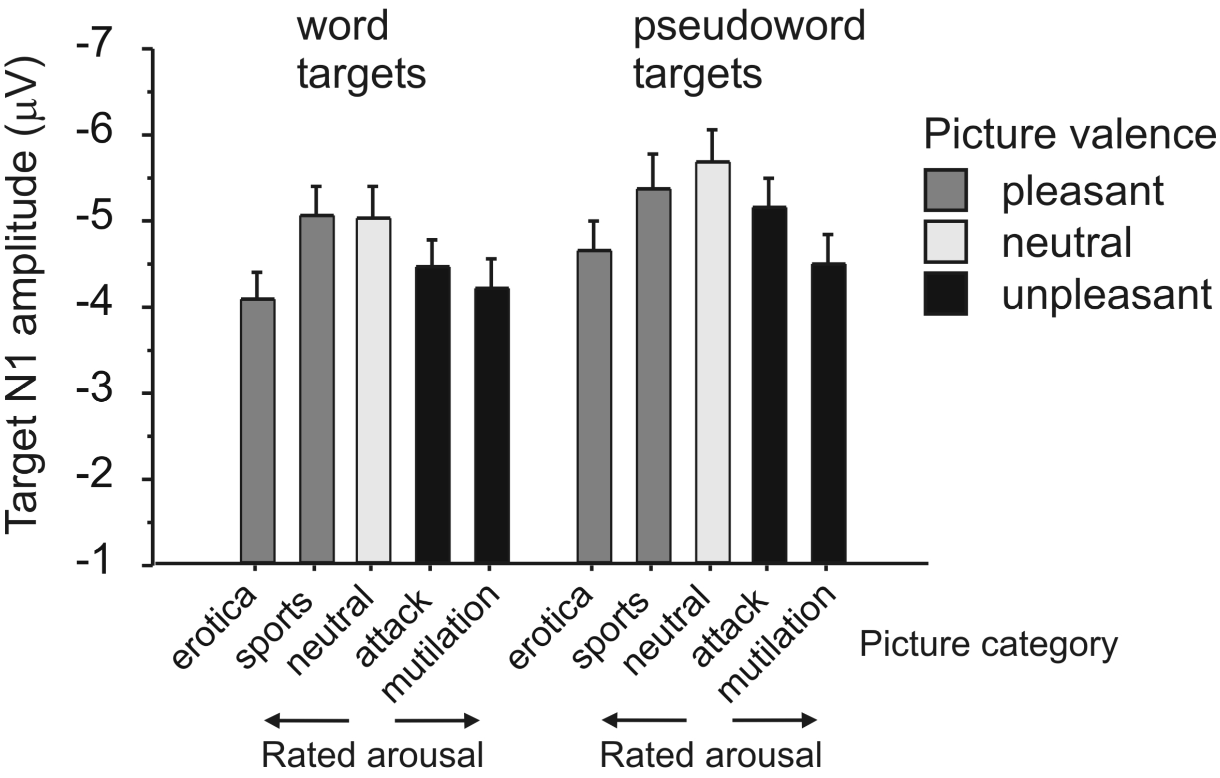

Figure 15. Mean target N1 amplitudes as a function of target type and category of the preceding picture. For both targets (words and pseudowords), the N1 decreased with increasing emotional intensity of the preceding picture, irrespective of picture valence. Error bars represent standard errors.

ANOVA on N1 peak latencies revealed that picture category also affected temporal aspects of processing, F(4, 84) = 4.9, $p < 0.01$. Across target types and hemispheres, N1 maxima were clearly delayed for erotica (mean latency: 241 ms, SE = 2.7) and mutilation (239 ms, SE = 2.3) pictures, compared to neutral pictures (229 ms, SE = 2.2). Again, interference, i.e., N1 latency increase, was less pronounced for sports (235 ms, SE = 2.2) and attack (234 ms, SE = 1.9) pictures, resulting in a significant quadratic trend: F(1, 21) = 9.0, p < 0.01. Notably, as opposed to the RT data, effects of emotion on the N1 were not sensitive to target lexicality showing reduced and delayed responses to words and pseudowords.

Target Late Positivity

Figure 16 depicts the target LP[6] at a representative left central sensor.

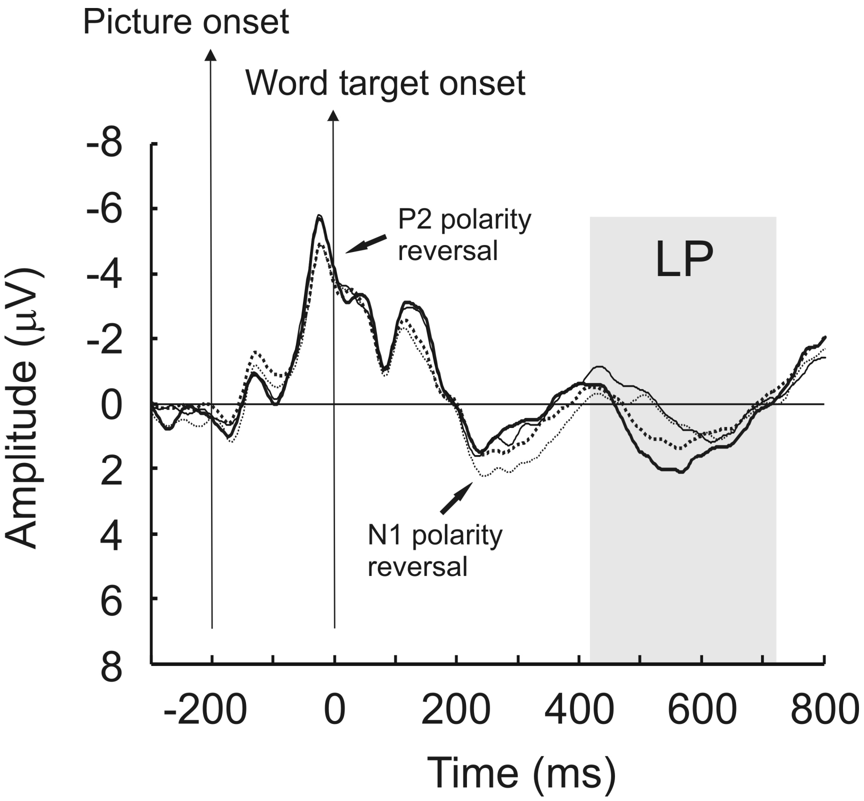

Figure 16. Target LP at a representative left central sensor (# 38). Waveforms illustrate grand mean ERP responses to words following neutral pictures (thick solid line), to pseudowords following neutral pictures (thin solid line), to words following unpleasant pictures (thick dotted line), and to pseudowords following unpleasant pictures (thin dotted line). The shaded bar marks the 300-ms time window used to quantify LP mean amplitudes around individual subjects' LP peaks.

ANOVA on LP amplitudes showed significant main effects of picture category, F(4, 84) = 3.9, p < 0.01, and target type, F(1, 21) = 23.8, p < 0.001, as well as an interaction between these factors F(4, 84) = 5.0, p < 0.01. Hemisphere also affected LP amplitudes, main

---

[6] Note that the target LP observed here most likely reflected late cognitive processing of neutral words/pseudowords. It is not to be confused with the late positive potential (LPP) elicited during affective picture viewing (see Experimental Series A), which may reflect resource allocation to motivational significant stimuli.

effect: F(1, 21) = 4.6, $p < 0.05$, hemisphere x target: 4.9, $p < 0.05$. As can be seen in Figure 17, the LP was larger for words than pseudowords, and in the left compared to the right hemisphere. Further, the difference between words and pseudowords was more pronounced in the left hemisphere. Trend analyses testing the effects of picture category showed that emotional pictures reduced word LPs in the predicted manner in the left, quadratic trend: F(1, 21) = 4.9, $p < 0.05$, and in the right hemisphere, F(1, 21) = 6.5, $p < 0.05$. The quadratic trend did not reach significance for pseudowords in either hemisphere (left: F(1, 21) = 3.3, $p = 0.08$, right: F(1, 21) = 4.3, $p = 0.05$. That is, similar to the RT data, interference was more evident for word compared to pseudoword targets.

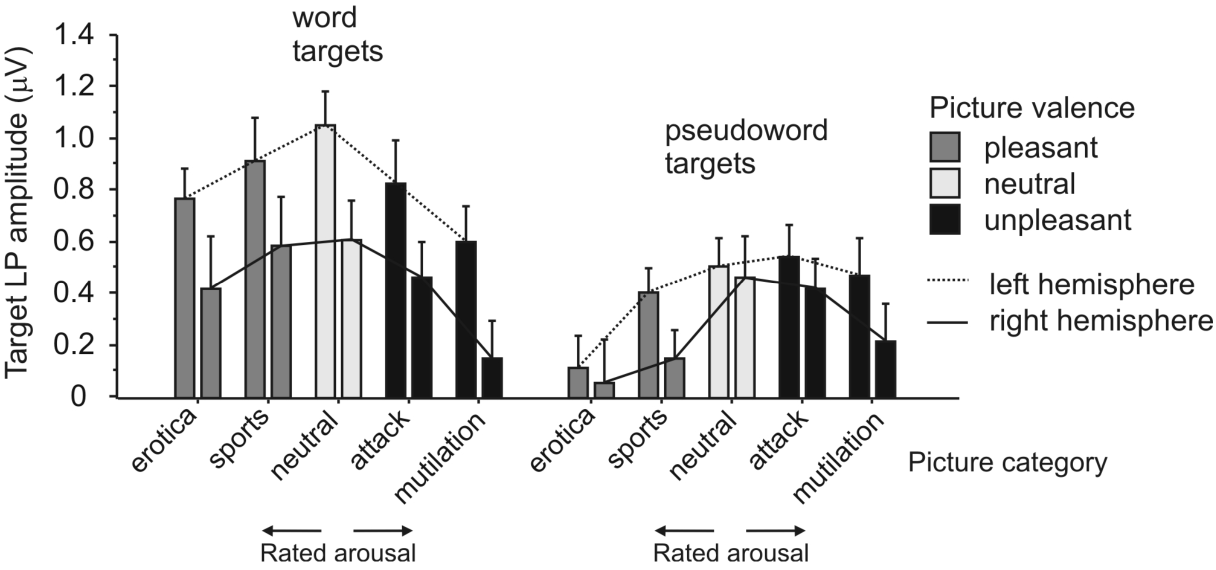

Figure 17. Mean target LP amplitudes as a function of target type, picture category, and hemisphere. The LP was reduced for targets subsequent to emotional information. For word targets, LP reduction followed the U-shaped function of picture arousal ratings. This pattern was less pronounced for pseudoword responses. Error bars indicate standard errors.

Peak latency analysis indicated that the LP reached maximum earlier for word (mean latency: 540 ms, SE = 5.2) than pseudoword targets (mean latency: 564 ms, SE = 5.6), target type: F(1, 21) = 6.9, $p < 0.05$. ANOVA also yielded significant interactions between picture category and target type, F(4, 84) = 3.1, $p < 0.05$, and between picture category, target type, and hemisphere, F(4, 84) = 3.8, $p < 0.01$. Trend analyses revealed a significant quadratic trend for word targets in the left, F(1, 21) = 8.7, $p < 0.01$, but not in the right hemisphere: $F < 1.0$. For pseudowords, quadratic trends did not reach significance, left: $F < 1.0$, right: F(1, 21) = 3.9, $p = 0.06$. In contrast to the N1 delay, the emotion effect on LP maxima was reflected in shorter latencies for (left-hemispheric) word responses following erotica (517 ms, SE = 19.0) and mutilation (527 ms, SE 14.0) pictures and slightly earlier peaks for word responses following sports (554 ms, SE = 9.6) and attack (546 ms, SE = 12.7) pictures, compared to the neutral category (556 ms, SE = 13.5). One might speculate that this finding reflects incomplete or shallower processing for targets subsequent to arousing pictures, resulting in earlier LP peak latencies and smaller LP amplitudes. Future work may employ approaches such as Independent Component Analysis to disentangle the differential contributions of stimulus processing, response selection, and response preparation in the current task.

Regional Source Modeling

Sensitivity maps of the regional dipoles (see Figure 18, left panel) demonstrated that the sources were spatially independent and reflected activity from non-overlapping cortical areas. Across conditions, the model accounted for more than 96.8 % of the variance in the grand average global field power. As illustrated in the time-varying activity plots (Figure 18, middle panel), source modeling yielded four activity peaks corresponding to picture P2, mask onset (see section Electrophysiological Data), target N1, and target LP. The sources demonstrated substantial differences in their overall contribution to those activity peaks and a distinct pattern of intensity changes across time. For the picture P2 and target N1, the model was dominated by bilateral activity in the pIT with moderate contributions of occipital, aIT, parietal, and frontal sources. The target LP was associated with activity mainly originating from left sources in the aIT region and minor contributions of parietal and frontal sources. Central sources did not show major intensity changes pertinent to picture or target onset.

The question arises whether and where the source waveforms for different picture categories displayed a pattern similar to the condition differences in the scalp potentials. As can be seen in Figure 18, at pIT, activity in the picture P2 window was more pronounced for neutral, compared to pleasant and unpleasant pictures. Conclusions from this result must be drawn with caution. First, seeding of pIT sources aimed at modeling language-related activity in areas of lexico-semantic analysis (see Methods) and not at finding a best-fit model of early negativities. Second, prior to the P2 window, right pIT and occipital source waveforms displayed the expected activity amplification for arousing relative to neutral pictures. This may reflect that the early enhancement for emotionally arousing stimuli may start with an increase of the N1 component (Keil et al., 2001), and then continues as a relative negative shift within a positive-going deflection (the P2).

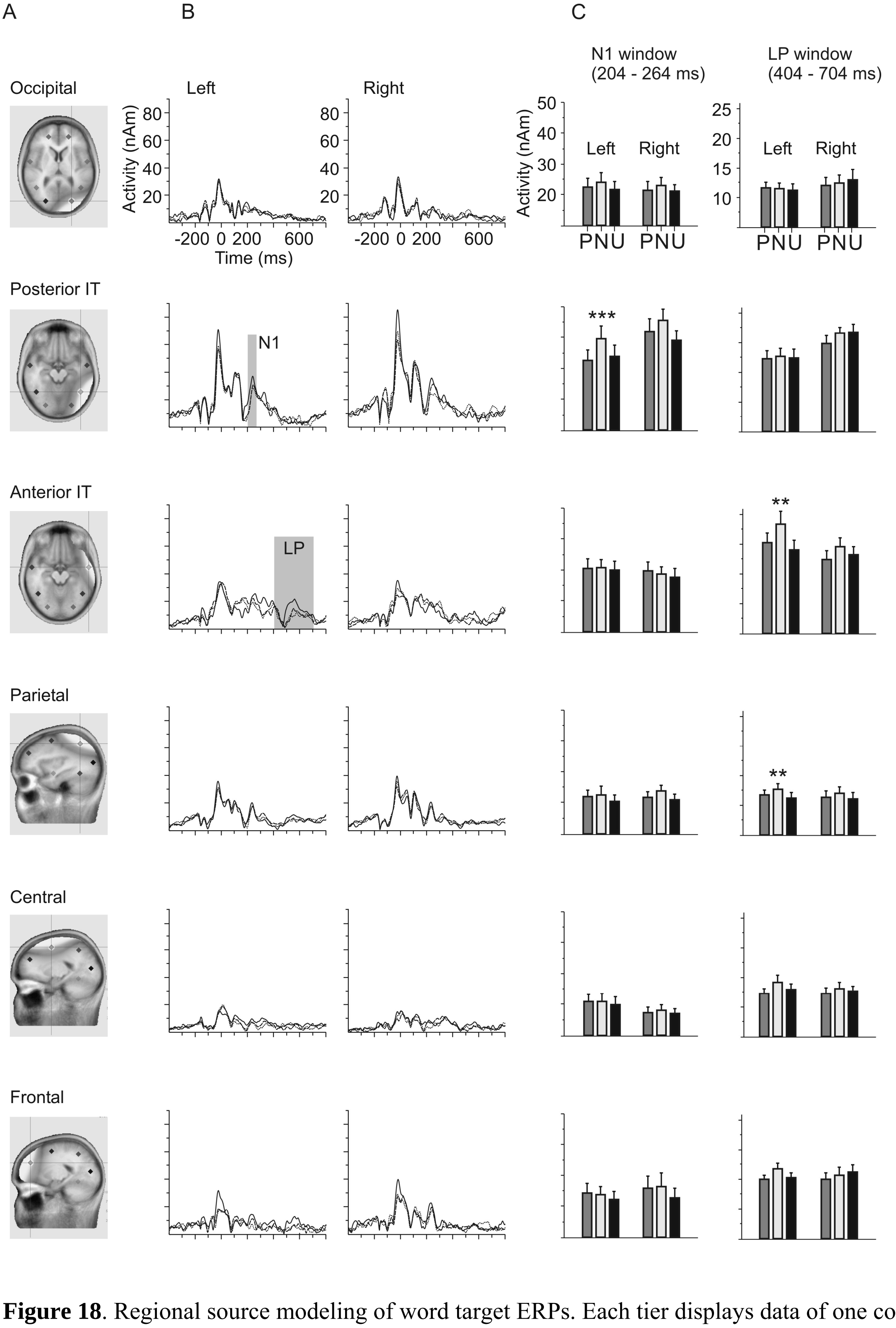

Figure 18. Regional source modeling of word target ERPs. Each tier displays data of one cortical region, i.e., bilateral symmetric sources in occipital, posterior inferior temporal (pIT), anterior inferior temporal (aIT), parietal, central, and frontal cortices. A: Sensitivity maps of the regional sources illustrating the spatial independence of the used dipoles. Only maps of left-hemispheric sources are shown. B: Time-varying activity plots of source responses to words following pleasant (dashed line), neutral (solid line), and unpleasant pictures (dotted line). Source waveforms were obtained by

applying the source montage to grand mean ERP data. Time zero represents word onset. C: Mean source intensity in the N1 and LP time window as a function of picture category (P = pleasant, N = neutral, U = unpleasant). Values were obtained by applying the source model to individual subjects' word ERPs and extracting mean activity in the respective time segment. In the N1 window, arousing pictures suppressed source activity in the left pIT. In the LP window, an effect of picture category was present in left aIT and parietal sources (*** $p < 0.001$, ** $p < 0.01$). Error bars represent standard errors.

As described in the Methods section, effects of picture category (pleasant, neutral, unpleasant) on word/pseudoword processing were statistically evaluated by calculating mean amplitudes in the target N1 (204-264 ms) and in the target LP (404-704 ms) time segment, for each source. In the early (N1) segment, a main effect of picture category emerged in the left pIT (see right panel of Figure 18), $F(2, 42) = 8.3$, $p < 0.001$, reflecting reduced activity for words following pleasant and unpleasant pictures, compared to words following neutral exemplars, Newman Keuls' ps < 0.01. For pseudoword targets, affective modulation in the N1 range was evident in the left pIT, $F(2, 42) = 6.5$, $p < 0.01$, and at the left occipital source, $F(2, 42) = 5.6$, $p < 0.01$. As for words, post-hoc tests showed a reduction for pseudowords preceded by pleasant, left: pIT $p < 0.01$, left occipital: $p < 0.05$, and unpleasant scenes, left pIT: $p < 0.05$, left occipital: $p < 0.01$, compared to neutral pictures. During this time window, the difference between targets following arousing versus non-arousing information did not reach the adjusted alpha-level of 1 % at any other source.

In the late (LP) time segment, significant effects of picture category were obtained for word targets in the left aIT, $F(2, 42) = 6.6$, $p < 0.01$, and left parietal cortex, $F(2, 42) = 6.6$, $p < 0.01$. Relative to neutral pictures, source activity was reduced for words following pleasant, ps < 0.05, and unpleasant scenes, ps < 0.01. Source activity for pseudowords did not vary as a function of affective picture category.

EXPERIMENT 2

Post-experimental Affective Ratings

Picture category modulated valence ratings, $F(2, 11) = 97.8$, $p < 0.001$, and arousal ratings, $F(2, 11) = 74.0$, $p < 0.001$. Confirming that participants perceived the pictures as intended, pleasant pictures yielded higher valence scores (mean valence: 6.2, SD = 1.2), Newman-Keuls $p < 0.05$, and unpleasant items obtained lower values (1.6, SD = 0.3), $p < 0.001$, compared to neutral stimuli (5.4, SD = 1.0). Both affective categories (pleasant: mean arousal: 4.8, SD = 1.5, unpleasant: 7.3, SD = 0.9) were rated as more arousing than the neutral category (2.1, SD = 1.2), ps < 0.001.

Lexical Decision

ANOVA on RTs yielded significant main effects of SOA, F(2, 22) = 28.4, and target type, F(1, 11) = 16.6, ps < 0.01, and a significant interaction between both factors, F(2, 22) = 4.1, p < 0.05. Post-hoc tests showed that for both target types, RTs were faster at the short SOA (80 ms) and slower at the long SOA (440 ms), compared to the 200 ms SOA, ps < 0.01. Importantly, we replicated the main effect of picture category, F(2, 22) = 10.5, and the interaction effect between picture category and target type, F(2, 22) = 5.9, p < 0.01, while ANOVA did not show any significant interactions involving SOA and picture category. That is, at all SOAs pleasant and unpleasant pictures elicited slower word responses than neutral pictures, ps < 0.01, while emotional content did not modulate pseudoword responses, ps > 0.49 (see Figure 19). These finding also suggest that the magnitude of interference exerted by erotica and mutilation pictures remains constant although task difficulty increases (indexed as an overall increase of RTs for shorter SOAs). ANOVA on error rates yielded a significant interaction between picture category and target type, F(2, 22) = 11.1, p < 0.001. Across SOAs, unpleasant pictures, compared to neutral exemplars, provoked higher error rates during word trials, p < 0.05. Errors on pseudoword trials were not modulated by picture category, ps > 0.05.

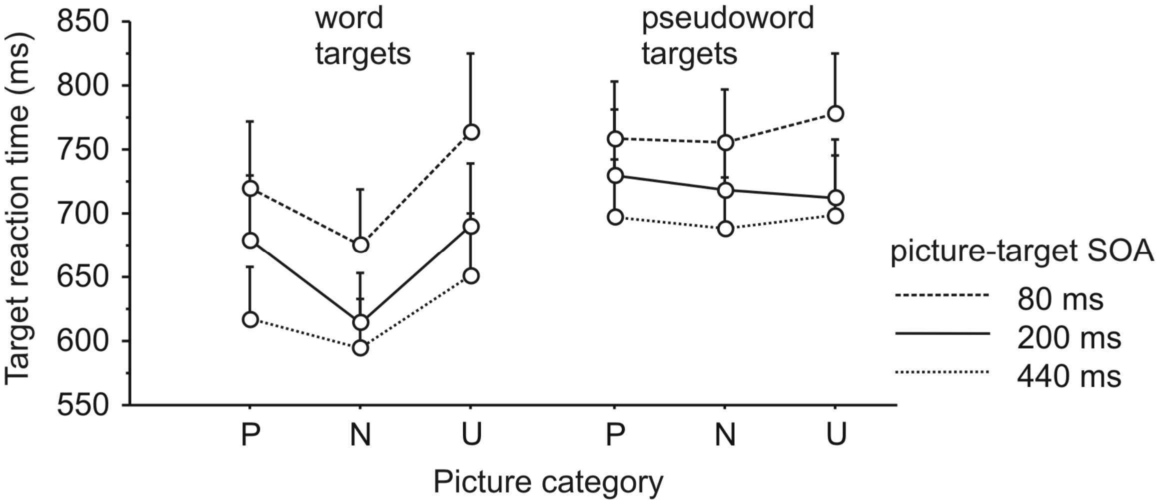

Figure 19. Experiment 2: Mean RTs as a function of picture-target SOA, target type, and category of the preceding picture (P = pleasant, N = neutral, U = unpleasant). At all SOAs, pleasant and unpleasant pictures impaired word but not pseudoword responses. Error bars indicate standard errors.

The implications of Experiment 2 are two-fold: i) Affective interference as observed here is not a short-lived/transient phenomenon or a mere result of stronger sensory forward masking of emotional pictures. Otherwise, performance deficits should have vanished at the latest for the 400 ms SOA. Instead, erotica and mutilation pictures appear to trigger an affective response having sustained negative consequences for subsequent processing. This assumption is also supported by work form our laboratory showing persisting emotional interference for up to 1000 ms during processing of RSVP streams. ii) Already shortly after the emotional picture, encoding of new information is impeded. The occurrence of facilitation (Phelps et al., 2006) versus interference (Most et al., 2005) may thus be determined by other task aspects, such as complexity or the degree of emotional arousal evoked by the lead stimuli.

EXPERIMENT 3

Post-experimental Affective Ratings

Ratings varied with picture category on both SAM dimensions, valence ratings: F(4, 14) = 219.0, arousal ratings: F(4, 14) = 65.3, ps < 0.001. As intended, both PH (mean valence = 6.3, SD = 1.0) and PL (7.1, SD = 0.7) pictures evoked higher pleasure scores than neutral items, ps < 0.001. Similarly, UH (1.5, SD = 0.5) and UL (2.6, SD = 0.6) pictures both were related to lower unpleasantness, compared to neutral pictures, ps < 0.001. For unpleasant items, subjectively rated arousal was higher for UH pictures (mean arousal = 7.6, SD = 0.9), compared to UL (5.7, SD = 1.2) and NL (2.8, SD = 0.9) pictures, ps < 0.001. PH pictures (5.1, SD = 1.6) were rated as more arousing, compared to PL (2.9, SD = 1.2) and NL items, ps < 0.001.

Lexical Decision

Error rates were higher for pseudowords than words, F(1, 15) = 17.1, p < 0.001, but not sensitive to influences of picture category, Fs < 2.0. Again, pseudoword responses were also slower than word responses, F(1, 15) = 17.1, p < 0.001. Further, ANOVA on RTs yielded a main effect of picture category, F(1, 15) = 6.7, p < 0.001. The significant interaction picture category x target type, F(4, 60) = 2.7, p < 0.05, indicated that, similar to the previous experiments, picture category affected word and pseudoword processing differently. While word responses showed an u-shaped interference pattern with most delayed RTs for the high-arousing categories (PH and UH), quadratic trend F(1, 15), = 14.8, p < 0.01, modeling of pseudoword RTs did not yield a significant parabolic relationship, F(1, 15) = 2.0, p = 0.17 (see Figure 20).

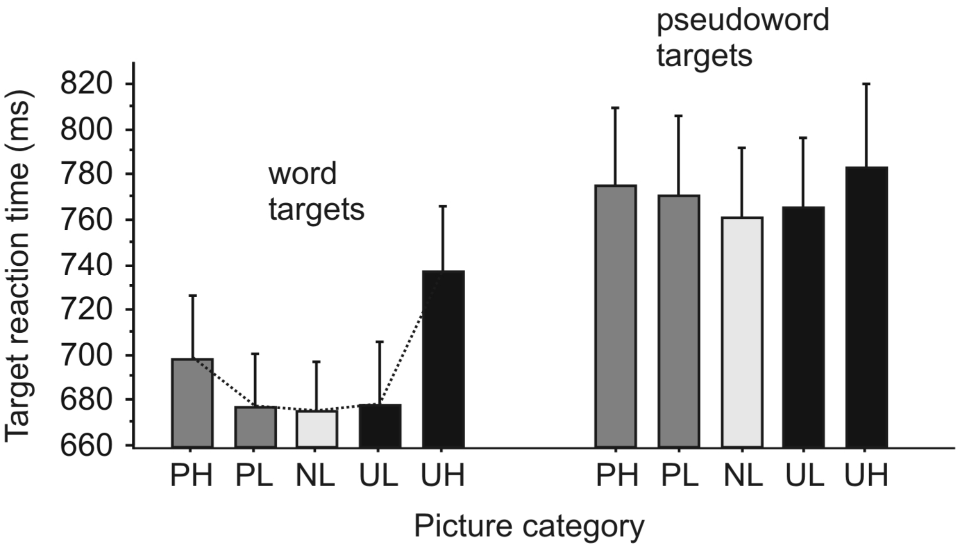

Figure 20. Experiment 3: Mean RTs as a function of target type and category of the preceding picture (PH = high-arousing pleasant, PL = low-arousing pleasant, NL = low-arousing neutral, UL = low-arousing unpleasant, UH = high-arousing unpleasant). Word responses were delayed for high-arousing (pleasant and unpleasant) pictures but not for low-arousing items. Error bars indicate standard errors.

Correlation Analyses

On the level of individual subjects, the differences between arousal scores for emotional (PH, PL, UL, UH) versus neutral pictures significantly correlated with the respective differences of word RTs, r = 0.33, p < 0.01. No such relationship was found for pseudowords, r = 0.04, p = 0.74. Similarly, in an item-level correlation analysis calculated for single pictures, mean arousal scores given to pictures predicted mean RTs in the word, r = 0.28, p < 0.001, but not in the pseudoword condition, r = 0.12, p = 0.11.

Taken together, both ANOVA and correlation analyses indicate a parametric relationship between picture arousal and response impairment, specifically when processing word targets. Although pleasantness/unpleasantness was comparable for PL and PH pictures, and UL and UH pictures, respectively, only the high-arousing items impeded responses. Further, as picture categories were semantically heterogeneous (e.g., both PH and PL pictures comprised erotica), Experiment 3 thus provided strong evidence against the potential that RT interference in the present design is caused by salient semantic category differences.

GENERAL DISCUSSION OF SERIES B

Enhanced perceptual processing of arousing versus non-arousing visual information is a pivotal finding in research on affective stimulus processing (e.g., Keil, Ihssen, & Heim, 2006; Schupp et al., 2003). In the present study, we aimed to track the consequences of such effects for a subsequent cognitive task. Verbal targets were presented after task-irrelevant picture stimuli, and participants executed a speeded lexical decision. Pictures varied in subjectively rated hedonic valence and emotional arousal.

At the level of behavioral output, we found delayed response latencies for word targets following high-arousing erotica and mutilation pictures at an SOA of 200 ms (Experiment 1). By employing different SOAs, in Experiment 2 the hypothesis was tested that behavioral interference versus facilitation is contingent on the temporal interval between picture and target. As a main result, emotionally arousing pictures interfered with lexical processing irrespective of SOA. Using the same SOA but different picture categories, Experiment 3 intended to disentangle the role of valence versus emotional intensity in the current paradigm. Emotional arousal now was varied within the aversive and appetitive categories which included semantically mixed stimulus exemplars. We demonstrated that responses were impaired for pleasant and unpleasant pictures rated as highly arousing but not for affective exemplars related to low arousal scores. Again, interference was stronger for lexical targets.

At the level of electrocortical scalp potentials (Experiment 1), the N1-P2 complex to picture stimuli displayed a negative amplitude shift for arousing pictures, especially for erotica and mutilation scenes (cf. Schupp et al., 2003). With regard to target processing, those stimuli reduced N1 amplitudes and delayed N1 peak latencies both for word and pseudowords responses. Inhibitory effects of emotional content were demonstrated for the target LP as well. In contrast to the target N1, LP amplitudes were found to vary as a function of target type and picture category showing stronger impairment for word responses.

Taken together, ERP amplitudes and RTs in the present work were supportive of detrimental effects of affective stimuli on subsequent processing. Further, across experiments the behavioral and electrocortical measures exploited here suggest interference as a function of emotional arousal. Facilitation as indicated by enhanced ERP components (Pourtois et al., 2004) or increased perceptual acuity (Phelps et al., 2006) have been recently reported for neutral information subsequent to emotional faces. Typically, judgments of emotional arousal are markedly lower for emotional faces, compared to images depicting emotional scenes (Schupp, Öhman et al., 2004). In addition to differences in experimental design, we propose that the emotional intensity of a preceding event is one crucial variable determining the occurrence of either facilitation or interference. This view is corroborated by the results of Experiment 3, which directly demonstrated that an increase of self-rated arousal is related to an increase of behavioral interference. Interference effects may result from resource-consuming analysis of emotionally arousing, complex visual scenes. In contrast, emotional faces may serve as salient cues for orienting attentional resources, exerting a facilitatory effect on subsequent processing. Future work may address the question how the complexity of the primary task, such as lexical decision versus orientation discrimination (Pourtois et al., 2004) or contrast detection (Phelps et al., 2006), mediate the divergent outcome of facilitation versus interference.

The gradual increase of interference with higher arousal scores is at variance with an alternative account drawing from theories on affective consistency. Faster RTs and larger N1 amplitudes for targets preceded by neutral versus emotional pictures may originate from facilitated processing for picture-target pairs having congruent hedonic valence (here, neutral-neutral versus pleasant-neutral and unpleasant-neutral pairs). This assumption would predict that effects of picture subcategories having the same valence (e.g., sport scences and erotica, see Methods) differ from effects of the neutral category to the same extent (Klauer & Musch, 2003). This, however, was not observed in the present study.

Target N1 and LP as Markers of Lexico-Semantic Analysis

Understanding emotional interference as observed here requires to consider possible functional roles of the target N1 and LP. A common finding in ERP studies on visual language processing pertains to a posterior negativity peaking around 200 ms and differentiating orthographic (words, pseudowords, consonant strings) from non-orthographic (forms, symbols) stimuli (e.g., Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier, 1999). In a similar time segment, Cohen et al. (2000) reported a negativity over left inferior temporal sensors to words and consonant strings. Intracranial recordings from the posterior fusiform gyrus (Nobre, Allison, & McCarthy, 1994) and brain imaging work (e.g., Petersen, Fox, Snyder, & Raichle, 1990) have stimulated the idea that these early negative deflections to orthographic material may index processing within specialized visual word form systems of the inferior temporal lobe (Cohen et al., 2000).

The present results from regional source modeling are consistent with the concept of inferior temporal language functions. Target processing in the N1 time window (204 - 264 ms) was associated with predominant activity at the pIT sources, which we had seeded according to proposed coordinates of the visual word form area (Cohen & Dehaene, 2004). We found pronounced pIT activity for both word and pseudoword targets. This supports the notion that the basal temporal cortex has a relative specificity for processing words and *word-like* stimuli (e.g., pseudowords), compared to random consonant strings or strings of non-alphabetic tokens (Petersen et al., 1990). However, our results neither corroborate the assumption that word-form processing happens *exclusively* in the left pIT nor do they imply that the *single* function of the left posterior fusiform gyrus relates to word-form processing (see also Price & Devlin, 2003). As illustrated in Figure 18, pIT activation was substantial during picture processing. Moreover, activity during N1 target processing was comparable in size for the left pIT and the homologous area in the right hemisphere. Minor activations in the N1 window were also found in parietal and frontal sources.

We provide suggestive evidence that occipitotemporal processing around 200 to 250 ms is not prelexical in nature as proposed by Dehaene and collaborators (Dehaene, Le Clec, Poline, Le Bihan, & Cohen, 2002). Specifically, the N1 amplitude was less negative to words than pseudowords. Lexicality effects in this latency range are a key finding of the Recognition Potential (RP) research (Martin-Loeches, Hinojosa, Gomez-Jarabo, & Rubia, 2001). The RP has been referred to a bilateral negative deflection at parieto-occipital sensors, typically PO7 and PO8, associated with conscious recognition of words or pictures (Rudell, 1992). Neuronal generators of the RP coincide with proposed locations of the visual word form area, namely lingualis and/or fusiform gyrus in basal extrastriate areas (Martin-Loeches et al., 2001). Interestingly, the RP has been linked to activation of left *and* right basal temporal cortices, which is consistent with our source modeling results showing bilateral pIT activation. In contrast to visual word-form activity, the RP has been shown sensitive to the meaningfulness of the stimulus with responses gradually increasing from random letter strings to pseudowords and words (Martin-Loeches et al., 2001). According to this perspective, the N1 as found here might reflect processes related to lexical access, i.e., activation of entries in a hypothetical mental lexicon.

Lexicality effects on the target LP were reflected by a more negative positivity to pseudowords compared to words at centroparietal sensors. This finding closely replicates those found in a forced-choice lexical decision paradigm by Hauk et al. (2006). It is also consistent with the N400 literature reporting increased negativity to pseudowords relative to words (Kutas & Federmeier, 2000). Activity in the LP window mainly originated from left aIT, suggesting that stimulus analysis in the late time segment had progressed in anterior direction. Within the left inferior temporal cortex, a directed processing stream proceeding anteriorly by continuous cortical processing has been demonstrated by Fernandez et al. (2001), who recorded intracranial ERPs during visual word processing. Using a similar methodology, Nobre and co-workers (1994) showed that anterior regions of the fusiform gyrus were sensitive to semantic stimulus dimensions, while posterior parts responded equally to words and non-words. Semantic encoding in anterior temporal areas is also suggested by studies using positron-emission tomography during visual word perception (Vandenberghe, Price, Wise, Josephs, & Frackowiak, 1996). In the LP window, we also observed activity at left parietal and frontal sources, which points to semantic representation encompassing distributed activation in large-scale neuronal networks (Pulvermüller, 1999).

Taken together, the componentry as found here suggests continuous processing along the ventral stream and a corresponding gradual increase in the complexity or depth of analysis. N1 effects may mark the final stages of initial stimulus analysis including lexical access but also mediate the transition from surface/form analysis to content/lexico-semantic processing. LP effects, on the other hand, may index word processing on a post-lexical level involving the formation of semantic associations and processes related to decision making and response planning (Hauk et al., 2006).

Emotional Interference in Network Models of Stimulus Representation

How do the proposed functions of target ERPs relate to the observed effects of emotional arousal? Source modeling showed that pIT, particularly in the left hemisphere, constituted the dominant site of emotional interference effects in the N1 time segment. Paralleling the N1 scalp data, inhibition at pIT sources was present for both words and pseudowords. During the LP window, emotional interference was evident at left aIT and parietal sources. Reduction effects, however, were limited to words. As indicated by RTs during pseudoword trials, interference in the early time segment (between 200 and 250 ms) does not necessarily manifest in overt response modulation. We speculate that a pseudoword decision can be achieved even if pIT activation is reduced to some amount. In contrast, word responses may be impaired because additional interference arises during later time windows (between 400 and 700 ms). Such post-lexical processing after initial word form access in pIT may be elicited automatically for words but not for pseudoword in more anterior regions of the ventral stream.

Another useful approach to explain emotional interference as observed here pertains to network models of stimulus representation (Anderson & Bower, 1973). In this perspective, words differ from pseudowords in the size and cortical topography of the underlying representational network (Pulvermüller, 1999). While pseudowords engage circumscribed cell assemblies representing physical, orthographic, and phonological properties, word representations additionally recruit connections to semantic memory comprising neurons in visual, motor, and multimodal association cortices. Activity in such large-scale semantic networks has recently been linked to centroparietal ERPs showing a positive-going peak around 400 to 500 ms (Hauk et al., 2006). In the same vein, the target LP displayed distributed centroparietal topography and was more pronounced for words, compared to pseudowords. A comparison of source waveforms in this time segment revealed that source strength differences between words and pseudowords were present across several cortical regions. Words showed larger activity than pseudowords in right anterior temporal, right posterior temporal, and bilateral parietal sources. As indexed by shorter RTs to words compared to pseudowords, activation of semantic associations facilitated motor responses. At the same time, it made the cognitive system vulnerable to interference from ongoing affective processing. Such interference may be viewed in terms of neurophysiological suppression, which mediates the rivalry between temporally overlapping activation patterns (Keysers & Perrett, 2002). Supporting the notion of neuronal competition as a basis of emotional interference, perceptual networks representing affectively arousing stimuli have been characterized by strong and widespread connections to memories, visceral, and motor systems (Lang et al., 1997). As described earlier, enhanced late positive potentials to arousing pictures may indicate continued activity within such networks across a time range of several seconds (Cuthbert et al., 2000). The present results suggest that late interference in semantic networks determines the presence of emotional interference at the level of motor output. Inhibited activity in sub-semantic network parts as observed for words *and* pseudowords in the N1 window may not be sufficient to modulate the speed of the manual response.

CONCLUSION

The findings of series B suggest that processing emotionally arousing information interferes with lexico-semantic analysis of temporally adjacent signals. The observed effects of picture emotional intensity and target meaning may help to disentangle the controversy of affective processing leading to either facilitated or impaired performance in a subsequent task. Emotional interference depends on the presence of highly arousing affective material. It increases with the amount of higher-order semantic encoding required for the primary task. Facilitatory effects, such as amplitude enhancement or response acceleration, may be specific to tasks with low perceptual complexity (e.g., orientation discrimination or contrast detection), which are presented subsequent to moderately arousing (face) stimuli. Further work is required to scrutinize the boundary conditions of emotion-induced facilitation versus interference for a given stimulus.

2.3 Emotional Processing during Rapid Serial Visual Presentation (Series C)

The previous lexical decision experiments (series B) corroborate the view that emotional stimuli capture attention and deplete processing resources more than neutral stimuli (e.g., Pratto & John, 1991). As a result, processing of information in spatial (Harris & Pashler, 2004) or temporal proximity (Most et al., 2005) to the affective cue is impaired. The body of research on emotional interference has been continuously growing across the last years, showing impaired processing of concurrent information for a variety of dependent variables including both accuracy (Most et al., 2005) *and* speed (Calvo & Castillo, 2005) measures. Though, several issues remain unclear that are examined in the studies of series C[7]. A first outstanding issue pertains to the temporal unfolding of interference across SOAs beyond 500 ms. In experimental series B, we demonstrated that briefly presented emotional pictures impeded lexical decisions for picture-targets SOAs ranging from 80 ms to 440 ms. In other studies, attention capture by emotional stimuli was observed at SOAs of 300 ms and below (see Fox et al., 2001; Hartikainen, Ogawa, & Knight, 2000; Most et al., 2005; Schimmack & Derryberry, 2005). At longer time intervals of, for instance 1000 ms (Calvo & Castillo, 2005) or 800 ms (Most et al., 2005), the disruptive effects vanished. Sustained emotional interference is found in studies using longer stimulus durations and explicit viewing conditions. For instance, Buodo and colleagues (2002) found a slowing of choice-reactions to acoustic stimuli as late as 4 seconds after onset of an arousing picture (see also Bradley, Cuthbert, & Lang, 1996). Resource occupation across a time range of several seconds is also suggested by work showing reduced P300 amplitudes to acoustic startle during affective picture viewing (Schupp, Cuthbert et al., 2004). However, these studies do not unravel whether a quick glimpse on an emotional stimulus and short distractor duration, respectively, is sufficient to induce sustained interference. Alternatively, impairment by briefly presented emotional information may be a comparatively short-lived phenomenon, not exceeding 400 - 500 ms.

A second unresolved question relates to the possibility of merged effects of behavioral facilitation (Keil & Ihssen, 2004) on the one hand and behavioral interference on the other

---

[7] The present studies are part of the manuscript “The Costs and Benefits of Attending to Emotional Stimuli during Rapid Serial Visual Presentation” by Niklas Ihssen and Andreas Keil, which has been submitted for publication in “Cognition and Emotion”. Data of Experiment 2 were collected by Thekla Witte-Uhl in the context of her diploma thesis.

hand. For instance, would a target stimulus presented subsequent to an arousing event suffer from capture effects when the target itself is emotionally arousing? Further, we might ask whether emotional attention capture is merely reflected in impaired performance on a target stimulus or also translates into improved performance on the attention capturing stimulus. Both issues, the time course of emotional interference and the relationship between emotional facilitation and interference, may be approached by the Rapid Serial Visual Presentation (RSVP) technique.

Rapid Serial Visual Presentation and the Attentional Blink

In RSVP experiments, stimuli are presented sequentially at the same spatial location and at a high rate of speed, for example 10 items per second (see Coltheart, 1999). In a typical experiment, participants search the stimulus stream for pre-specified target items. At RSVP rates between 8 and 12 Hz, attending to a target amidst a distractor stream leads to a transient impairment in detecting or identifying a second target stimulus. In the visual modality, this so-called “Attentional Blink” (AB, Raymond, Shapiro, & Arnell, 1992) has been demonstrated with a variety of stimuli such as symbols, letters, digits, and words (e.g., Raymond, 2003). Report rates of the second target are usually reduced for inter-target-intervals between 200 and 500 ms. Although recent studies (Di Lollo, Kawahara, Shahab Ghorashi, & Enns, 2005; Olivers, van der Stigchel, & Hulleman, 2005) questioned the validity of resource-limitation accounts, the majority of AB theories attribute the accuracy impairment to a decreased availability of memory encoding (e.g., Chun & Potter, 1995; Jolicoeur, 1998) or attentional selection (Nieuwenstein & Potter, 2006) resources due to encoding/selection of the first target (T1). As a consequence, the second target (T2) is not transformed into a durable and reportable working memory representation.

Effects of Emotional Content during RSVP

As delineated in chapter 1.3.1, a growing number of studies have reported enhanced report rates for emotional relative to neutral T2s during the AB period. In their initial study, Anderson and Phelps (2001) found that word T2s denoting aversive content were reported with higher accuracy than neutral T2s at inter-target intervals below 400 ms. Anderson (2005) and Keil and Ihssen (2004) extended these findings by showing that the relevant dimension for T2 facilitation was not the valence but the emotional intensity of the stimulus, i.e., the AB effect was reduced for high-arousing pleasant and unpleasant words, compared to neutral ones. Subsequent studies showed enhanced performance during the AB using facial expressions of emotion (Milders et al., 2006) or arousing IAPS pictures (Trippe et al., 2007). Recently, facilitated identification of emotional T2s was linked to enhancement of sensory processing (Keil et al., 2006). As measured by the electrocortical response at posterior sensors, pleasant and unpleasant T2s elicited rapid amplitude enhancement (120-270 ms after T2 onset).

Compared to emotional facilitation, only few studies documented emotional interference during RSVP. In the work by Most et al. (2005, Experiment 1), participants monitored a 10 Hz picture stream for a single rotated image. A negative or neutral (non-target) picture preceded the target by either two or eight items. At the short (200 ms) but not at the long (800 ms) distractor-target interval, detection was worse for targets following negative distractors, compared to neutral items. Interference during RSVP may also be obtained by emotional word distractors. In a study by Barnard and co-workers (Barnard, Ramponi, Battye, & Mackintosh, 2005), participants had to identify target words belonging to the semantic category “job or profession”. For participants with high state anxiety, performance was impaired when the target word was preceded by a non-target physical threat word. Similar to Most et al. (2005), emotional interference was restricted to a specific time window (lag 4, i.e., 400 ms after onset of the threat word).

Overview of the Studies of Series C

The fleetingness of emotional interference in the outlined RSVP studies raises the question whether the impact of emotional RSVP items may be increased and prolonged, respectively, by increasing their task-relevance. To shed some light on this issue, Experiment 1 utilized an Attentional Blink paradigm where emotional versus neutral words were presented as first targets preceding a second (neutral) target at varying temporal distances. Crucially for the present series of experiments, Shapiro and co-workers (Shapiro, Schmitz, Martens, Hommel, & Schnitzler, 2006) recently found a direct relationship between the magnitude of the T2 deficit and the amount of attentional resources allocated to the T1 task. Thus, the more resources were allocated to T1 (as indexed by increased electrocortical activation), the less likely T2s were reported. The rationale of our experiment was to increase resource allocation to T1 by increasing its motivational significance. Emotional arousing (positive and negative) T1s were predicted to prompt an enhanced attentional response and distract more processing resources away from the T2 task than neutral (non-arousing) T1s. By varying the position of T2 relative to T1, affective interference could be tracked at different time ranges. Targetness was defined by font color, making T1s easily discriminable from the

RSVP stream. Thus, emotional interference was probed under conditions of maximized task-relevance and full explicit attention to the emotional stimulus. In contrast to the studies by Most et al. (2005) and Barnard et al. (2005), this setup also allowed for simultaneously measuring potential facilitatory effects (as reflected by enhanced T1 report rates for emotional stimuli) and interference effects (as reflected by reduced T2 report rates for T2s subsequent to emotional T1s) of emotional content.

In Experiment 2, we applied a reversed strategy. Emotional versus neutral items were presented as RSVP non-target fillers. In contrast to previous studies (see Most et al., 2005, and Barnard et al., 2005, Experiment 1), we did not draw emotional distractors and targets from the same stimulus class. Participants in the Most et al. study searched for pictures preceded by picture distractors, in the study by Barnard et al. they searched for words preceded by word distractors. Hence, we cannot safely infer that the emotional fillers did not undergo explicit processing. In Experiment 2, we used the same paradigm as in Experiment 1 but replaced the emotional/neutral T1s with emotional/ neutral picture stimuli. Choosing picture instead of word distractors minimized the perceptual correspondence to the pre-specified target template (a green word). Therefore, as opposed to Most et al. and Barnard et al., participants had no *need* to process the meaning of the emotional distractor. Differences in the amount of interference elicited by emotional vs. neutral pictures thus would index the amount of automatic and stimulus-driven interference.

The goal of Experiment 3 was to investigate combined effects of emotional facilitation and inhibition within one experimental setup. This was done by varying emotional content of T1 *and* T2 as independent factors in a fully crossed design. Given the finding by Shapiro et al. (2006) that dual-target identification during RSVP is subject to inter- and intra-individually variable resource sharing strategies, variations in T2 emotional content may modify interference effects of T1 content. Finally, Experiment 4 set out to unravel the effects of affective congruence between T1 and T2.

EXPERIMENT 1

In an Attentional Blink paradigm with word stimuli, T1s denoting high-arousing positive, low-arousing neutral, or high-arousing negative content were presented at varying SOAs to a trailing second target. T2s and fillers were neutral in content. Reflecting an Attentional Blink, T2 accuracy was predicted to increase with increasing T1-T2 SOA. Emotional content was hypothesized to facilitate T1 identification. On the other hand, emotional T1s were assumed to deplete more processing resources than neutral T1s. As a result, accuracy for T2s subsequent to neutral T1s should exceed accuracy for T2s following emotional T1s.

Methods

Participants

16 university students consented to participate. 11 participants were female. The mean age was 23.6 years (SD = 4.3). All participants were native speakers of German and had normal or corrected-to-normal vision. Participants received course credits or a financial bonus of 7.50 Euros.

Stimuli

Stimuli in all four experiments were German verbs. Selection of stimuli was based on aggregation of 10 affective rating studies. Using a paper-and-pencil or a computer-based version of SAM (Bradley & Lang, 1994), the ratings were conducted with 215 participants from the University of Konstanz, either in single rating sessions or subsequent to other experiments. Participants rated 370 verbs on the dimensions emotional valence and arousal. After selection, T1s in Experiment 1 comprised 30 high-arousing pleasant, 30 low-arousing neutral, and 30 high-arousing unpleasant verbs. Mean valence and arousal scores according to the pre-study ratings are shown in Table 4.

Table 4. Affective and linguistic parameters of the stimuli used in Experiment 1. Values represent means and standard deviations (standard error for word frequency) across all verbs per category (arousal: 1 = low-arousing, 9 = high-arousing; valence: 1 = unpleasant, 9 = pleasant).

Stimulus Category	Arousal	Valence	Number of letters	Number of syllables	Word frequency
T1 pleasant	6.9 (0.8)	7.9 (0.6)	8.2 (1.6)	2.7 (0.7)	163.9 (60.6)
T1 neutral	3.0 (0.5)	5.1 (0.4)	8.1 (1.8)	2.5 (0.6)	150.2 (55.7)
T1 unpleasant	7.7 (0.5)	1.6 (0.4)	8.3 (2.2)	2.6 (0.8)	143.0 (47.5)
T2 (neutral)	4.2 (0.7)	5.5 (0.6)	7.7 (2.1)	2.5 (0.7)	241.6 (47.7)
Fillers	4.3 (0.7)	5.3 (0.5)	8.2 (2.0)	2.8 (0.7)	325.9 (88.2)

Note. The T1 and T2 set each comprised 30 verbs, fillers comprised 50 verbs.

To obtain stimulus samples sufficient in size and homogenous in terms of spoken/written word frequency, word length, and valence/arousal, we did not use so-called taboo words (curses, colloquials for sexual behavior etc.) as employed for example by MacKay et al. (2004) and Anderson (2005). Maximizing the rated emotional arousal of pleasant and unpleasant target words, we selected verbs denoting emotional reactions (e.g., to cheer, to execrate) or actions being typically judged as either positive or negative (to win, to abuse, to torture), in addition to neutral words. The three T1 categories were matched for phonological word length (mean number of syllables), which is assumed to be a relevant variable in AB tasks (Olson, Chun, & Anderson, 2001), for mean number of letters, and for mean word (lemma) frequency using the CELEX data base (Baayen et al., 1995). ANOVAs yielded no differences between the three T1 categories on these dimensions, all Fs < 0.5. T2 items (30 verbs) and fillers (50 verbs) were neutral in content.

Procedure

After being screened for clinical problems and photic epilepsy, participants read the written instructions and signed the informed consent form. Participants were then seated comfortably in a sound attenuated, dimly lit room. Stimuli were presented on a computer monitor located at a distance of 70 cm in front of the participants, who kept their chin on a rest. Monitor retrace frequency was set to 60 Hz, which resulted in a frame duration of 16.7 ms. RSVP stimulation was generated by displaying a stream of words in which each item was shown for 4 refresh cycles (66.7 ms) and followed by a blank screen of 3 refresh cycles duration (50 ms). This resulted in a RSVP frequency of 8.6 Hz. All words were centrally shown in Helvetica 26 Font subtending a vertical angle of 0.6°. Words were displayed in upper-case characters on a black background. Participants were instructed to monitor a stream of white words for two green target items and to report the identity of the targets after each trial.

An initial blank screen displayed for 1000 ms at the beginning of each trial preceded the following sequence of events : a) A pre-target baseline with a number of filler words randomly varying between 5 and 25, shown in white color, b) the first target (either a pleasant, neutral, or unpleasant word), shown in green color, c) 1 (SOA 233 ms, lag 2), 3 (SOA 467 ms, lag 4), or 5 (SOA 700 ms, lag 6) intervening white fillers[8], d) the second target,

---

[8] Selection of T1-T2 SOAs was based on the results of a pilot study with 8 participants (4 women, mean age: 25.6, SD = 4.1 years). Using the same RSVP procedure as in Experiments 1 - 4, neutral T2s were presented at positions 1 to 8 after neutral T1s. T2 identification was most strongly impaired at lag 2 (percentage of correct responses: 30.8, SE = 3.8) and report rates reached asymptotic recovery at lag 6 (82.8, SE = 4.2).

shown in green color, e) depending on the SOA condition, 13 (lag 6), 15 (lag 4) or 17 (lag 2) white fillers. The schematic of one RSVP trial is shown in Figure 21.

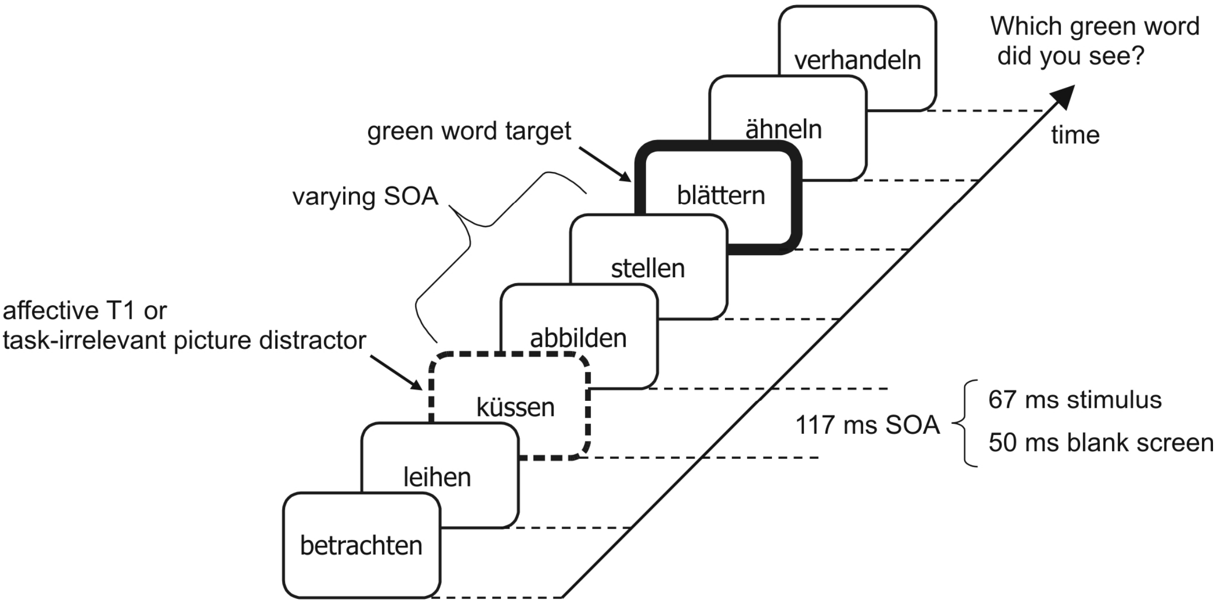

Figure 21. Schematic of one RSVP trial. Participants' task was to identify a target word shown in green color and embedded in a stream of (white) neutral filler words. At varying SOAs (233 ms, 467 ms, and 700 ms), the target was preceded by another target (T1, Experiments 1, 3, and 4) or task-irrelevant picture (Experiment 2) that varied in emotional content. betrachten = to look at, leihen = to borrow, küssen = to kiss, abbilden = to map, stellen = to place, blättern = to browse, ähneln = to resemble, verhandeln = to negotiate.

Filler words were not repeated within one trial. Each word within the three T1 category sets (pleasant, neutral, and unpleasant) appeared once in each lag condition and hence was repeated three times in the experiment. There were no repetitions of any T1-T2 pairs. The combination of T1 category and T1-T2 SOA resulted in a two-factorial design with 3 (T1 categories) x 3 (T1-T2 SOAs) conditions and 30 words x 9 conditions = 270 trials. The experiment was divided by a break into two equally sized blocks of 135 trials. The sequence of conditions was randomized with the constraint of neither the same condition nor the same word appearing in consecutive trials.

Responses were recorded after each trial. Participants were prompted by instructions on the monitor to type the first letter of each green word on a keyboard. Additionally, they were asked to report aloud the target words using a microphone located in the experimental chamber. Participants then initiated the next trial with a button press. Participants performed 4 practice trials prior to the experimental session. Different targets were used in these practice trials than in the experimental trials. At the end of the experimental session, ratings of emotional arousal and valence were obtained for each target item.

Data Analysis

For each participant and target category, we calculated mean ratings of emotional valence and arousal. Mean ratings were submitted to repeated-measurement Analysis of Variance (ANOVA) having the within-subject factor CATEGORY (4; T1 pleasant, T1 neutral, T1 unpleasant, T2 [neutral]). With regard to the RSVP task, two dependent measures were calculated for each participant: T1 report was computed as the percentage of correct T1 responses in each condition. T2 accuracy was assessed as the percentage of correct T2 responses for those trials in which T1 had been identified as correct (T2 | T1). The two dependent measures were evaluated by means of a two-factorial repeated measurement ANOVA using T1-T2 SOA (3; 233 ms, 467 ms, 700 ms) and T1 CATEGORY (3; pleasant, neutral, unpleasant) as within-subject factors. Significant effects were followed by Newman-Keuls tests.

Results

Subjective Ratings

ANOVA showed a significant main effect of CATEGORY both for the arousal and the valence dimension, arousal: $F(3, 45) = 125.3$, $p < 0.001$, valence: $F(3, 45) = 575.1$, $p < 0.001$. Unpleasant T1s (mean arousal rating: 7.6, $SD = 1.0$) were rated as more arousing than pleasant T1s (6.5, $SD = 1.1$), Newman-Keuls $p < 0.01$, neutral T1s (2.6, $SD = 1.4$), $p < 0.001$, and neutral T2s (3.3, $SD = 1.4$), $p < 0.001$. Pleasant T1s obtained higher arousal scores than the neutral categories, ps < 0.001. Neutral T2s had slightly higher arousal ratings than neutral T1s, $p = 0.03$. Regarding valence, subjectively rated pleasure was higher for pleasant T1s (mean valence rating: 7.9, $SD = 0.7$), compared to neutral T2s (5.4, $SD = 0.4$) and neutral T1s (5.1, $SD = 0.3$), ps < 0.001. Unpleasant T1s (1.5, $SD = 0.3$) obtained lower valence values than the neutral categories, ps < 0.001. Valence was slightly higher for neutral T2s compared to neutral T1s, $p = 0.04$.

T1 Identification

T1 report rates were high, resulting in a mean of 85.6 percent correct responses ($SE = 0.9$) across T1-T2 SOAs and T1 affective categories (see Figure 22, left panel). T1 accuracy did not vary with T1-T2 SOA, $F(2, 30) = 0.7$. ANOVA showed a main effect of T1

CATEGORY, F(2, 30) = 3.4, $p < 0.05$, reflecting facilitated identification for emotional T1s. Accordingly, post-hoc tests yielded slightly higher report rates for pleasant, p = 0.047, and unpleasant, p = 0.059, compared to neutral T1s. Accuracy for pleasant and unpleasant T1s did not differ from each other, p = 0.75. The effect of T1 CATEGORY did not interact with T1-T2 SOA, F(4, 60) = 0.4.

T2 Identification

Averaged across conditions, T2s were reported as correct in 49.0 percent of the trials (SE = 2.5). We observed a significant main effect of T1-T2 SOA, F(2, 30) = 131.6, $p < 0.01$. As illustrated in Figure 22 (right panel), T2 performance improved with increasing distance to T1. Accuracy was higher at lag 4 compared to lag 2 and at lag 6 compared to lag 4, ps < 0.001. A significant main effect of T1 CATEGORY, F(2, 30) = 7.4, $p < 0.01$, indicated modulation of T2 report by T1 affective category. Compared to T2s following neutral T1s, T2s were worse identified when preceded by pleasant, p = 0.02, or unpleasant T1s, $p < 0.01$. Pleasant and unpleasant T1s did not evoke different response patterns, p = 0.18. Further, there was no interaction between T1-T2 SOA and T1 CATEGORY F(4, 60) = 1.3.

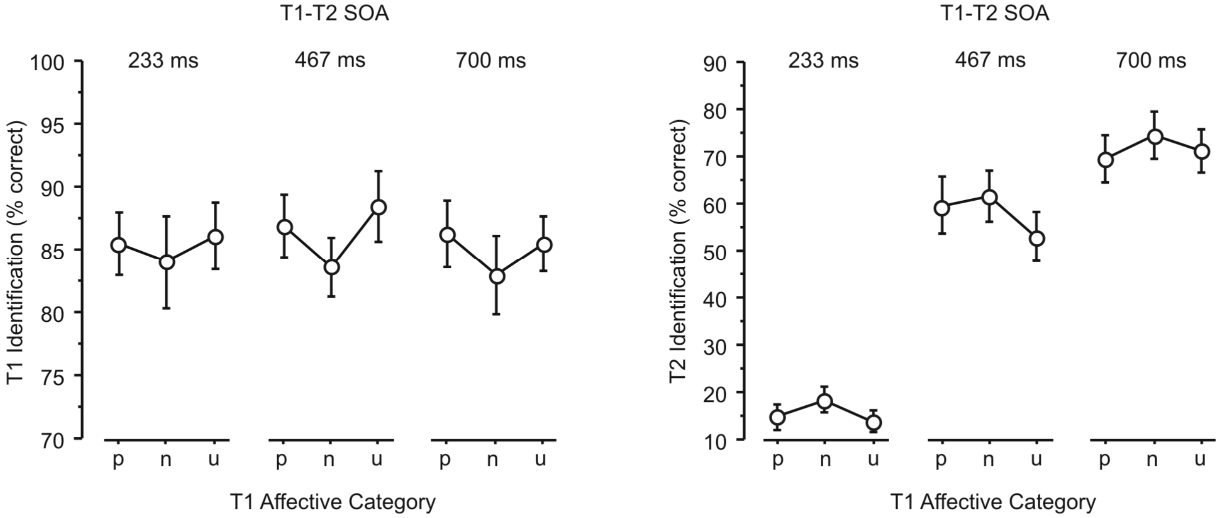

Figure 22. Mean identification accuracy (percent correct) for first targets (T1s, left panel) and second targets (T2s, right panel) as a function of T1-T2 SOA and T1 affective category in Experiment 1. Whereas fillers and T2s were neutral in content, affective category of T1 varied between high-arousing pleasant (p), low-arousing neutral (n), and high-arousing unpleasant (u). T2 followed T1 after one (SOA 233 ms), three (SOA 467 ms), or five (SOA 700 ms) intervening filler words. Error bars indicate standard errors.

Discussion

In Experiment 1, participants' task was to search for two green-colored targets amidst a stream of white filler words and to report the targets after each trial. Varying emotional content of the first target affected identification accuracy in two ways. Irrespective of valence, arousing content facilitated report of the first target. Accordingly, identification accuracy was higher for pleasant and unpleasant compared to neutral T1s. This finding extends previous work observing enhanced accuracy for arousing T2s compared to neutral T2s (Anderson, 2005; Keil et al., 2006). These authors also demonstrated that facilitation effects for emotional stimuli during RSVP do not originate from differences in semantic distinctiveness, orthographic and contextual factors, or the unexpectedness of the arousing material.

As a second main result of Experiment 1, emotional content of T1 impaired encoding of the second target. Reflecting an Attentional Blink, T2 identification was mainly determined by the temporal distance to T1, i.e., performance increased markedly with increasing T1-T2 SOA. At all T1-T2 SOAs (233 ms, 467 ms, 700 ms), however, T2s following pleasant or unpleasant T1s were less accurately reported then T2s subsequent to neutral T1s. The lack of an interaction T1-T2 SOA x affective T1 category suggests that motivated attention to emotional arousing information and instruction-based volitional attention to T1 affected T2 performance in an additive manner. This is consistent with the additive effects of emotional and spatial selective attention seen on a neuronal level (Keil et al., 2005). The finding that emotional impairment in Experiment 1 did not recover at longer T1-T2 SOAs complements the results of series B, where briefly presented arousing pictures impaired lexical decision responses at SOAs between 80 and 440 ms. It also extends several other studies using brief distractor durations that showed relatively short-lived emotional interference, for instance in tasks that required two-choice motor responses to target words following negative words (Calvo & Castillo, 2005) or the detection of rotated pictures following negative pictures (Most et al., 2005). In both studies, performance decrements were restricted to short distractor-target intervals of 200 or 400 ms. At longer SOAs, accuracy was not different for emotional versus neutral distractors. Two mechanisms could have contributed to the divergence of results. First, detrimental effects of affective content may persist longer when aspects of working memory consolidation and semantic encoding are tapped in the interfered task. Consistent with this idea, experimental series B indicated a relation between the amount of emotional interference and the amount of lexico-semantic processing, i.e., processing a word versus pseudoword, required for the target.

A second factor determining the dynamics of emotional interference could pertain to the task-relevance of the emotional stimulus. We found that a target RSVP item that is voluntarily selected out of the stimulus stream captures additional attentional resources. Data from a variety of paradigms imply that interference does not require explicit selection of these stimuli. Rather, emotionally charged signals grab processing resources in an automatic, implicit manner (Pratto & John, 1991). Effects of emotional content and task-relevance during RSVP have also been addressed in the study by Barnard et al. (2005, Experiment 2). There, participants had to identify semantically defined targets preceded by either threat-related or neutral words. Distractors also varied in semantic similarity to the searched target category. Interestingly, for distractors with high similarity to the targets, i.e., high task-relevance, non-anxious participants exhibited a delay of interference effects being maximal around 500 ms. In contrast, interference for distractors with low task-relevance was most pronounced at a shorter lag (300 ms). This suggests that task-relevance may determine the point of time where attention is bound to the emotional stimulus. Using the same (T2) identification task as in the previous study, Experiment 2 examined the time course of interference elicited by task-irrelevant emotional information.

EXPERIMENT 2

The current study investigated automatic attention grabbing by task-irrelevant emotional items during RSVP. We used the same paradigm as in Experiment 1 except for emotional/neutral word T1s that were replaced by emotional/neutral non-target pictures. By presenting pictures instead of word distractors, we aimed to reduce the perceptual overlap with the target template (a green word). Not corresponding to participants' attentional set, picture distractors were assumed to elicit less explicit processing, compared to the emotional distractors as used in previous studies (Barnard et al., 2005; Most et al., 2005). Building on the findings of experimental series B, we predicted that high-arousing pleasant and unpleasant pictures trigger an automatic attentional response resulting in impaired target identification.

Methods

Participants

22 university students consented to participate in Experiment 3 (14 female, mean age: 23.7, SD = 3.0). All participants were native speakers of German and had normal or corrected-to-normal vision. They received course credits or a financial bonus of 7.50 Euros.

Stimuli

Stimuli were identical to Experiment 1 with the exception that the verbal T1 stimuli were replaced by colored pictures (30 high-arousing pleasant, 30 low-arousing neutral, and 30 high-arousing unpleasant pictures) selected from the IAPS (Lang et al., 2005). According to the normative IAPS ratings, pleasant pictures (erotic couples, pets, sport scenes; mean arousal: 5.5) and unpleasant pictures (mutilated bodies and attack scenes; mean arousal: 6.2) were higher in emotional intensity, compared to neutral items (neutral persons and objects; mean arousal: 3.4). On the valence dimension, normative ratings were higher for pleasant (mean valence: 7.0) compared to neutral (4.8) and unpleasant material (2.5).

Procedure

The procedure was the same as in Experiment 1 with the exception that there was no T1 task. Instead of a first attended target, T2 now was preceded by a task-irrelevant picture distractor (pleasant, neutral, or unpleasant, see above) that was presented at the same three T1-T2 SOAs as in Experiment 1. Pictures subtended a visual angle of 4.3° vertically and 5.8° horizontally. Participants were instructed to monitor the RSVP on a single green word and to report the identity of the target after each trial. After the RSVP experiment, each picture distractor was presented on the computer monitor, and participants rated its emotional arousal and valence.

Data Analysis

Percentages of correct responses to the word targets were evaluated with ANOVA having the within-subject factors T1-T2 SOA (3; 233 ms, 467 ms, 700 ms) and DISTRACTOR CATEGORY (3; pleasant, neutral, unpleasant). Newman-Keuls tests were used to follow significant effects. Mean SAM ratings were analyzed as described in Experiment 1.

Results

Subjective Ratings

ANOVA showed a main effect of CATEGORY both for the arousal, $F(2, 42) = 112.6$ $p < 0.001$, and the valence dimension, $F(2, 42) = 230.4$, $p < 0.001$. On the arousal dimension, SAM ratings were higher for unpleasant (6.5, $SD = 1.2$) compared to pleasant (mean arousal: 3.9, $SD = 1.5$), and for pleasant compared to neutral pictures (2.5, $SD = 1.1$), $ps < 0.001$.

Pleasure scores decreased from pleasant (mean valence: 6.7, SD = 0.9) to neutral (5.5, SD = 0.8), and from neutral to unpleasant items (2.3, SD = 0.5), ps < 0.001.

Target Identification

Collapsed across conditions, targets were identified as correct in 89.9 percent of the trials (SE = 0.7). As reflected by a significant main effect of CATEGORY, F(2, 42) = 5.7, p < 0.01, identification accuracy varied with affective content of the preceding picture distractor. Figure 23 illustrates that, in comparison to preceding neutral pictures, report rates were reduced for pleasant , p < 0.01, and unpleasant distractors , p = 0.016. In contrast, the difference between pleasant and unpleasant pictures was not significant, p = 0.47. ANOVA showed neither a main effect of T1-T2 SOA, F(2, 42) = 0.1, nor an interaction T1-T2 SOA x CATEGORY, F(4, 84) = 1.7.

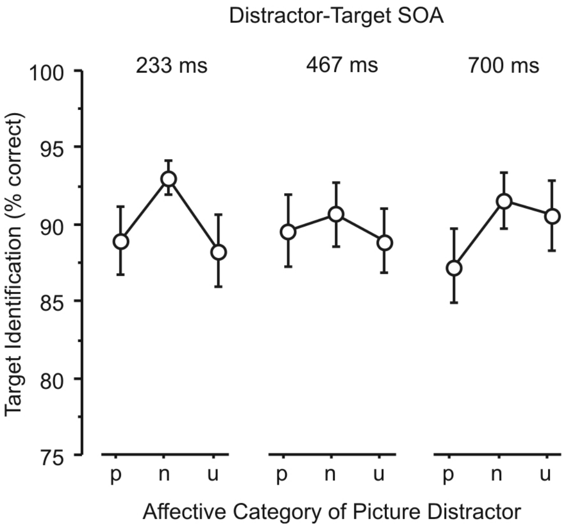

Figure 23. Mean target identification accuracy in Experiment 2. Participants searched for a single green word target that was preceded by a task-irrelevant picture (high-arousing pleasant (p), low-arousing neutral (n), or high arousing unpleasant(u)) at three distractor-target SOAs (233 ms, 467 ms, 700 ms). Filler words and targets were neutral in content. Error bars indicate standard errors.

Discussion

High identification rates in the present study indicate that participants did well at selecting a single target word from the RSVP stream. The overall high performance also suggests that, although the picture distractors were perceptually salient amidst word stimuli, in most trials participants had no trouble ignoring them. Yet, failures to report the target were more frequent when the word was preceded by emotional compared to neutral pictures, i.e., emotional content interfered with discerning the identity of the target word. Experiment 2 thus verified that interference effects are not contingent on prior capture by perceptual target cues such as font color. Previous work demonstrated such non-target emotional interference exerted by word distractors inside a word stream (Barnard et al., 2005) or picture distractors inside a picture stream (Most et al., 2005). We add to these findings that automatic emotional interference occurs for items that are clearly distinct from the searched target and do not evoke any voluntary attempts of semantic processing.

Again, we found no interaction between distractor-target lag and distractor category, i.e., interference effects were sustained also for non-target emotional stimuli. This suggests that the fleetingness of attention capture found in the study by Most et al. (2005) did not result from the task-irrelevance of emotional RSVP items. Rather, we propose that the nature of the interfered task or the processes required for accomplishing the task, respectively, determines the dynamics of performance deficits. Whereas interference between affective processing and low-level analysis, such as orientation detection (Most et al., 2005), may be restricted to short distractor-target distances, detrimental effects of emotion on subsequent lexico-semantic analysis may persist longer.

In the first two experiments, we showed attention capture for stimuli of both valence categories (positive and negative). Certainly, disentangling the role of valence versus arousal would require a fully crossed design where, for instance, high-arousing negative distractors are contrasted with low-arousing negative distractors. Yet, building upon the findings of the experiments in series B, impaired processing during RSVP most likely varies as a function of stimulus' emotional intensity and not hedonic valence.

The question arises whether emotional RSVP items produce interference with a subsequent target task that involves affective processing, too. Previous work showed that emotional T2s are encoded with higher accuracy than neutral T2s (e.g., Keil & Ihssen, 2004). The above results demonstrated that emotional T1s inhibit encoding of neutral T2s. By fully crossing the factors of T1 and T2 emotional content, Experiment 3 examined the possibility of combined effects of T1 and T2 affective category within one RSVP stream.

EXPERIMENT 3

In Experiment 3, emotional content of T1 and emotional content of T2 were varied as independent factors. To keep the duration of the experiment within acceptable limits, only low-arousing neutral and high-arousing unpleasant targets were used.

Methods

Participants

Informed consent for participation was obtained from 19 healthy university students. Data from one subject was excluded due to equipment problems. The remaining sample (n = 18) consisted of 13 female participants, having a mean age of 25.2 years (SD = 6.4). SAM ratings of two additional participants could not be used because of non-compliance with the rating procedure. Thus, the remaining sample for the rating data (n = 16) included 11 women, the mean age was 25.2 years (SD = 6.8). None of the respondents had been participating in other studies involving RSVP with emotional material. All subjects were native speakers of German and had normal or corrected-to-normal vision. Participants received course credits or a financial bonus of 7.50 Euros.

Table 5. Affective and linguistic parameters of the stimuli used in Experiments 3 and 4. Values represent means and standard deviations (standard error for word frequency) across all verbs per category (arousal: 1 = low-arousing, 9 = high-arousing; valence: 1 = unpleasant, 9 = pleasant).

Stimulus category	Arousal	Valence	Number of letters	Word frequency
Neutral targets (30 verbs)	3.3 (0.5)	5.2 (0.4)	8.1 (1.4)	73.0 (15.1)
Unpleasant targets (30 verbs)	7.5 (0.7)	1.5 (0.2)	8.0 (2.0)	70.5 (21.8)
Pleasant targets (30 verbs)	6.6 (0.9)	7.8 (0.6)	7.5 (1.3)	72.9 (17.7)
Fillers (60 verbs)	4.3 (0.8)	5.3 (0.5)	9.1 (2.3)	858.7 (256.0)

Note. Unpleasant targets were only used in Experiment 3, pleasant targets only in Experiment 4.

Stimuli

Based on the pre-experimental ratings (see methods section Experiment 1), three sets of German verbs were selected: low-arousing neutral fillers (60 verbs), low-arousing neutral targets (30 verbs), and high-arousing unpleasant targets (30 verbs). Neutral and unpleasant targets did not differ in word frequency as assessed with the lemma frequency of the CELEX data base (Baayen et al., 1995), $t(29) = .087$, $p > 0.9$. With regard to phonological word length, both target categories comprised 20 2-syllables and 10 3-syllables words. There were no differences in the mean number of letters, t(29) < 0.5. Word length of fillers was kept larger to guarantee sufficient filler-target masking, which is considered to be a key condition of the AB (Giesbrecht & Di Lollo, 1998). Affective and linguistic parameters of the stimulus classes used in Experiment 3 and 4 are shown in Table 5.

Procedure

The procedure was similar to Experiment 1. Each RSVP trial began with a blank screen (500 ms) followed by 5 - 25 fillers and T1. T1 and T2 could be either a neutral low-arousing, or an unpleasant high-arousing item. T2 appeared at lag 2 (T1-T2 SOA 233 ms), or lag 6 (T1-T2 SOA 700 ms) after T1, and 10 fillers followed T2 offset. All words were shown centrally in Times New Roman 26 Font and displayed in lower-case characters. The combination of T1 category, T2 category, and T1-T2 SOA resulted in a three-factorial design with 2 (T1 category: pleasant and unpleasant) x 2 (T2 category: pleasant and unpleasant) x 2 (T1-T2 SOA: 233 and 700 ms) = 8 within-subject conditions. Each word of the two target sets (30 verbs) appeared once in each condition resulting in 8 repetitions for each target word and 8 conditions x 30 words = 240 trials for the whole experiment. The experiment was divided by a break into two equally sized blocks of 120 trials. The sequence of conditions was constructed by introducing smaller blocks of 8 trials. In each block, each of the 8 conditions appeared once at a randomized position, and at the block transitions the same condition never appeared in consecutive trials. In trials with affectively congruent target categories (T1 neutral - T2 neutral, T1 unpleasant - T2 unpleasant), T1 and T2 were never the same words. Furthermore, specific targets were not repeated within one of the 8-conditions blocks (see above) nor did specific T1-T2 pairs occur more than once across the whole session. As in the previous experiments, the stream of fillers contained no repetitions of words.

Responses were registered after each trial. Participants were prompted via an entry form on the monitor to type each identified green word on a keyboard. They ran 5 practice trials prior to the experimental session. Only neutral stimuli and items that were not included in the experimental set served as targets in the practice trials. Upon completing the AB trials, ratings of the target words were obtained as described in Experiment 1. In addition to the arousal and valence ratings, concreteness of the words was assessed using a modification of the SAM scale with the poles concrete - abstract. Target concreteness was rated in order to control for possible confoundation with the affective dimensions.

Data Analysis

Report rates were calculated in the same way as in Experiment 1. T1 and T2 report rates were submitted separately to a three-factorial repeated measurement ANOVA using T1-T2 SOA (2, 233 ms vs. 700 ms), T1 CATEGORY (2, T1 neutral vs. T1 unpleasant), and T2 CATEGORY (2, T2 neutral vs. T2 unpleasant) as within-subject factors. In terms of SAM ratings, differences between pleasant and unpleasant targets were tested with t-tests.

Results

Subjective Ratings

Arousal scores for neutral targets (mean: 2.99, SD = 1.34) were lower than for unpleasant targets (7.43, SD = 0.93), t(15) = 9.8, $p < 0.001$. Neutral targets (5.42, SD = 0.76) obtained higher valence scores than unpleasant targets (1.71, SD = 0.64), t(15) = 13.5, $p < 0.001$. In terms of concreteness, ratings for neutral targets (6.34, SD = 1.34) did not differ from unpleasant targets (6.74, SD = 1.57), t(15) < 1.0. Thus, it is unlikely that concreteness differences between the affective categories can account for possible differences in the accuracy data.

T1 Identification

Across conditions, mean T1 identification accuracy was 82.2 percent correct (SE = 1.2). ANOVA showed no main effects of T1-T2 SOA or T2 CATEGORY, Fs < 0.5. However, there was a significant main effect of T1 CATEGORY, F (1, 17) = 12.6, $p < 0.01$. As illustrated in Figure 24 (top panel), unpleasant T1s were reported with higher accuracy than neutral T1s. There were no significant two-way or three-way interactions between the factors, Fs < 2.5.

T2 Identification

T2s were successfully reported with a mean of 49.2 percent (SE = 2.7). ANOVA on T2 accuracy yielded a significant main effect of T1-T2 SOA, F(1, 17) = 143.5, $p < 0.001$, reflecting reduced accuracy at 233 ms SOA compared to 700 ms SOA. A strong main effect of T2 CATEGORY, F(1, 17) = 54.0, $p < 0.001$, indicated facilitated identification of unpleasant T2s compared to neutral T2s. The main effect of T1 CATEGORY, F(1, 17) = 0.7, was not significant. Further, ANOVA showed a weak interaction T1-T2 SOA x T2 CATEGORY, F (1, 17) = 4.0, $p = 0.062$. As illustrated in Figure 24 (bottom panel), the benefit for unpleasant compared to neutral T2s was stronger at the short compared to the long T1-T2 SOA. All other interactions did not reach significance, all Fs < 2.6.

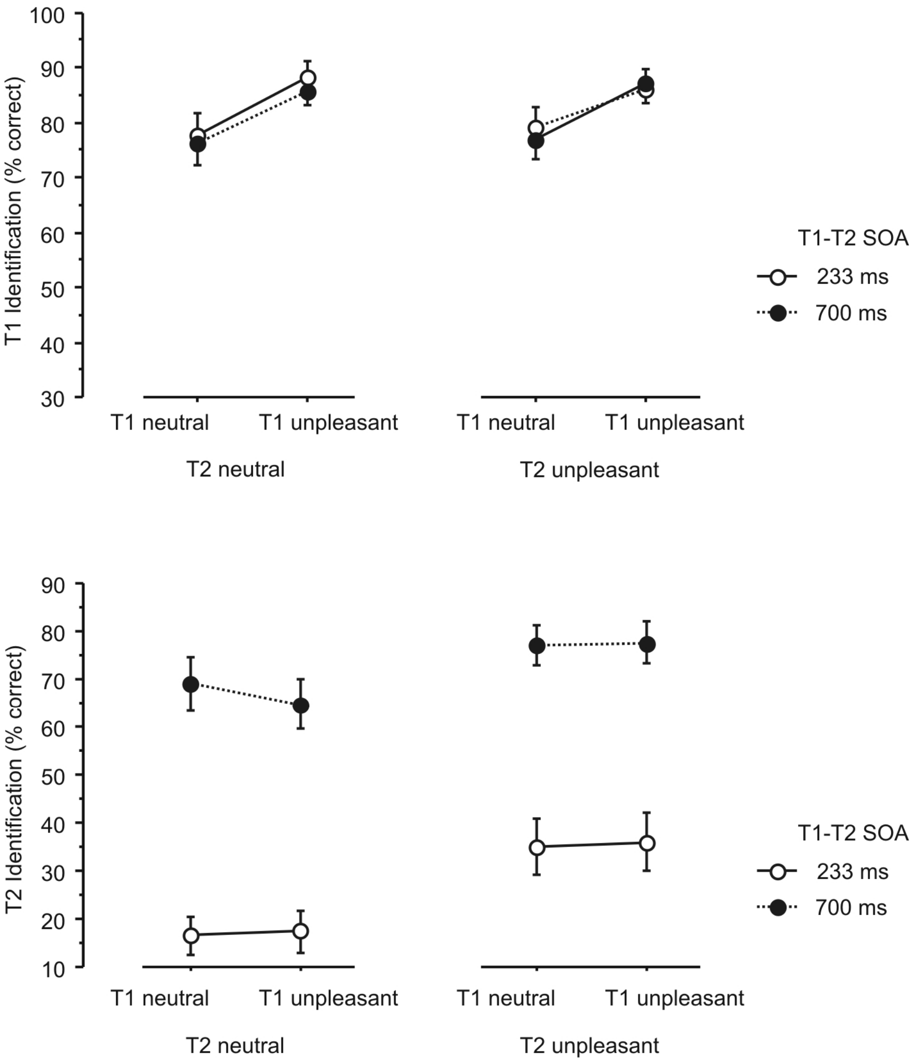

Figure 24. Mean identification accuracy (percent correct) for first targets (T1s, top panel) and second target (T2s, bottom panel) as a function of T1-T2 SOA, T1 affective category, and T2 affective category in Experiment 3. Error bars indicate standard errors.

Discussion

In Experiment 3, we measured report rates for two RSVP targets that could both vary in emotional content (low-arousing neutral vs. high-arousing unpleasant). As a replication of Experiment 1, unpleasant T1s were identified with higher accuracy compared to neutral T1s. T1 facilitation was not affected by the position or emotional category of T2. Similar to T1 report, T2 report was enhanced for unpleasant compared to neutral T2s (see Keil & Ihssen, 2004). Though present at both T1-T2-lags, the benefit for arousing T2s was more pronounced at the short lag (233 ms). Importantly, in the present study emotional T1s did *not* reduce T2 report rates. There was neither a main effect of T1 category nor an interaction T1 category x T2 category. For both T2 types, T2s following unpleasant versus neutral T1s were reported with equal accuracy. What may underlie the lack of T1 inhibition in the current experiment? We propose that variations of emotional content in the T2 task weakened the inhibitory force of emotional T1s. According to this view, participants' knowledge that the second target in some trials denoted motivational relevant content resulted in strategic retention of resources for the T2 task. Withholding a fixed amount of processing capacity for T2 then overrode the effects of emotional T1 enhancement. This interpretation is consistent with the notion of Shapiro and colleagues (2006) that dual-target identification during RSVP does not reflect an immutable serial bottleneck but is a result of individual resource sharing strategies. For example, T2 report rates grow when participants are informed about the temporal distance between the two targets in the next trial (Martens & Johnson, 2005). Taken together, the current findings highlight the sensitivity of RSVP tasks towards processing strategies or context effects arising from differences in stimulus set or experimental design.

Apart from resource interference and perceptual enhancement, another potential factor influencing report of two emotional RSVP targets pertains to T1-T2 affective congruency. Effects of emotional congruency between two stimuli have been extensively investigated in affective priming studies (for a review see Klauer & Musch, 2003). A common finding is facilitated processing of a probe stimulus when it is preceded by a prime stimulus having the same emotional valence. Affective priming in the present experiments would predict higher identification rates for pleasant-pleasant and unpleasant-unpleasant T1-T2 pairs, compared to target configurations involving a mixture of valence categories. In the AB paradigm, improved report has been found for T2s following semantically related T1s (Potter et al., 2005) or semantically related post-T1 distractors (Maki, Frigen, & Paulson, 1997). Experiment 4 was carried out to examine whether valence congruency affects dual-target identification during RSVP.

EXPERIMENT 4

Experiment 4 aimed at quantifying effects of T1-T2 affective congruency. By introducing targets that differed strongly in their valence (pleasant and unpleasant) but only marginally in their arousal ratings, valence congruency was manipulated while arousal effects (perceptual enhancement and resource interference) were held constant, i.e., were expected to occur in each congruency condition with the same magnitude.

Methods

Participants

15 university students (11 women; mean age 24.0, SD = 7.5 years) gave informed consent to participate in this study. Respondents had not been participating in the previous experiment. All participants had normal or corrected-to-normal visual acuity and were native speakers of German. For participation, they received course credits or 7.50 Euros.

Stimuli

Three sets of verbs were used in Experiment 4: 60 low-arousing neutral fillers, 30 high-arousing pleasant targets, and 30 high-arousing unpleasant targets. Fillers and unpleasant targets were the same as in Experiment 3. Pleasant and unpleasant targets showed no difference in word frequency as assessed with the lemma frequency of the CELEX data base (Baayen et al., 1995), $t(29) < 0.5$. In terms of phonological word length, both target categories comprised 20 2-syllables and 10 3-syllables words. There were no differences in the mean number of letters, $t(29) < 1.0$. Affective and linguistic parameters of the stimulus classes used in Experiment 4 are shown in Table 5 (p. 98).

Procedure

The procedure was identical to Experiment 3 with the exception that the neutral target category was replaced by the pleasant target category. Hence, T1 and T2 could be either a pleasant or an unpleasant verb.

Data Analysis

Similar to Experiment 3, T1 and T2 accuracy were submitted to repeated measurement ANOVA using T1-T2 SOA (2, 233 ms vs. 700 ms), T1 CATEGORY (2, T1 pleasant vs. T1 unpleasant), and T2 CATEGORY (2, T2 pleasant vs. T2 unpleasant) as factors. Effects of affective congruency on T1 and T2 report were evaluated by calculating for each participant the percentages of correct responses in trials with congruent (pleasant-pleasant, unpleasant-unpleasant) and incongruent (pleasant-unpleasant, unpleasant-pleasant) T1-T2 pairs. Report rates were then submitted to two-factorial ANOVA using T1-T2 SOA (2; 233 ms vs. 700 ms) and CONGRUENCY (2; congruent vs. incongruent) as factors. Differences in affective ratings were tested with t-tests.

Results

Subjective Ratings

Pleasant targets (mean arousal score: 6.3, SD = 1.2) and unpleasant targets (6.7, SD = 1.4) showed no difference on the arousal dimension, t(14) < 1.0. Pleasant targets had higher valence scores (7.9, SD = 0.5) than unpleasant targets (1.6, SD = 0.6), t(14) = 29.7, p < 0.001. On the concreteness dimension, pleasant (6.1, SD = 1.4) and unpleasant targets (6.7, SD = 1.1) obtained comparable scores, t(14) < 2.0.

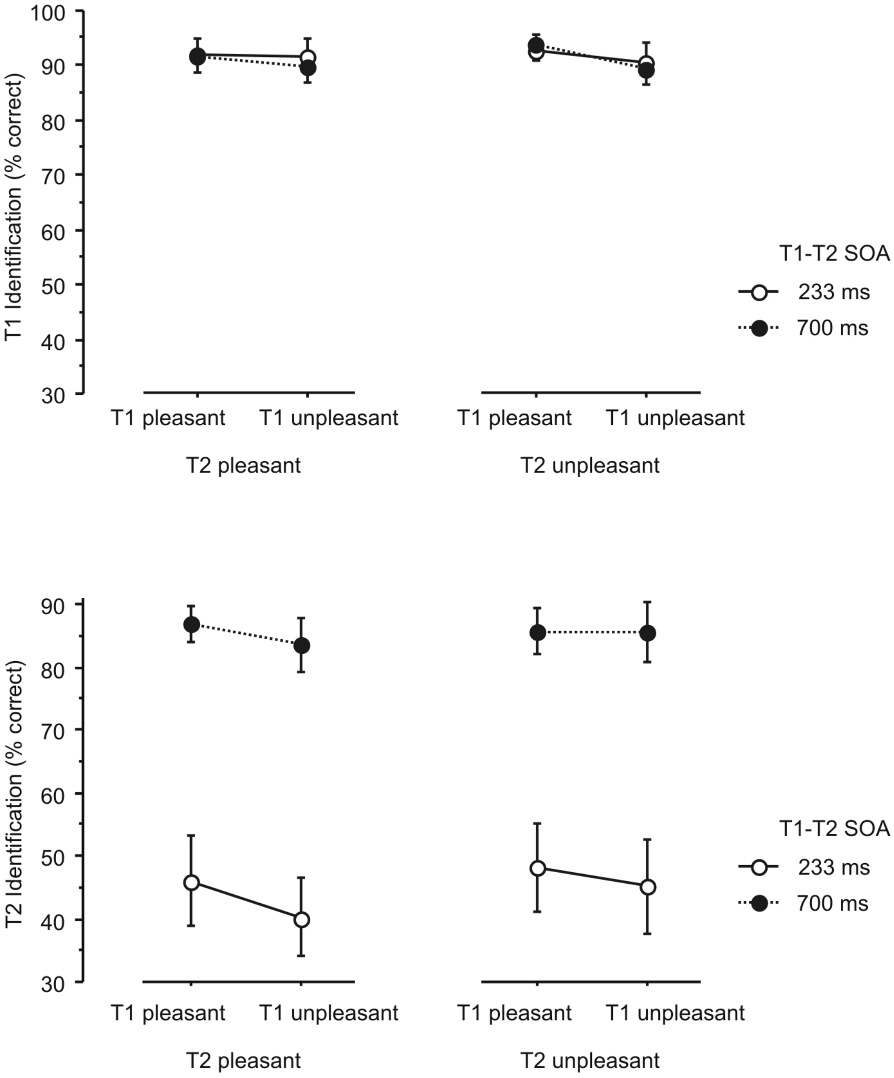

Figure 25. Mean identification accuracy (percent correct) for first targets (T1s, top panel) and second targets (T2s, bottom panel) as a function of T1-T2 SOA, T1 affective category, and T2 affective category in Experiment 4. Error bars indicate standard errors.

T1 Identification

Across conditions, T1s were reported as correct in 91.4 percent of the trials (SE = 1.0). T1 report rates did not vary with T1-T2 SOA, T1 or T2 CATEGOY, or any combination of these factors, all Fs < 2.5. T1 accuracy was not sensitive to affective congruence between T1 and T2. ANOVA using SOA and CONGRUENCE as factors yielded no significant effects, all Fs < 1.6.

T2 Identification

Mean T2 identification accuracy was 65.1 percent (SE = 2.7). ANOVA showed a main effect of T1-T2 SOA, F(1, 14) = 56.1, p < 0.001, reflecting higher report rates at the 700 ms SOA, compared to the 233 ms SOA . No significant main effect of T2 CATEGORY or any interaction of T2 CATEGORY with T1-T2 SOA or T1 CATEGORY emerged, all Fs < 1.6. However, ANOVAs showed a significant main effect of T1 CATEGORY, F(1, 14) = 10.7, p < 0.01. As illustrated in Figure 25, T2s following unpleasant T1s were identified less frequently, compared to T2s following pleasant items.

ANOVA using affective congruency instead of target category as factor showed a main effect of SOA, F(1, 14) = 56.7, p < 0.001. CONGRUENCY or the interaction SOA x CONGRUENCY did not affect report rates, Fs < 0.8. Hence, T2 report was not facilitated for affectively congruent versus incongruent T1-T2 pairs (see Figure 26).

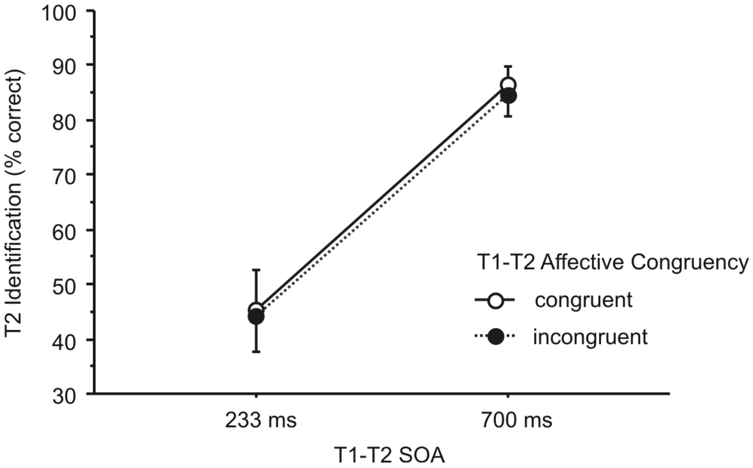

Figure 26. Mean identification accuracy (percent correct) for second targets (T2s) as a function of T1-T2 SOA and T1-T2 affective congruency in Experiment 4. T2 report rates were equal for T1-T2 pairs having congruent affective valence (pleasant-pleasant, unpleasant-unpleasant) and pairs having incongruent valence (pleasant-unpleasant, unpleasant-pleasant).

Discussion

Neither T1 nor T2 report were significantly enhanced for congruently valenced (pleasant-pleasant or unpleasant-unpleasant) compared to incongruent (pleasant-unpleasant and vice versa) T1-T2 pairs. Thus, spreading activation or response competition mechanisms related to affective priming (Klauer & Musch, 2003) are unlikely to influence dual-target identification during RSVP. The absence of affective priming in the present study corroborates the view that congruency effects occur reliably only in speeded motor-response tasks or in tasks involving evaluative responses (Hermans, De Houwer, & Eelen, 2001). It also invalidates an alternative account regarding the findings of Experiment 1, namely that reduced accuracy for neutral T2s preceded by pleasant or unpleasant T1s, compared to neutral-neutral target pairs, was driven by facilitatory effects of affective congruency for the latter stimulus configuration.

Further, Experiment 4 showed that improved T2 identification is contingent on (strong) differences in emotional intensity between target stimuli and does result from valence differences: Both T2 categories (pleasant and unpleasant) were associated with comparable scores on the arousal dimension and comparable report rates. As T1 accuracy was near ceiling (mean percent correct responses 91.4, SE = 1.3), we cannot safely infer that there were no effects of emotional category on T1 report. We also found that T2 report was inhibited for T2s following unpleasant T1s, compared to T2s subsequent to pleasant T1s. This result is surprising, as we expected resource interference (similar to facilitation effects) varying as a function of changes along the arousal dimension. Again, a possible cause for the presence of T1 inhibition effects in the current study is the nature of the T2 task. Similar to Experiment 1 but unlike Experiment 3, the T2 task comprised *no* variations in motivational relevance of the second target (both target categories obtained similar arousal scores). In contrast to Experiment 3, where T2 was either low-arousing neutral or high-arousing unpleasant, participants felt not obliged to withhold a certain amount of resources for the T2 task. As resources were unbound, the processing system became sensitive to small differences in the relevance of the T1 target. Statistical analysis of the pre-study rating scores indicated that there were indeed some differences in the arousal scores between the pleasant and unpleasant target category, yielding higher rated arousal for aversive compared to appetitive words, $t(29) = 4.0$, $p < 0.001$. Yet, these differences were not present in the post-experimental rating. The question if the inhibitory effects in Experiment 4 are related to valence-sensitive processes of T1 selection or reflect differential resource allocation based on small differences in arousal is subject to future work.

GENERAL DISCUSSION OF SERIES C

Emotional significant objects "soak up processing resources" (Bower, 1992, p. 17,), i.e., they grab attention away from concurrently or subsequently presented stimuli. The majority of studies demonstrating these interference effects have relied on tasks requiring simple motor responses (e.g., Hartikainen et al., 2000), and in case of successive tasks, on short distractor-target intervals (e.g., Calvo & Castillo, 2005). The present series of studies shows that during viewing of rapid word sequences identification of a colored target is impaired when the pre-target RSVP stream contains a stimulus associated with high self-rated emotional arousal. Further, we demonstrate that the costs of attending to briefly presented emotional information can extend across longer temporal distances between affective stimulus and primary task.

In Experiment 1, high-arousing (pleasant and unpleasant) first targets were reported as correct more frequently than neutral T1s. At the same time, emotional T1s were related to a more frequent failure to report a subsequent second (neutral) target. Reduced T2 identification accuracy was present across all T1-T2 SOAs (233 ms, 467 ms, and 700 ms). Using the same RSVP paradigm, in Experiment 2 sustained interference was observed for emotional stimuli clearly irrelevant to the task (pleasant and unpleasant pictures). Thus, temporally extended interference as elicited by a briefly presented emotional distractor does not require explicit attentional selection of the distracting stimulus. Rather, we propose that the time course of resource interference is determined by the complexity of the interfered task. Attention capture effects reported in tasks demanding low-level, pre-semantic processing, e.g., detection of picture rotation (Most et al., 2005), are transient, i.e., they vanish 200 or 300 ms after distractor onset. Conversely, resource drawing by briefly shown emotional distractors may have longer effects when the task requires word identification and working memory consolidation. Finally, lexical decision tasks (word-non-word-discrimination) may yield temporal interference effects having an intermediate duration. That is, they are impaired for a longer time window than low-level perceptual discriminations (see experimental series B) but for a shorter duration than the higher-order processes tapped here.

Apart from showing relative persistence of emotional interference, the present studies help to disentangle the role of negative valence versus emotional intensity/arousal in eliciting attention capture. The majority of previous work on emotional interference has examined the effects of aversive versus neutral material (Barnard et al., 2005; Pratto & John, 1991), often in the context of clinical or sub-clinical populations associated with high scores on anxiety or threat-sensitivity. Extending these approaches, we show that emotional interference is reliably obtained for participants recruited from a non-clinical student population and for stimuli from both valence categories (negative *and* positive). These results corroborate behavioral and physiological evidence showing a dominant role of emotional arousal in perceptual (Schupp et al., 2003), memory (Bradley et al., 1992), and interference (Schimmack & Derryberry, 2005) tasks (see also results of experimental series B). It should be noted that in the present experiments we demonstrate reliable effects of non-taboo emotional words, which have been suggested as too less arousing in order to elicit measurable behavioral responses (Hadley & Mackay, 2006). An integrative view on the CRT results of experimental series A and the results here suggests that target accuracy measures (e.g., Attentional Blink performance) may be more sensitive than target reaction times to reveal subtle processing differences between emotionally charged and neutral linguistic stimuli. The applied paradigm (identifying a word without prior knowledge about possible response alternatives) also overcomes some problems associated with two-choice motor responses. Zeelenberg et al. (2006) have argued that those measures are prone to response biases making inferences about emotional effects on stimulus processing problematic.

In Experiment 3 emotional interference was examined with regard to a primary task that itself elicited affective processing. Interestingly, unpleasant versus neutral T1s did not inhibit T2 identification when emotional intensity of the second target varied across trials. In contrast, for a task where T2 was *always* an arousing (pleasant or unpleasant) stimulus (Experiment 4), again, T2 report was sensitive towards emotional category of T1: T2s subsequent to pleasant T1s were reported with higher accuracy than T2s subsequent to (slightly higher arousing) unpleasant T1s. This differential pattern of interference effects supports the view that resource sharing strategies geared to the current experimental setup impact dual-target identification during RSVP. That is, emotional distractors are less likely to hamper a subsequent task when respondents expect having to distinguish between emotional significant and non-significant information. As soon as participants learn that stimulus' emotional relevance in the primary task does not change across trials, they become vulnerable to even small differences in resource allocation to the antecedent stimulus. Experiment 4 also demonstrated that report rates for T1-T2 pairs with congruent (pleasant-pleasant, unpleasant-unpleasant) and incongruent hedonic valence (pleasant unpleasant and vice versa) are equal. This ruled out the possible confounding of affective priming with the observed facilitation and interference effects.

Difficulties arise when one attempts to reconcile the present results with predominant models of dual-target performance, such as serial bottleneck accounts (e.g., Jolicoeur, 1998). Emotional interference effects did not show the typical time course associated with the AB, i.e., decreasing interference with increasing T1-T2 lag. Instead, we found comparable attention capture at each T1-T2 SOA. A challenging issue also concerns the finding of enhanced identification of emotional T1s in combination with T2 rates that were either reduced (Experiment 1) or did not vary with T1 category (Experiment 3). Several authors showed that higher report rates in the T1 task are associated with higher accuracy in the T2 task (Christmann & Leuthold, 2004; Chun & Potter, 1995), which reflects a pattern reversed to our results. According to the serial bottleneck framework, T1s that are easily discriminated from the distractor stream elicit a smaller delay of T2 memory consolidation and a decreased AB (Chun & Potter, 1995). The present results point out the importance of distinguishing between manipulations of difficulty (the ease of identification) and processes of resource allocation (see also McLaughlin, Shore, & Klein, 2001; Shore, McLaughlin, & Klein, 2001). Although arousing T1s were more easily identified, they did not save resources or allowed for earlier and more accurate T2 processing. Instead, they withdrew processing capacity available for performing the T2 task (given that participants did not expect emotional variations in the T2 task). This finding indicates that during RSVP resource allocation based on motivational significance might be functionally dissociated from resource allocation based on stimulus difficulty. Shore et al. (2001) showed that difficulty effects in AB paradigms depend on preparatory resource allocation in advance of each RSVP trial. This requires that difficulty conditions are blocked and participants become aware of the difficulty contingencies. In contrast, the results of experimental series C show a relation between the relevance of T1 and the amount of T2 impairment with non-blocked and randomly intermixed T1 conditions, which suggests that differential resource allocation occurred on-line during the RSVP stimulation.

Apart from resource models, another useful framework to account for the present findings is the notion that affective processing may be described in terms of distributed networks (Lang et al., 1997). As described above (see General Discussion of Series B), networks representing emotionally arousing stimuli are characterized by strong connections to memories, visceral systems, and motor systems that mediate executive functions, response preparation, and action towards the affective stimulus. During the course of competition with other activated RSVP stimulus representations, networks representing arousing T1s or T2s may thus benefit from higher connectivity and larger extension of their connections. In this view, facilitated identification relates to a decreased susceptibility of arousing targets' initial representations to distractor noise. Similar to the experiments of series B, neuronal competition between temporally overlapping activation patterns may also underlie emotional interference in the present task. Specifically, inhibition of T2 report by arousing T1s may be due to neurophysiological suppression mechanisms (Keysers & Perrett, 2002).

2.4 Emotional Processing and Audiovisual Integration (Series D)

In the fourth empirical part of this thesis, the relation between affective stimulus processing and task performance is explored with respect to multimodal aspects of perception[9]. During natural face-to-face conversation, the human brain effortlessly integrates the auditory speech signal with visual information conveyed by the speaker's talking lips and face into a unified percept. One approach to examine the process of audiovisual speech integration capitalizes on presenting discrepant linguistic input simultaneously to both modalities (de Gelder & Bertelson, 2003). As demonstrated by the McGurk illusion, auditory speech perception can be modified by the discrepant information from the visual channel (see Massaro, 1998). In the prototypical configuration, an auditory /ba/ frequently results in the perception of /da/ when it is synchronously dubbed on a speaker's face articulating /ga/ (McGurk & MacDonald, 1976).

Human communication also involves extraction of non-linguistic information on personal aspects of the speaker, such as identity, gender, and emotional state (Abercrombie, 1967). This information is communicated via both the auditory and visual modality (Lachs & Pisoni, 2004). Accordingly, several studies have demonstrated cross-modal interaction for non-linguistic emotional signals, a finding which is sometimes referred to as an "emotional McGurk effect" (de Gelder & Bertelson, 2003). For instance, emotional classification of a voice tone was biased towards the valence of a simultaneously presented picture showing an emotional face (e.g., de Gelder & Vroomen, 2000). In a related study, Pourtois and colleagues (Pourtois, de Gelder, Vroomen, Rossion, & Crommelinck, 2000) reported that affectively congruent face-voice pairs elicited cortical enhancement as early as 110 ms following onset of auditory stimulation, relative to incongruent stimulus pairs.

The question arises whether emotional processes can also modulate audiovisual integration of speech, i.e. phonetic classification. The McGurk paradigm provides an avenue to empirically address this issue. In this task, it is possible to manipulate affective parameters regarding facial expressions of emotion (visual channel) and with respect to affective prosody (auditory channel). The experiments of series D aimed to explore putative differences between processing audiovisual emotional speech and audiovisual neutral speech. To this end, we employed affectively consistent stimuli containing happy, neutral, and angry content in both channels. We expected that audiovisual fusion triggered by incongruent speech

---

[9] Data of experimental series D are included in the manuscript entitled "Look How is Shouting? Emotional Intensity Enhances Audiovisual Fusion", which is being prepared for publication.

stimulation (i.e., the McGurk effect) may be modulated by presentation of (consistent) affective information on the level of overt behavior.

EXPERIMENT 1

The purpose of Experiment 1 was to explore whether processing of discrepant audiovisual speech material (auditory /aba/ synchronized on visual /ada/ or /aga/) is modified by facial and vocal signatures of a happy or angry emotional state, relative to a neutral state, displayed by the speaker. Effects of emotion on audiovisual fusion were assessed by comparing the frequency of McGurk illusions for each affective category. Audiovisual speech integration was assumed when the forced-choice response (“aba”, “ada”, “aga”, or “other”) did not reflect the actual auditory stimulus (see Methods). Perceptual intelligibility of the stimulus material was tested in unimodal and bimodal congruent control trials intermixed with the McGurk stimuli. Further, post-hoc valence and arousal ratings were implemented to verify the validity of the affective manipulations of stimuli.

Methods

Participants

Thirty-one university students (15 females, 16 males; mean age 22.1 years, Standard Deviation = 2.5 years) gave informed consent to participate in Experiment 1. All subjects were native speakers of German, reported normal vision and hearing, and were naïve in terms of the goal of the study.

Stimuli

Stimulus selection was based on digital recordings of a professional male actor (age 63) who was filmed in frontal whole-head view against a blue background. Repeated utterances of three different vowel-consonant-vowel stimuli were video-taped: /aba/, /ada/, and /aga/. First, the actor pronounced the disyllables in neutral voice with a neutral facial expression. In a second and third run, the actor was instructed to pronounce the disyllables with an angry (happy) voice and corresponding angry (happy) facial expressions.

Three audio tokens of /aba/ were selected of each affective category (happy, neutral, and angry) and dubbed on two deviant visual tokens (/ada/ and /aga/), which were from the same affective category as the auditory part. Thus, the critical McGurk stimuli comprised 3 (auditory /aba/ items) x 2 (visual /ada/ and /aga/ items) x 3 (affective categories) different exemplars. Each of these incongruent stimuli was repeated four times across the whole session, resulting in 72 experimental trials. All stimuli were presented unimodally (visual and auditory alone) and in their bimodal congruent version to control for effects of intelligibility. As a result, the experiment comprised 3 (unimodal auditory, unimodal visual, bimodal congruent) x 27 (3 /aba/, 3 /ada/, 3 /aga/ x 3 affective categories) control stimuli. The sequence of experimental and control trials was randomized with identical stimuli never occurring in consecutive trials.

Procedure

After instruction, participants were seated in a dimly lit, sound attenuated experimental room, with their heads approximately 70 cm in front of a 19-inch monitor. Sound was delivered over two loudspeakers placed at the left and right side of the monitor. Each trial started with a blank screen presented for 1250 ms followed by a central fixation cross, displayed for 200 ms in order to ensure visual attention at the beginning of the trial. With regard to McGurk stimuli and bimodal/visual control stimuli, the fixation cross was replaced either by the video token (mean duration: 3190 ms), or in case of unimodal auditory stimuli by a blank screen with the same duration. After another blank screen of 500 ms duration participants were prompted via an entry form on the monitor to indicate what they had perceived. Using mouse cursor and left button, participants could choose between four alternative responses: “aba”, “ada”, “aga”, and “other“. Responses automatically started the next trial. Before starting the experimental session, participants performed 9 practice trials involving only neutral stimuli, which were not part of the experimental set.

Upon completing the session, subjects were asked to rate the stimuli in terms of their emotional valence (pleasure) and arousal. Rating was performed using a computer-based version of the SAM (Bradley & Lang, 1994) and encompassed the entire set of McGurk stimuli, unimodal auditory /aba/ items, and unimodal visual /ada/ and /aga/ exemplars.

Differences were evaluated using repeated measurement analyses of variance (ANOVA). Significant main effects pertaining to affective category were followed by linear and quadratic trend analyses, significant interaction effects were examined in detail employing Newman-Keuls post-hoc tests.

Results

Emotional Valence and Arousal Ratings

Mean affective ratings for each stimulus category (McGurk stimuli, unimodal auditory, and unimodal visual) are shown in Table 6. Within each stimulus category, pleasure scores varied with affective category (happy, neutral, and angry), $Fs(2, 60) > 194.6$, $ps < 0.001$; trend analysis confirmed a linear decrease from happy to neutral and angry exemplars, $ps < 0.001$. ANOVAs indicated a significant stimulus type (V/ada/ and V/aga/) by affective category interaction for the visual items, $F(2, 60) = 4.4$, $p < 0.05$, reflecting that angry V/aga/ stimuli were perceived more aversively than angry V/ada/ items, Newman-Keuls $p < 0.01$.

Affective categories also differed in terms of arousal ratings, $Fs(2, 60) > 85.5$, $ps < 0.001$. Quadratic trends for all stimulus categories, $ps < 0.001$, revealed that happy and angry stimuli yielded higher arousal scores than the neutral category. The stimulus type by affective category interaction reached significance for the McGurk items, $F(2, 60) = 4.6$, $p < 0.05$, and for the visual items, $F(2, 60) = 4.0$, $p < 0.05$. Higher arousal scores were found for happy A/aba/V/aga/ items compared to happy A/aba/V/ada/, for happy V/aga/ compared to happy V/ada/, and for angry V/aga/ compared to angry V/ada/, $ps > 0.05$.

Table 6. Affective ratings (mean [SE]) of the stimuli presented during Experiment 1.

Stimulus		Affective category					
		Happy	Neutral	Angry	Happy	Neutral	Angry
			Pleasure			Arousal	
McGurk stimuli	A/aba/V/ada/	7.3 (0.2)	4.3 (0.1)	2.2 (0.2)	3.5 (0.3)	3.1 (0.2)	6.3 (0.3)
	A/aba/V/aga/	7.3 (0.2)	4.4 (0.1)	2.1 (0.1)	3.9 (0.3)	3.1 (0.2)	6.3 (0.3)
Unimodal visual stimuli	V/ada/	7.2 (0.2)	4.3 (0.1)	2.3 (0.2)	3.4 (0.3)	3.3 (0.2)	6.0 (0.3)
	V/aga/	7.2 (0.2)	4.4 (0.1)	2.1 (0.2)	3.7 (0.3)	3.1 (0.2)	6.3 (0.3)
Unimodal auditory stimuli	A/aba/	6.4 (0.2)	4.7 (0.1)	2.3 (0.1)	3.6 (0.3)	3.1 (0.2)	5.8 (0.3)

Phonetic Classification: Experimental Trials

Effects of affective category on the McGurk illusion were evaluated by submitting percentages of “aba” responses in McGurk trials to a stimulus type (A/aba/V/ada/ and A/aba/V/aga/) x affective category (happy, neutral, and angry) repeated measurement ANOVA. Because “aba” responses for the McGurk stimuli corresponded to the auditory token, lower percentages were assumed to reflect enhanced audiovisual fusion. “aba” reports occurred in 12.7 % of McGurk trials (Standard Error SE = 2.0). A statistical trend towards modulation of “aba” responses by affective category was observed, F(2, 60) = 2.7, p = 0.076. Quadratic trend analysis confirmed that participants favored the auditory channel in trials with neutral stimuli as opposed to trials with happy or angry items (see Figure 27, left panel), F(1, 30) = 4.2, p < 0.05. No effect of stimulus type was evident.

Closer examination of the response pattern revealed that the dominant response for incongruent stimuli was “ada” (73.8 %, SE = 2.5), which is consistent with previous research (e.g., Shigeno, 2000). Analyses of “ada” responses showed main effects of stimulus type, F(1, 30) = 4.9, p < 0.05, and of affective category, F(2, 60) = 5.9, p < 0.01. Linear trend analysis verified that the proportion of “ada” responses increased from happy to neutral and from neutral to angry items, F(1, 30) = 11.8, p < 0.01 (see Figure 27, right panel).

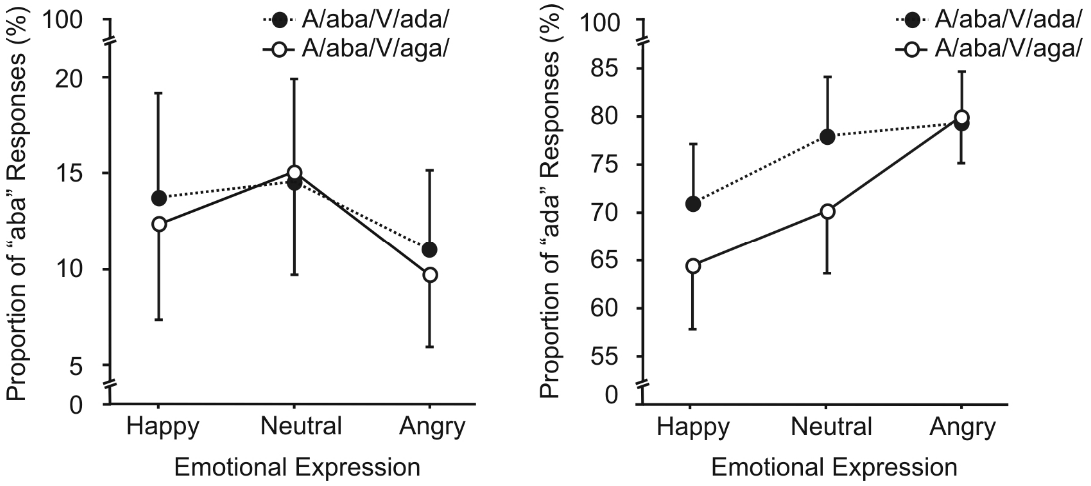

Figure 27. Results of the McGurk trials in Experiment 1. McGurk stimuli varied in terms of their visual component (auditory /aba/ + visual /ada/ versus auditory /aba/ + visual /aga/) and affective content. Left panel: Mean proportions of percepts corresponding to stimulation in the auditory modality (“aba” responses). Right panel: Mean proportions of the dominant illusionary response (“ada” responses). Error bars indicate standard errors.

“aga” and “other” were less favored response classes (4.5 %, SE = 0.6 for “aga” and 8.9 %, SE = 1.4 for “other”). For “aga” responses, ANOVAs demonstrated main effects of stimulus type, F(1, 30) = 14.6, p < 0.01, as well as affective category, F(2, 60) = 6.8, p < 0.01, and an interaction between both factors, F(2, 60) = 7.4, p < 0.01. With respect to the A/aba/V/aga/ configuration, post-hoc analyses indicated higher proportions of “aga” for happy compared to neutral and for neutral compared to angry stimuli, ps < 0.05. Such differences were not found for the A/aba/V/ada/ configuration. “other” responses were neither altered by stimulus type nor by affective category.

Identification Accuracy: Control Trials

Intelligibility was tested for bimodal /aba/ and auditory /aba/ in separate ANOVAs using percentage of correct responses as the dependent variable and affective category as the within-subject factor. To control for differences in visual salience, effects of affective category on lip-reading performance with visual /ada/ and /aga/ stimuli were evaluated. Mean percentages of correct responses are shown in the top panel of Figure 28.

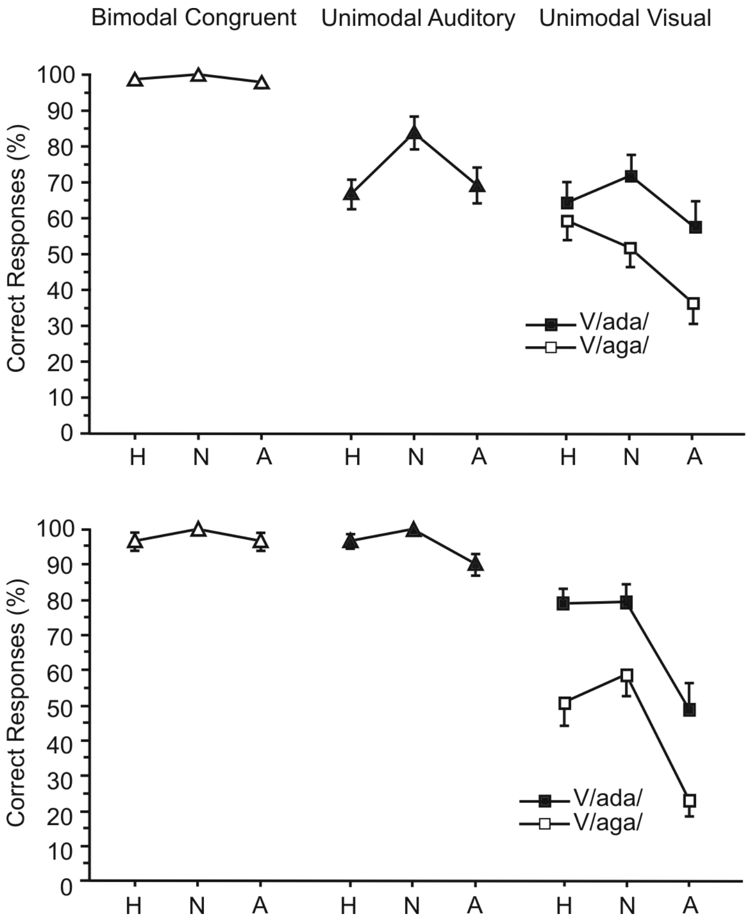

Figure 28. Results of the control trials for the two experiments. Upper graphs show mean percentages of correct responses to control stimuli (H = happy, N = neutral, A = angry expression), which were randomly intermixed with McGurk trials, in Experiment 1. Bottom graphs illustrate identification accuracy in Experiment 2, in which unimodal control trials were delivered prior to the McGurk trials. Error bars indicate standard errors.

Perception of /aba/ items in bimodal congruent trials was close to ceiling and did not covary with affective category, F(2, 60) = 1.0. Identification accuracy of auditory /aba/ items was sensitive to affective category, F(2, 60) = 4.8, $p < 0.05$, showing lower percentages for happy and angry exemplars compared to neutral ones, F(1, 30) = 9.7, $p < 0.01$. Performance was poor in trials with unimodal visual presentation. Lip-reading was influenced by stimulus type (V/ada/ and V/aga/), $F(1, 30) = 5.9$, $p < 0.05$, and affective category, $F(2, 60) = 6.1$, $p < 0.01$. In detail, angry items were identified less accurately compared to happy and neutral items, $F(1, 30) = 10.0$, $p < 0.01$.

Discussion

Experiment 1 demonstrates that fusion of audiovisual stimuli may depend on facial and prosodic signs of emotional state displayed by a speaker. Auditory /aba/ dubbed on incongruent lip-movements of /ada/ or /aga/ led listeners to neglect the auditory information (i.e., /aba/) when the speaker appeared happy or angry. Thus, fusion is facilitated. One methodological question arises from findings in control trials, however. The acoustic /aba/ items were identified with high accuracy and regardless of valence when presented with the /aba/ lip-movements in bimodal congruent control trials. On the other hand, the unimodal auditory /aba/ condition was related to lower identification accuracy for happy and angry compared to neutral items. Hence, the decrease of auditory responses and the corresponding increase of McGurk illusions for affective, incongruent stimuli might be directly linked to reduced perceptual intelligibility of some of the affective auditory /aba/ exemplars. A number of research teams have identified several temporal-spectral features, such as modulation of the fundamental frequency, being associated with the vocal expression of anger or happiness (see Bachorowski, 1999). While increasing the reliability with which listeners recognize and categorize emotive expressions (Banse & Scherer, 1996), these features can have detrimental effects on perception of the speech signal (Mullennix, Bihon, Bricklemyer, Gaston, & Keener, 2002). As a consequence, we designed a second experiment to examine whether affective modulation of the McGurk illusion occurred when perceptual intelligibility of the auditory component was controlled for.

EXPERIMENT 2

The present experiment tested unimodal identification accuracy of each auditory item prior to the McGurk trials. This provided a more direct measure of stimulus intelligibility that was unaffected by context effects arising when McGurk and control trials were randomly intermixed (Shigeno, 2002). For each participant, only those McGurk stimuli were evaluated, which involved auditory parts being correctly identified in the initial unimodal trials.

Methods

Participants

Data from 19 participants (7 females, 12 males; mean age 24.6 years, Standard Deviation = 3.4 years) were collected for Experiment 2. All subjects spoke German as their native language, reported normal vision and hearing, and were naïve in terms of the goal of the study.

Stimuli and Procedure

Stimuli, trial characteristics, and instruction were identical to Experiment 1. In contrast to Experiment 1, unimodal control and McGurk trials were presented in two blocks, separated by a break. In the pre-experimental first block, all unimodal auditory and all unimodal visual stimuli (27 stimuli in each case, see Experiment 1) were presented twice, resulting in 108 identification trials. The subsequent experimental block comprised 72 bimodal incongruent (McGurk) and 27 bimodal congruent stimuli (see Experiment 1) in random order. To maintain visual attention across the session, 9 lip-reading stimuli (3 happy, neutral, and angry tokens, respectively) were interspersed as catch trials. Again the session closed by emotional valence and arousal ratings of the stimuli.

Results and Discussion

Emotional Valence and Arousal Ratings

Table 7. Affective ratings (mean [SE]) of the stimuli presented during Experiment 2.

Stimulus		Affective category					
		Happy	Neutral	Angry	Happy	Neutral	Angry
		Pleasure			Arousal		
McGurk stimuli	A/aba/V/ada/	7.2 (0.2)	4.4 (0.2)	2.2 (0.1)	3.5 (0.3)	3.5 (0.5)	6.3 (0.3)
	A/aba/V/aga/	7.3 (0.2)	4.5 (0.2)	2.8 (0.4)	3.6 (0.3)	3.7 (0.5)	6.5 (0.3)
Unimodal visual stimuli	V/ada/	7.1 (0.2)	4.5 (0.2)	2.4 (0.2)	3.3 (0.3)	3.6 (0.5)	6.2 (0.3)
	V/aga/	6.9 (0.3)	4.3 (0.2)	2.3 (0.3)	3.0 (0.3)	3.6 (0.5)	6.8 (0.4)
Unimodal auditory stimuli	A/aba/	6.1 (0.2)	5.0 (0.1)	2.7 (0.2)	3.5 (0.3)	2.9 (0.4)	6.2 (0.3)

Results were similar to the judgments obtained after Experiment 1 (see Table 7). For each stimulus category, pleasure scores linearly decreased from angry to neutral and happy items, ps < 0.001. For unimodal auditory /aba/ items, happy and angry stimuli were associated with higher arousal compared to neutral items, $p < 0.001$. For McGurk and visual items, however, emotional arousal was higher for angry exemplars than for neutral and happy items, $ps < 0.001$. Unimodal visual /aga/ was rated as more arousing than visual /ada/, $F(2, 36) = 6.4$, $p < 0.01$, when angry tokens were considered, $p < 0.01$.

Identification Accuracy: Pre-Experimental Control Trials

Similar to Experiment 1, lip-reading performance was inferior for angry items compared to happy and neutral items, $p < 0.001$ (see Figure 28, bottom panel, p. 116). Again, identification for unimodal auditory items was most accurate for neutral items, $p < 0.01$. However, identification accuracy for auditory items in general was higher in Experiment 2 (95.5 %, SE = 1.3) than in Experiment 1 (73.3 %, SE = 2.8). Thus, in contrast to unimodal visual identification, classification of auditory items depended on the experimental context. In a similar vein, Bertelson, Vroomen and de Gelder (2003) demonstrated that frequent exposure to ambiguous auditory items dubbed on visual /aba/ or /ada/ can recalibrate auditory speech identification during post-experimental testing in favor of the visual component. Other authors reported that in the context of audiovisual stimulus classification, unimodal auditory items are sensitive to selective adaptation and anchoring (Roberts & Summerfield, 1981; Saldana & Rosenblum, 1994; Shigeno, 2002). In Experiment 1, unimodal auditory trials were embedded in incongruent A/aba/V/ada/ and A/aba/V/aga/ trials, while in Experiment 2 auditory items were delivered alone in a pre-experimental block. It is therefore conceivable that low auditory identification rates in Experiment 1 reflect an anchoring process, which typically occurs when a salient reference stimulus (anchor) is randomly scattered among the target stimuli (Shigeno, 2002).

Phonetic Classification: Experimental Trials

For each participant, proportions of “aba” responses were calculated for those McGurk stimuli whose auditory components were identified as correct during both unimodal presentations. In total, calculation of accuracy scores was based on at least 16 trials in each affective category.

Findings concerning phonetic classification of McGurk stimuli replicated those from Experiment 1. Averaged across conditions, incongruent stimuli elicited “aba” responses in 14.2 % (SE = 2.4) of the trials. “aba” reports were modulated by affective category, $F(2, 36) = 7.0$, $p < 0.01$, being reduced for happy and angry stimuli, $F(1, 18) = 8.8$, $p < 0.01$ (see Figure 29, left panel). Affective information also affected “ada” reports, which again constituted the dominant illusionary response class (81.0 %, SE = 2.6). Main effects were found for stimulus type, F(1, 18) = 7.6, $p < 0.05$, and affective category, F(2, 36) = 4.5, $p < 0.05$. Similar to Experiment 1, angry items elicited more “ada” responses than neutral and happy items, F(1, 18) = 5.1, $p < 0.05$ (see Figure 29, right panel). No effects of affective category or stimulus type were evident for the atypical illusionary responses “aga” (4.2 %, SE = 1.1) and “other” (0.6 %, SE = 0.3).

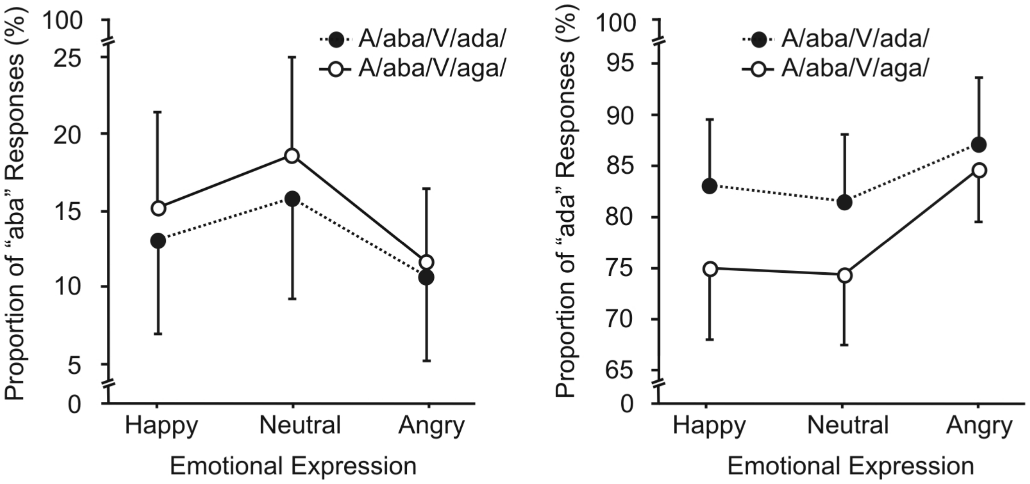

Figure 29. Results of the McGurk trials in Experiment 2. The graphs represent mean proportions of “aba” responses (top panel) and “ada” responses (bottom panel) for McGurk stimuli controlled for perceptual clarity of the auditory component. Error bars indicate standard errors.

Taken together, the results from Experiment 2 suggest that affective modulation of the McGurk illusion does not depend on perceptual factors. Pre-experimental unimodal data provided evidence that affective content induces perceptual interference on visual as well as auditory identification. However, controlling for perceptual ambiguity at the individual subjects level, we found that perceptual characteristics of the speech signal cannot account for affective facilitation of audiovisual fusion.

GENERAL DISCUSSION OF SERIES D

Two experiments demonstrated that processing of audiovisual speech stimuli being composed of discordant information can be modulated by affective content. A reduced frequency of responses corresponding to the auditory modality (i.e., “aba”) indicated facilitated intermodal binding specifically for affectively arousing stimuli. Phonetic classification in McGurk trials revealed that aversive information (i.e., angry) facilitated the typical illusionary fusion (i.e., “ada” for A/aba/V/ada/ and A/aba/V/aga/). Thus, fusion probability was modulated by emotional intensity or arousal, whereas valence-specific mechanisms mediated the particular outcome of the cross-modal integration process. Experiment 2 provided evidence that facilitation of audiovisual integration by affective parameters may occur independently of perceptual factors. Here, fusion was enhanced even when for each participant only McGurk stimuli having non-ambiguous auditory components were considered.

In both experiments, affective interference was also demonstrated for unimodal identification of visual items. Lip-movements of angry /ada/ and /aga/ were categorized with lower accuracy than lip-movements of neutral and happy exemplars. It is conceivable that the reduced perceptual clarity of angry visual items counteracted the diminished intelligibility of angry auditory items, resulting in comparable a priori probabilities of a McGurk illusion for neutral and angry stimulus configurations. This further supports the notion that affective modulation of audiovisual integration cannot be attributed to differences in perceptual clarity.

Affective facilitation of intermodal fusion might be due to propagation of perceptual enhancement (Schupp et al., 2003) to higher-order cortical regions of audiovisual binding, such as the superior temporal sulcus (Calvert & Campbell, 2003). Motivated attention (Lang et al., 1997) would thus act as a pacemaker of the fusion process. This account is in line with evidence showing that manipulation of intramodal endogenous visual attention is capable of modulating more downstream audiovisual integration (Tippana, Andersen, & M., 2004). Furthermore, van Wassenhove, Grant, and Poeppel (2005) have shown that visual information speeds up the cortical processing of auditory speech in a McGurk paradigm.

To conclude, audiovisual scene integration may be optimized in situations that are indexed as emotionally arousing by visual or auditory cues. The present approach allows studying the influence of different modalities on speech fusion and may be extended to measure context-specific aspects of communication in the future.

3. SYNTHESIS AND IMPLICATIONS

In the previous chapters, data of four experimental series have been reported and discussed that addressed the impact of affective processing on task performance. Emotion-performance interactions were approached using different levels of analysis (behavioral responses, surface and source electrocortical activation patterns), different stimulus channels (verbal versus pictorial material), different modalities (auditory, visual, audiovisual), different task demands (categorization, lexical decision, identification), and different response modes (speeded motor responses, non-speeded accuracy responses). Notwithstanding this apparent diversity, the pattern of results indicates a number of regularities mediating stimulus-response relations. In the following chapter, an attempt is made to extract the key dimensions, which may determine whether affective processing is associated with behavioral costs or benefits. A hierarchical model (see Figure 30) is delineated including i) one higher-order variable that influences behavioral outcome in a dichotomous way, i.e., determines the direction of behavioral effects (facilitation versus impairment), ii) intervening mechanisms that lead to divergent behavioral outcome in speeded versus non-speeded tasks, and iii) one dimension affecting the magnitude of facilitation or interference, respectively, once the direction of behavioral outcome has been determined. In addition, several subordinate variables are discussed that may exert minor influences on performance and are probably mediated by the first three dimensions (chapter 3.2). This thesis closes with a discussion of the model's limitations, outstanding issues, and possible future directions (chapter 3.3).

3. 1 Key Dimensions Predicting Task Performance during Affective Processing

3.1.1 The Direction of Behavioral Effects: Common versus Separate Networks

Network models of stimulus representation have become an attractive theoretical alternative to modular perspectives on stimulus processing (e.g., Fodor, 1983). In the network view, widely distributed but interconnected neural circuits form the neuronal basis of cognitive units such as words (Pulvermüller, 1999) or visual objects (von der Malsburg, 1995). These cell assemblies (Hebb, 1949) group single perceptual elements to organized wholes by synchronous activity of the network parts, which may be reflected in high-frequency oscillations (Keil, Müller, Ray, Gruber, & Elbert, 1999). They may comprise connections to a variety of brain regions, including sensory and motor areas (Pulvermüller, 1999). Emotion theories have adopted these notions arguing that emotion networks comprise connections to the appetitive or defensive motive system (see chapter 1.1.4), i.e., they are linked to specialized subcortical structures such as the amygdala. When applied to the level of single stimuli or stimulus features, the network perspective can be exploited to predict behavioral interference and facilitation as examined in the current context. In this view, interference occurs only for those stimuli or stimulus features that are represented in a network distinct from the network underlying the affective stimulus. Vive versa, a prerequisite for behavioral facilitation is that target stimuli/stimulus attributes are part of (or linked to) the network representing the emotional stimulus. The present facilitation effects are consistent with this view. In series C, we found enhanced identification rates for emotional T1s and T2s relative to neutral targets. Word identity/meaning may be considered as one element of the T1 representational network, which incorporates links to the motive circuits. Conscious access to the semantic branches of the network may be facilitated by excitatory influence of the motivational parts, possibly induced by spreading activation (Posner & Snyder, 1975). Facilitated audiovisual integration of emotional McGurk stimuli can be explained in a similar vein assuming a single polymodal network that connects auditory, visual and affective components. Conversely, in two experimental series yielding emotional interference effects (series B and C), emotional and target stimuli were clearly distinguishable as different information entities both by perceptual factors (e.g., pictures versus letter strings) and temporal segregation. Given that temporal segregation is not too large (approximately < 1 sec), impaired processing of stimuli following emotional (pleasant or unpleasant) information (e.g., pictures or RSVP T1s) may thus be attributed to a temporal overlap between two activated stimulus networks. More specifically, co-activation of stimulus networks will likely result in neuronal competition, which is biased in favor of the motivationally significant stimulus. A possible mechanism leading to reduction of concurrent network activity is neurophysiological suppression. Suppression may mediate effects of both attention (Desimone & Duncan, 1995) and masking (Keysers & Perrett, 2002). It can take place within a given processing level (lateral inhibition) or between processing stages, for instance, via descending connections from higher-order areas to lower processing stages (feedback inhibition). Series B offers evidence for neurophysiological suppression on a macroscopic cortical level showing reduction of neuronal activity in the extended visual cortices, namely in posterior and anterior inferior temporal regions. The temporal unfolding of interference across SOAs beyond 500 ms (as found in the RSVP experiments of series C) suggests that activity of emotional networks is upheld comparatively long and even for brief stimulus durations (below 120 ms). Such sustained network activation elicited by briefly presented emotional signals is also implicated by ERP studies. For instance, emotional pictures presented for 120 ms are related to LPP enhancement extending up to 700 ms after stimulus onset (Schupp, Junghöfer et al., 2004), which supports the idea that motivationally relevant information evokes continued perceptual analysis (Cuthbert et al., 2000). Similarly to temporal separation, neuronal competition (and thus interference) may also be elicited by spatial separation between affective and target stimulus (e.g., Erthal et al., 2005). However, as will be discussed in chapter 3.3, emotion-performance in spatial attention tasks may be qualitatively different to the mechanisms under consideration here.

An intriguing question is whether a separate target stimulus and an affective stimulus may be experimentally presented in a way that enables their integration into one perceptual entity and thus prevents network competition. Drawing from the Gestalt laws (Wertheimer, 1923), perceptual similarity, spatial vicinity, and temporal proximity of target and affective stimuli are three hypothetical conditions for the occurrence of perceptual integration/grouping and behavioral facilitation, respectively. It may be speculated that these conditions were met in the study by Phelps and colleagues (2006), who demonstrated that fearful versus neutral faces have the capacity to facilitate subsequent discrimination of grating patterns. In this study, target stimuli (grey-scale Gabor patches) and preceding affective stimuli (grey-scale faces from Ekman’s “pictures of facial affect”) were similar regarding both size and shape and presented within a time window of 125 ms. Experimentally testing the outlined hypotheses would help to strengthen the present network account in relation to a more parsimonious explanation, namely that interference versus facilitation depends on the role of affective stimuli being either an attended target or a (non-attended) distractor stimulus (see introduction). As a key advantage, the network approach adopted here allows more specific predictions based on stimulus and task properties.

3.1.2 Intervening Mechanisms: Threat Slowing in Speeded Tasks

The network account outlined above predicts beneficial effects of emotional content for accessing stimulus aspects that are activated as a part of the affective stimulus network. These aspects may include stimulus identity but also stimulus category, i.e., the stimulus’ association with semantic classes (e.g., animal versus human). In the choice reaction time experiments 1 - 3 of series A we showed that pleasant pictures are categorized faster and more accurately than neutral pictures (as pictures versus non-pictures or as items containing one or two persons). Positive affect also speeds up identification (Gallegos & Tranel, 2005) or evaluative discrimination of faces (Feyereisen et al., 1986; Leppänen et al., 2003). As will be delineated in chapter 3.1.3, these effects may be due to sensitivity towards certain perceptual features of pleasant stimuli. They thus do not directly relate to an activated appetitive motive system that modulates responses. However, on a descriptive level, for pleasant stimuli the network-variable appears sufficient to predict behavioral outcome, i.e., enhanced accuracy/faster responses related to aspects within the affective stimulus network and impaired performance for information presented in a separate network. The question arises why speed measures frequently fail to show facilitation effects for negative target stimuli (e.g., Bradley et al., 1992; De Cesarei & Codispoti, 2006). Similarly, according to the network account both pleasant *and* unpleasant pictures should have facilitated categorization. Further, in case of the emotional Stroop task one might argue that print color meets the above criteria of perceptual integration, i.e., font color is most likely part of the stimulus network representing dimensions of the displayed affective word. Thus, *faster* and not slower responses may be hypothesized for negative Stroop stimuli. Interestingly, in the few studies employing pleasant Stroop stimuli, these items yield equal or even marginally faster latencies relative to neutral stimuli (McKenna & Sharma, 1995; Richards, French, Johnson, Naparstek, & Williams, 1992; Rutherford, MacLeod, & Campbell, 2004).

Consequently, in case of aversive stimulus processing it appears useful to supplement the network variable by a second intervening factor, namely a threat-driven mechanism which slows processing during speeded tasks. In the discussion of series A, it was elaborated how this mechanism may delay information processing related to aversive events. Algom and co-workers have suggested a similar generic threat-sensitive mechanism effecting response slowing in the Stroop task (Algom et al., 2004). Worth mentioning, they showed delayed RTs also for lexical decisions and reading of threat words. Functionally, a pervasive slowdown of discriminative and/or response selection processes may serve to avoid erroneous or premature decisions having harmful consequences in case of possible threats. Parallels may also be drawn to the defense response towards aversive stimuli described by Lang (1997, see chapter 2.1). Slowdown of certain cognitive processes (such as response selection), however, does not preclude an increase of sensory intake and the gathering of stimulus or contextual information. Instead, as indicated by enhancement of late positive potentials, these processes may be commanded in a similar vein by both arousing pleasant and arousing unpleasant events. As a result, appetitive *and* aversive stimuli increase accuracy in non-speeded task, while they are related to diverging behavioral effects under speeded conditions. Consistent with this view, we found enhanced T1 accuracy and audiovisual integration for arousing stimuli (pleasant and unpleasant) in the non-speeded experiments of series C and D, and different effects for positive versus negative stimuli in the RT studies of series A. The outlined predictions are also in agreement with other behavioral findings: Non-speeded identification or recognition tasks frequently show facilitated performance for both positive and negative stimuli including affective words (Zeelenberg et al., 2006) and pictures (Trippe et al., 2007). As opposed to this, RT benefits in speeded paradigms are more consistently reported for pleasant stimuli (e.g., De Cesarei & Codispoti, 2006). A suitable future approach to test these assumptions may be the use of delayed responding instructions. Thus, withholding a speeded response up to the occurrence of a cue signal (e.g., 1000 ms after stimulus onset) may “bridge” the threat slowing period and lead to speed benefits also for aversive stimuli.

As detailed in the next sections, for pleasant stimuli the speed-versus-accuracy variable may represent an intervening factor with regard to another aspect: It may not mediate the direction of behavioral outcome (facilitation - interference) but determine whether emotional arousal increases or decreases the magnitude of behavioral modulation, once the direction is established (see Figure 30).

3.1.3 Determining the Magnitude of Facilitation/ Interference: Emotional Arousal

A common problem of experimental manipulations of emotion is that subjective ratings of valence and arousal are not independent. In the two-dimensional affective space, the correlation between arousal and valence reveals as two gradients reflecting the degree of appetitive and defensive motivation. Thus, in studies selecting stimuli on the basis of pleasantness/unpleasantness solely, effects of emotional arousal and valence extremity are hard to disentangle. In series A (Experiment 4) and B (Experiment 3), stimulus selection aimed at approximating an experimental design in which the factors of valence and arousal were crossed, i.e., responses to high-arousing pleasant/unpleasant stimuli should be contrasted with responses to low-arousing pleasant/unpleasant stimuli. Figure 30 illustrates the affective spaces formed by the stimuli employed in the CRT and lexical decision experiments. It can be seen that weak gradients reflecting more unpleasantness for arousing stimuli were still present for negative stimuli. Conversely, for pleasant stimuli the clusters of low- and high-arousing stimuli appeared comparable with respect to valence scores suggesting a weaker valence-arousal correlation for appetitive material. Thus, demonstrating arousal-modulated interference for pleasant stimuli provides strong evidence for a dominant role of emotional arousal in altering task performance, irrespective of valence extremity (see Schimmack & Derryberry, 2005). Such a relation between subjectively rated arousal and the magnitude of response impairment was shown in the lexical decision experiments of series B using the categorical analyses illustrated above. For both pleasant and unpleasant stimuli, the predictive value of emotional arousal with respect to the degree of interference was also confirmed by correlation analyses performed on an item- and subject-level. Arousal-related modulation also qualified choice reaction interference elicited by negative pictures in experimental series A. Taken together, these findings support the view that emotional arousal rather than valence extremity is associated with increased impairment of behavioral responses. Likewise, Schimmack and Derryberry (2005) demonstrated that arousal but not categorical negativity or evolutionary threat increases interference with executive aspects of attention. With respect to facilitation effects, a parametric (positive) relationship between rated arousal and performance has been shown in Attentional Blink studies (e.g., Keil & Ihssen, 2004) and memory tasks (e.g., Buchanan, Etzel, Adolphs, & Tranel, 2006). Further, arousal-related facilitation is also implicated by the present RSVP and McGurk experiments showing increased T1 accuracy and audiovisual fusion for high-arousing (pleasant and unpleasant) compared to low-arousing neutral stimuli.

Crucially, the current data indicate that for speeded responses to pleasant stimuli arousal and behavioral facilitation may be inversely related. In series A, CRTs were accelerated for mildly-arousing pleasant pictures specifically.[10] These findings are consistent with the speed benefits for pleasant words and happy faces (e.g., Leppänen et al., 2003), typically associated with moderate degrees of arousal. Two theoretical alternatives may be put forward to account for the relation between pleasant content and RT acceleration (see also chapter 2.1):

---

[10] At first glance, the CRT pattern (significant modulation for low- but not for high-arousing positive content and for high- but not for low-arousing negative content) shows an analogy to the processing asymmetries predicted for positive and negative affect at different levels of activation (Cacioppo & Berntson, 1999). The notions of a "positivity offset" and a "negativity bias" have been widely used by social psychologist to account for the complex interaction between emotion and task behavior (e.g., Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Crawford & Cacioppo, 2002; Rozin & Royzman, 2001; Taylor, 1991). They hypothesize stronger (motivational) responses to positive input at zero to low levels of activation and stronger responses to negativity at high arousal. However, this prediction compares response strength for positive and negative input *at the same* arousal level. Instead, the CRT pattern here concerns processing differences within a given valence category but *at different* arousal levels.

i) High-arousing (e.g., erotica) pleasant stimuli elicit a co-activation of the appetitive and defensive system. Although the Lang model postulates a reciprocal relationship between appetitive and defensive activation, it does not preclude the possibility of co-activation or dual motives (Lang, 1995), for instance during processing of ambiguous stimuli. In this view, erotica may evoke strong appetitive and a certain degree of defensive motivation. Appetitive motivation is related to facilitatory effects on processing speed. These benefits are, however, counteracted by the co-activated defensive system, which slows processing.

ii) The brain is highly sensitive towards appetitive object features (e.g., a smiling mouth) and enables fast processing of stimuli containing these features. As soon as a pleasant stimulus activates the appetitive system (i.e., the stimulus is arousing), these perceptual benefits are attenuated by a slowing mechanism similar to the one proposed for defensive motivation. Future work must examine whether sensitivity to appetitive features is due to *perceptual* salience of these cues (e.g., figure-ground segregation, see Bradley et al., 2007). Such feature differences have been proposed as a possible source of the happy-face-advantage during affective categorization (see Leppänen et al., 2003). Alternatively, sensitivity to appetitive features may reflect a (learned or hard-wired) property of the primate perceptual system. The net outcome of these opponent processes – sensitivity to pleasant features and response slowing during appetitive motivation – was that RTs to high-arousing pleasant pictures were only marginally faster relative to RTs elicited by neutral pictures. Yet, yielding mean CRTs still below the neutral baseline, one might describe emotional arousal as a dimension that does not change the direction but alters the magnitude of performance modulation.

Figure 30 summarizes how behavioral responses in presence of unpleasant (Figure 30.1) and pleasant (Figure 30.2) stimuli may be predicted using the three dimensions detailed above. Different routes are proposed corresponding to basic task/stimulus properties, with emotional arousal determining the amount of facilitation/interference[11]. Effects on the magnitude of facilitation/interference but not on the direction of effects are also assumed for the variables described in chapter 3.2.

---

[11] An alternative to dichotomous routes is the view that facilitation/interference reflects the *net effect* of the proposed processes, which exert potentially opponent influences. For instance, lexical decision RTs to negative words may reflect the sum of both facilitatory mechanisms (perceptual enhancement within the affective stimulus network) and inhibitory mechanisms (threat slowing of response selection during speeded tasks). This view may explain the diverging results in studies using this task, ranging from null effects (MacKay et al., 2004) to facilitation (Williamson et al., 1991) and interference (Algom et al., 2004).

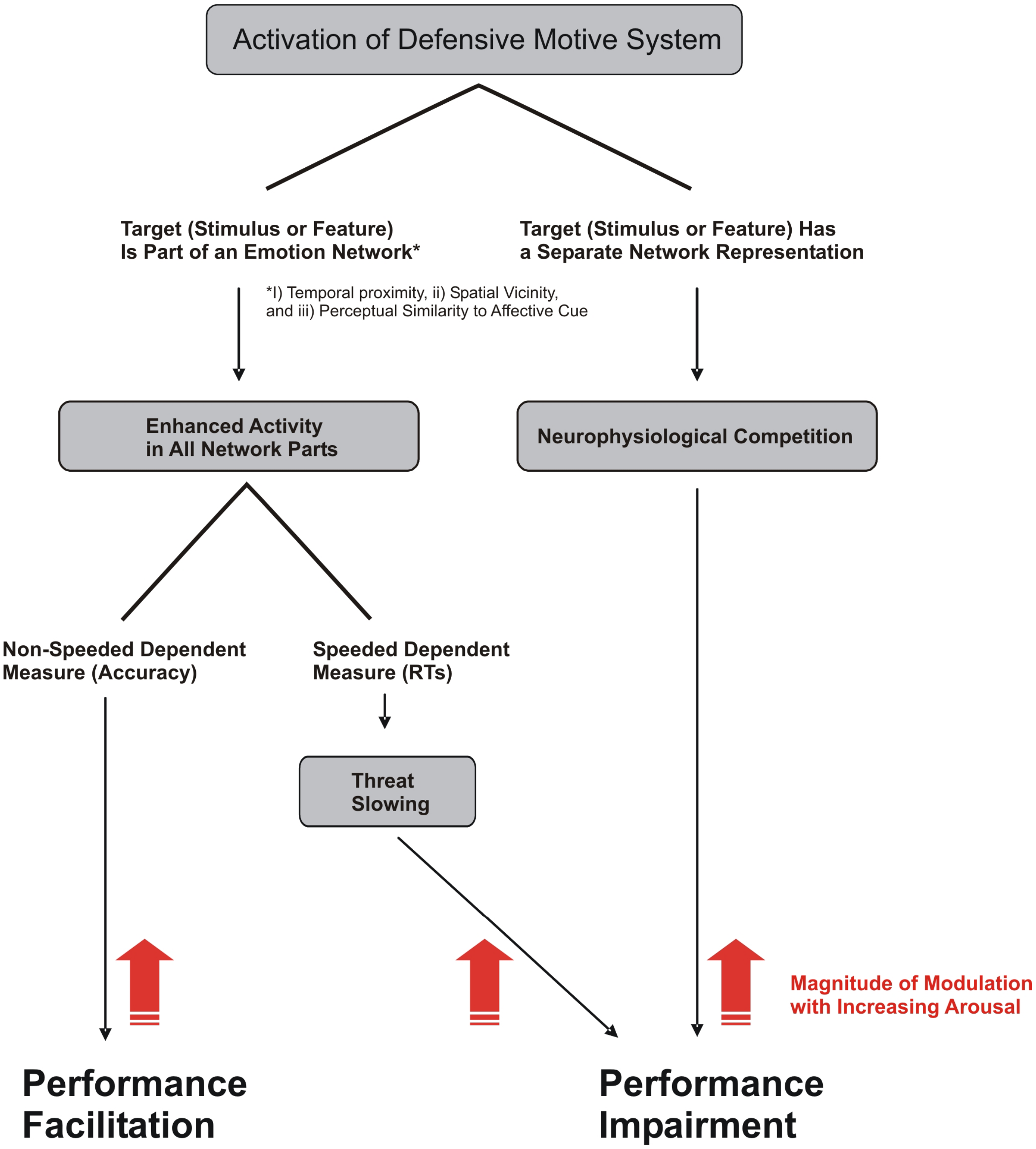

Figure 30.1. A hierarchical model predicting behavioral outcome (facilitation versus interference) in (non-spatial) perceptual/cognitive tasks with unpleasant stimuli as a function of two dichotomous variables: i) The target stimulus is integrated or not integrated in an emotion network (i.e., linked to hypothetical motive systems), ii) the task requires speeded or non-speeded processing. It is proposed that depending on these task and stimulus properties distinct mechanisms operate during information processing (mechanisms are illustrated in grey boxes). Emotional arousal is assumed to increase the magnitude of facilitation/interference once the direction of behavioral effects has been determined.

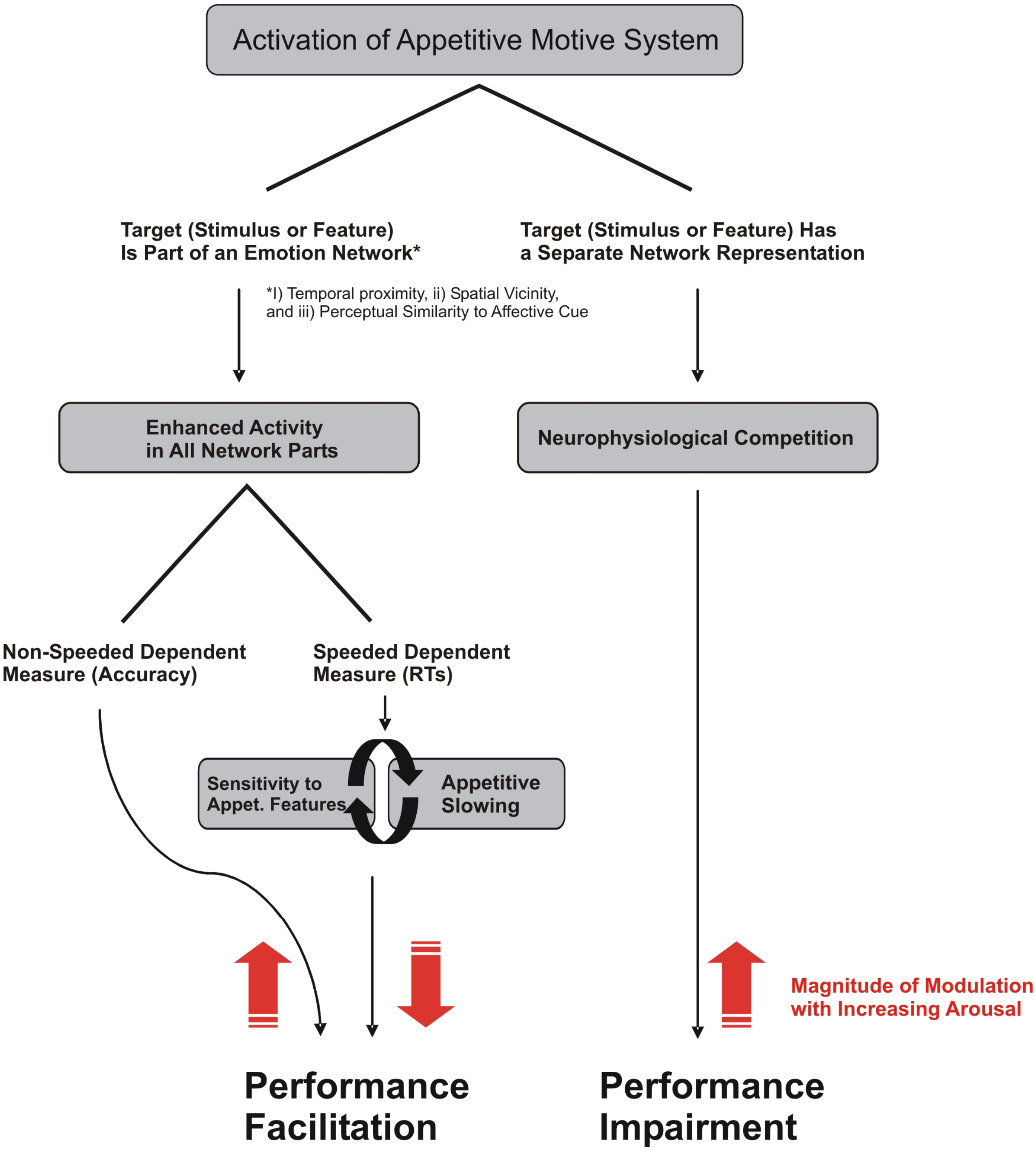

Figure 30.1. A hierarchical model predicting behavioral outcome (facilitation versus interference) in (non-spatial) perceptual/cognitive tasks with pleasant stimuli. It is proposed that speeded processing of pleasant stimuli involves two opponent mechanisms: i) High perceptual sensitivity to (and thus fast detection of) appetitive features (e.g., a smiling mouth) and ii) slowing of processing elicited by appetitive motivation.

3. 2 Subordinate Variables and Other Regularities

3.2.1 Stimulus Channel

The within-group comparison of CRTs to pictures versus words in series A (Experiment 1) indicates that affective modulation of task performance is stronger for content conveyed by photographic versus linguistic material. This difference is attributable to the correlation between the dimension of information/stimulus channel and the arousal dimension. Accordingly, stimulus channel may also affect the magnitude but not the direction of emotional response modulation. In addition, differences between emotional pictures and emotional words will be more evident in speeded tasks: Pictures elicit fast activation of object representations allowing emotional aspects to influence processing within 300 ms (see Keil, 2006). In contrast, emotional word content exerts a modulatory influence not before word form analysis is completed. For non-speeded tasks, the difference between verbal and picture emotion effects may thus be marginal. In speeded tasks, reliable emotion effects are predicted for pictures, while verbal material may depend on special classes of stimuli (e.g., taboo words) or specific populations (e.g., anxious persons) to elicit overt response modulation. In agreement with this, we demonstrated consistent variations of word responses in the non-speeded RSVP studies but not in the CRT studies.

3.2.2 Emotional Variations in the Target Task

The magnitude of the Attentional Blink is subject to individual resource allocation strategies tailored to changing difficulty conditions (e.g., Martens & Johnson, 2005). Series C demonstrates that these strategies may include motivational aspects. Interference induced by arousing T1s was mitigated when the T2 task varied in emotional content as well, reflecting strategic retention of resources. In a broader sense, emotional distractors are less likely to impair a subsequent task when participants expect having to distinguish between emotional significant and non-significant information. In contrast, when the category of the target is fixed across trials (as the case, for instance, in the lexical decision series), they may become more sensitive to emotional properties of the antecedent stimulus, i.e., interference increases. Similar to stimulus channel, learning of stimulus contingencies thus appears capable to modulate the magnitude of interference.

3.2.3 Semantic versus Non-semantic Processing

A key finding of series B was stronger emotional interference for word relative to pseudoword responses. On an electrocortical level, words but not pseudowords were related to reduced source strength in anterior inferior temporal regions suggesting suppressive effects for semantic processing specifically. Preliminary analysis of a behavioral follow-up experiment (N = 24) supports this assumption only partially. Using the same pictures and SOA (200 ms) as in the original study but varying verbal target tasks, emotion-driven response impairment increased from vowel detection (vowel versus consonant) to lexical decision, and semantic categorization (animal versus plant words). However, an orientation discrimination task (tilted versus upright letter strings) yielded the same degree of interference as the semantic task. Thus, not only semantic task aspects determined the amount of response impairment in this study. Another critical aspect relates to the experimental design used in the follow-up study. Tasks were presented in a blocked and not randomized experimental design. Though task order was balanced across subjects, it cannot be ruled out that the tasks were differentially affected by habituation.

3.2.4 Controlled versus Automatic Processing

Apart from semantic and habituation aspects, the tasks described above might have differed with respect to the complexity of stimulus-to-response-code matching. In many trials of the orientation task, participants were required to map left-tilted letter strings to the right mouse button and right-tilted strings to the left mouse button (compare the Simon task for similar task demands). Emotional interference may be weaker for concurrent tasks tapping automatic detection processes, such as the vowel task. A qualitative distinction between automatic and controlled forms of processing has been proposed in seminal work of Schneider and Shiffrin (1977; Shiffrin & Schneider, 1977). According to the dual-process model of information processing, highly skilled cognitive tasks such as word decoding primarily engage automatic processes, which are not capacity limited and allow for parallel processing (e.g., carrying out dual tasks). Controlled, serial, and capacity-limited forms of processing come into play during unskilled tasks or handling new information. In the current context, it may be argued that the magnitude of emotional interference in a given task increases with the amount of controlled processing required.[12] In turn, a reversed gradient may be predicted for emotional facilitation showing stronger benefits for highly automatic processes, such as

---

[12] A similar gradient of interference has been demonstrated for the color Stroop task (Klein, 1964). Color naming latencies increased as follows: random letter strings, pseudowords, low-frequency words, high-frequency words, words with color association, and color words.

polymodal binding examined in series D or word identification in series C. The distinction between emotion effects on automatic versus controlled processes may also account for the different time course of emotional interference in the identification tasks of series B (up to 700 ms) and in the detection task (< 300 ms) in the study by Most and colleagues (2005).

3. 3 Limitations, Outstanding Issues, and Future Directions

In the final sections of this thesis, several points are discussed showing limitations of the current approach and indicating possible directions of future work. First, it should be noted that the above list of variables, of course, is not exhaustive. The list represents an attempt to find regularities i) within the current data and ii) between the present results and the plethora of findings provided by other studies examining behavioral responses to affective stimuli. Further, the routes from appetitive/defensive motivation to behavioral facilitation/interference as proposed here are not immune against other intervening processes, (e.g., affective habituation/sensitization), let alone “noise” arising from inter-individual differences (e.g., anxiety, socio-cultural background) or situational factors. This notion is also pertinent with respect to a distinction between long-term strategies and short-term tactics in emotion (Lang et al., 1990). In this view, each emotional stimulus activates the appetitive or defensive motive system to some degree, reflecting a strategic disposition towards defense or approach behavior. However, the actual response is mainly determined by tactics adapted to the situational demands and specific (social and environmental) context of motivational activation. As discussed in the next section, another critical issue concerns the scope of the outlined predictions.

3.3.1 A special Role of Spatial Attention

Conflicting with the notion of threat slowing, unpleasant stimuli (e.g., faces or fear-relevant stimuli) are typically detected faster in spatial arrays of other stimuli (Öhman, Lundqvist et al., 2001). Further, unpleasant spatial cues do not slow but *accelerate* processing of subsequent probes (Stormark et al., 1995) although the probe stimuli may involve concurrent network activity (see above). Thus, the outlined predictions do not hold for tasks manipulating aspects of spatial attention, including visual search, spatial cueing, and dot-probe designs. However, affective modulation of visuospatial processing may be mediated by mechanisms different than those examined in the current studies. Thus, it appears that we can quickly say where a threat is but not what the threat is. This view is consistent with functional theories of visual encoding (Kahneman, Treisman, & Gibbs, 1992) suggesting that visual events are decomposed in visual types (What representations stored in long-term memory) and spatiotemporal tokens (Where and When representations). Related distinctions are made in anatomical and physiological studies showing that the cortical organization of vision comprises two functionally and anatomically differentiated processing streams (Ungerleider & Mishkin, 1982): A ventral "What"-stream projecting from occipital to anterior temporal regions and mediating object recognition, and a dorsal "Where-stream" extending from occipital to parietal regions and having visuomotor and spatial functions. The behavioral tasks discussed here are likely to engage activation of the ventral stream primarily. Further, as shown in Series B, the ventral processing systems, for instance inferior temporal cortical areas, are also target regions of the outlined regulatory mechanisms (e.g., neurophysiological suppression of concurrent network activity).

3.3.2 Increasing the Adaptiveness of Behavioral Responses

Emotions can be conceptualized as action dispositions (Lang, 1995, see chapter 1.1). In this theoretical framework, actions refer to adaptive responses executed in motivationally significant situational contexts and regulated by the defensive or appetitive motive systems. In situations that signal danger, the defensive system coordinates appropriate withdrawal or attack actions. Conversely, the appetitive system mediates a repertoire of approach actions, such as courtship, procreation/copulation, ingestion, nurturance and caregiving (Bradley et al., 2001). As opposed to these behaviors, responses in the present studies (e.g., pressing the left or right mouse button) evidently had no adaptive value. The question arises whether different results are obtained in an experimental setup that increases the adaptiveness of task behavior, i.e., when task responses come closer to attack/withdrawal or approach-related actions. A possible experimental avenue to pursue this question is offered in the studies of Chen and Bargh (1999) and Wentura et al. (Wentura, Rothermund, & Bak, 2000). In this research, different response modes were used to demonstrate that positive and negative word stimuli elicit approach and avoidance-related motor behavior, respectively. As a main finding, responses to negative stimuli were faster when participants pushed the response lever away (avoidance response) than when pulling it towards them (approach response). An opposite pattern was found for positive stimuli showing faster pulling responses relative to pushing responses. Remarkably, these results were obtained using both evaluative (good-bad classification) and non-evaluative (responses to stimulus onset, i.e., simple reaction times) task instructions. Similar findings have been reported in the study by Wentura (2000) showing facilitated and impaired lexical decision to emotional words depending on the required response type (pushing or releasing a response key). As argued by Flykt (2006), another response type that may come closer to adaptive behavior is spatial detection in visual search tasks. Here, the RT benefit for fear-relevant stimuli was attributed to faster action preparation elicited by threatening events.

With respect to the current data, an intriguing question is whether threat slowing to aversive stimuli may be eliminated or reversed using adaptive response modes, i.e., an avoid-condition, similar to the paradigm developed by Chen and Bargh (1999). Analogously, the speed benefits for pleasant pictures (see chapter 2.1) may further increase with approach-related response types.

3.3.3 Scrutinizing the Subprocesses of Emotion-Performance Interactions

The above results highlight the need to specify the impact of stimulus' emotional content on different subprocesses in the information processing chain. One outstanding issue pertains to the locus of emotional modulation observed during choice reactions (in series B). According to the subtractive logic (Donders, 1868), an useful extension of the present SRT and CRT paradigm may be the inclusion of a "c reaction task". In contemporary experimental terms, the "c reaction task" is a go-no-go paradigm. Go-no-go RTs involve no response selection processes. In combination with choice RTs, they thus offer the possibility to distinguish emotion effects on categorization versus response-code mapping processes.

In series A, the absence of threat slowing in the simple reaction times data indicated that motor inhibition does not account for threat slowing. However, future work is required to examine more thoroughly a possible motor-related origin of emotion-performance interactions. For instance, a posturographic study by Facchinetti and co-workers (Facchinetti, Imbiriba, Azevedo, Vargas, & Volchan, 2006) recently showed that viewing of unpleasant mutilation pictures as used here elicited freezing-like motor reactions in humans (immobility and rigidity behavior) that have been typically reported in animal studies. In another line of research, transcranial magnetic stimulation was employed to show enhancement of motor-evoked potentials during viewing of emotional pictures (Hajcak et al., 2007). In the current RT paradigms, an appropriate means to track emotion effects on motor processes might be the recording of lateralized readiness-potentials (LRPs), which decompose the stimulus-response chain in premotoric and motoric stages. Thus, stimulus-locked LRPs estimate the time required for information extraction (up to and including response selection), while response-locked LRPs indicate the duration of motor preparation and -execution in supplementary and primary motor cortices.

Taken together, the present thesis demonstrates that behavioral responses in tasks using emotional stimuli reflect a complex interplay among a variety of processes and dimensions, ranging from stimulus properties to experimental demands and individual processing strategies. Three key dimensions and underlying mechanisms were proposed that may help to predict emotional facilitation or emotional interference effects in a given cognitive task. Additional data obtained from different experimental approaches and at different levels of analysis are required to decipher these mechanisms more thoroughly. Thus, as noted by Algom and colleagues (2004, p. 336): “Science progresses by a never-ending process of differentiation and refinement. Phenomena considered identical or unitary are discovered to tap separate processes. Nuances and subtleties are noticed and elaborated. Eventually, the distinctions attained are reflected in the pertinent terminology and theory.”

REFERENCES

Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: Edinburgh University Press.

Adolphs, R., Cahill, L., Schul, R., & Babinsky, R. (1997). Impaired declarative memory for emotional material following bilateral amygdala damage in humans. *Learning and Memory, 4*(3), 291-300.

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*(3), 323-338.

Amaral, D. G., Behniea, H., & Kelly, J. L. (2003). Topographic organization of projections from the amygdala to the visual cortex in the macaque monkey. *Neuroscience, 118*(4), 1099-1120.

Anderson, A. K. (2005). Affective influences on the attentional dynamics supporting awareness. *Journal of Experimental Psychology: General, 134*(2), 258-281.

Anderson, A. K., & Phelps, E. A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature, 411*(6835), 305-309.

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory*. Washington, DC: Winston.

Armony, J. L., & Dolan, R. J. (2002). Modulation of spatial attention by fear-conditioned stimuli: an event-related fMRI study. *Neuropsychologia, 40*(7), 817-826.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The Celex lexical database (Release 2) [CD-ROM]*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania [Distributor].

Bachorowski, J. A. (1999). Vocal expression and perception of emotion. *Current directions in psychological science, 8*, 53-57.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*(3), 614-636.

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology, 62*(6), 893-912.

Barnard, P. J., Ramponi, C., Battye, G., & Mackintosh, B. (2005). Anxiety and the deployment of visual attention over time. *Visual Cognition, 12*(1), 181-211.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*(4), 323-370.

Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: time course and scalp distribution. *Journal of Cognitive Neuroscience, 11*(3), 235-260.

Berg, P., & Scherg, M. (1994a). A fast method for forward computation of multiple-shell spherical head models. *Electroencephalography and Clinical Neurophysiology, 90*(1), 58-64.

Berg, P., & Scherg, M. (1994b). A multiple source approach to the correction of eye artifacts. *Electroencephalography and Clinical Neurophysiology, 90*(3), 229-241.

Bertelson, P., Vroomen, J., & De Gelder, B. (2003). Visual recalibration of auditory speech identification: a McGurk aftereffect. *Psychological Science, 14*(6), 592-597.

Binder, J. R., McKiernan, K. A., Parsons, M. E., Westbury, C. F., Possing, E. T., Kaufman, J. N., & Buchanan, L. (2003). Neural correlates of lexical access during visual word recognition. *Journal of Cognitive Neuroscience, 15*(3), 372-393.

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*(8), 1484-1504.

Bower, G. H. (1992). How might emotions affect learning? In S. Christianson (Ed.), *Handbook of emotion and memory: Research and theory* (pp. 3-32). Hillsdale, NJ: Erlbaum.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: defensive and appetitive reactions in picture processing. *Emotion, 1*(3), 276-298.

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1996). Picture media and emotion: effects of a sustained affective context. *Psychophysiology, 33*(6), 662-670.

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, and Cognition, 18*(2), 379-390.

Bradley, M. M., Hamby, S., Low, A., & Lang, P. J. (2007). Brain potentials in perception: Picture complexity and emotional arousal. *Psychophysiology, 44*(3), 364-373.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the Self-Assessment Manikin and the Semantic Differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49-59.

Bradley, M. M., Sabatinelli, D., Lang, P. J., Fitzsimmons, J. R., King, W., & Desai, P. (2003). Activation of the visual cortex in motivated attention. *Behavioral Neuroscience, 117*(2), 369-380.

Buchanan, T. W., Denburg, N. L., Tranel, D., & Adolphs, R. (2001). Verbal and nonverbal emotional memory following unilateral amygdala damage. *Learning and Memory, 8*(6), 326-335.

Buchanan, T. W., Etzel, J. A., Adolphs, R., & Tranel, D. (2006). The influence of autonomic arousal and semantic relatedness on memory for emotional words. *International Journal of Psychophysiology, 61*(1), 26-33.

Buodo, G., Sarlo, M., & Palomba, D. (2002). Attentional resources measured by reaction times highlight differences within pleasant and unpleasant arousing stimuli. *Motivation and Emotion, 26*(2), 123-138.

Cacioppo, J. T., & Berntson, G. G. (1999). The affect system: Architecture and operating characteristics. *Current Directions in Psychological Science, 8*(5), 133-137.

Cacioppo, J. T., Crites Jr., S. L., Berntson, G. G., & Coles, M. G. H. (1993). If attitudes affect how stimuli are processed, should they not affect the event-related brain potential? *Psychological Science, 4*, 108-112.

Calvert, G. A., & Campbell, R. (2003). Reading speech from still and moving faces: the neural substrates of visible speech. *Journal of Cognitive Neuroscience, 15*(1), 57-70.

Calvo, M. G., & Castillo, M. D. (2005). Foveal vs. parafoveal attention-grabbing power of threat-related information. *Experimental Psychology, 52*(2), 150-162.

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience, 7*(3), 308-313.

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin, 25*(2), 215-224.

Christmann, C., & Leuthold, H. (2004). The attentional blink is susceptible to concurrent perceptual processing demands. *Quarterly Journal of Experimental Psychology A, 57*(2), 357-377.

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance, 21*(1), 109-127.

Codispoti, M., Ferrari, V., & Bradley, M. M. (2007). Repetition and Event-related Potentials: Distinguishing Early and Late Processes in Affective Picture Perception. *Journal of Cognitive Neuroscience, 19*(4), 577-586.

Codispoti, M., Ferrari, V., De Cesarei, A., & Cardinale, R. (2006). Implicit and explicit categorization of natural scenes. *Progress in Brain Research, 156*, 53-65.

Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: the case for the visual word form area. *Neuroimage, 22*(1), 466-476.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., & Michel, F. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain, 123 (Pt 2)*, 291-307.

Coltheart, V. (1999). Perceiving and remembering brief visual stimuli. In V. Coltheart (Ed.), *Fleeting Memories: Cognition of brief visual stimuli*. Cambridge, Massachusetts: MIT Press.

Crawford, L. E., & Cacioppo, J. T. (2002). Learning where to look for danger: Integrating affective and spatial information. *Psychological Science, 13*(5), 449-453.

Cuthbert, B. N., Schupp, H. T., Bradley, M., McManis, M., & Lang, P. J. (1998). Probing affective pictures: attended startle and tone probes. *Psychophysiology, 35*(3), 344-347.

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: covariation with autonomic arousal and affective report. *Biological Psychology, 52*(2), 95-111.

Damasio, A. R. (1994). *Descartes error: emotion, reason and the human Brain*. New York: GP Putnam’s Sons.

De Cesarei, A., & Codispoti, M. (2006). When size does not matter? Effects of stimulus size on affective modulation. *Psychophysiology, 43*, 207-215.

de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in Cognitive Science, 7*(10), 460-467.

de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion, 14*(3), 289-311.

Dehaene, S., Le Clec, H. G., Poline, J. B., Le Bihan, D., & Cohen, L. (2002). The visual word form area: a prelexical representation of visual words in the fusiform gyrus. *Neuroreport, 13*(3), 321-325.

Derryberry, D. (1991). The immediate effects of positive and negative feedback signals. *Journal of Personality and Social Psychology, 61*, 267–278.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193-222.

Di Lollo, V., Kawahara, J., Shahab Ghorashi, S. M., & Enns, J. T. (2005). The attentional blink: Resource depletion or temporary loss of control? *Psychological Research, 69*(3), 191-200.

Dickinson, A., & Dearing, M. F. (1979). Appetitive-aversive interactions and inhibitory processes. In A. Dickinson & R. A. Boakes (Eds.), *Mechanisms of Learning and Motivation* (pp. 203–231). Hillsdale, NJ: Erlbaum.

Dijksterhuis, A., & Aarts, H. (2003). On wildebeests and humans: the preferential detection of negative stimuli. *Psychological Science, 14*(1), 14-18.

Dijksterhuis, A., & Smith, P. K. (2002). Affective habituation: subliminal exposure to extreme stimuli decreases their extremity. *Emotion, 2*(3), 203-214.

Donders, F. C. (1868). Die Schnelligkeit psychischer Prozesse. *Archiv für Anatomie und Physiologie und wissenschaftliche Medizin*, 657-681.

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Perception and Psychophysics, 63*(6), 1004-1013.

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2003). Negative facial expression captures attention and disrupts performance. *Perception and Psychophysics, 65*(3), 352-358.

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*(4), 388-395.

Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: is it really attention free? *Journal of Experimental Psychology: Human Perception & Performance, 31*(6), 1476-1492.

Everitt, B. J., Cador, M., & Robbins, T. W. (1989). Interactions between the amygdala and ventral striatum in stimulus–reward associations: studies using a second-order schedule of sexual reinforcement. *Neuroscience, 30*, 63–75.

Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. (2001). A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience, 13*(2), 171-180.

Facchinetti, L. D., Imbiriba, L. A., Azevedo, T. M., Vargas, C. D., & Volchan, E. (2006). Postural modulation induced by pictures depicting prosocial or dangerous contexts. *Neuroscience Letters, 410*(1), 52-56.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*(2), 229-238.

Fernandez, G., Heitkemper, P., Grunwald, T., Van Roost, D., Urbach, H., Pezer, N., Lehnertz, K., & Elger, C. E. (2001). Inferior temporal stream for word processing with integrated mnemonic function. *Human Brain Mapping, 14*(4), 251-260.

Feyereisen, P., Malet, C., & Martin, Y. (1986). Is the faster processing of expressions of happiness modality-specific? In H. D. Ellis, M. A. Jeeves, F. Newcombe & A. Young (Eds.), *Aspects of face processing* (pp. 349–355). Boston: Martinus Nijhoff.

Fischler, I., & Bradley, M. (2006). Event-related potential studies of language and emotion: words, phrases, and task effects. *Progress in Brain Research, 156*, 185-203.

Flykt, A. (2006). Preparedness for action: responding to the snake in the grass. *American Journal of Psychology, 119*(1), 29-43.

Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge, Massachusetts: MIT Press.

Fox, E., & Damjanovic, L. (2006). The eyes are sufficient to produce a threat superiority effect. *Emotion, 6*(3), 534-539.

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*(4), 681-700.

Fredrikson, M., Wik, G., Greitz, T., Eriksson, L., Stone-Elander, S., Ericson, K., & Sedvall, G. (1993). Regional cerebral blood flow during experimental phobic fear. *Psychophysiology, 30*(1), 126-130.

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

Gaillard, R., Del Cul, A., Naccache, L., Vinckier, F., Cohen, L., & Dehaene, S. (2006). Nonconscious semantic processing of emotional words modulates conscious access. *Proceedings of the National Academy of Science USA, 103*(19), 7524-7529.

Gallegos, D. R., & Tranel, D. (2005). Positive facial affect facilitates the identification of famous faces. *Brain and Language, 93*(3), 338-348.

Giesbrecht, B., & Di Lollo, V. (1998). Beyond the attentional blink: visual masking by object substitution. *Journal of Experimental Psychology: Human Perception and Performance, 24*(5), 1454-1466.

Gray, J. A. (1983). A theory of anxiety: the role of the limbic system. *L'Encephale, 9*(4 Suppl 2), 161B-166B.

Greenwald, A. G., Klinger, M. R., & Liu, T. J. (1989). Unconscious processing of dichotically masked words. *Memory and Cognition, 17*, 35-46.

Hadley, C. B., & Mackay, D. G. (2006). Does emotion help or hinder immediate memory? Arousal versus priority-binding mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(1), 79-88.

Hajcak, G., Molnar, C., George, M. S., Bolger, K., Koola, J., & Nahas, Z. (2007). Emotion facilitates action: A transcranial magnetic stimulation study of motor cortex excitability during picture viewing. *Psychophysiology, 44*(1), 91-97.

Hansen, C. H., & Hansen, R. D. (1988). Finding the face in the crowd: an anger superiority effect. *Journal of Personality and Social Psychology, 54*(6), 917-924.

Harris, C. R., & Pashler, H. (2004). Attention and the processing of emotional words and names: not so special after all. *Psychological Science, 15*(3), 171-178.

Harris, C. R., Pashler, H. E., & Coburn, N. (2004). Moray revisited: high-priority affective stimuli and visual search. *Quarterly Journal of Experimental Psychology, 57*(1), 1-31.

Hartikainen, K. M., Ogawa, K. H., & Knight, R. T. (2000). Transient interference of right hemispheric function due to automatic emotional processing. *Neuropsychologia, 38*(12), 1576-1580.

Hauk, O., Davis, M. H., Ford, M., Pulvermüller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage, 30*(4), 1383-1400.

Hebb, D. O. (1949). *The organization of behavior. A neuropsychological theory*. New York: John Wiley.

Hermans, D., De Houwer, J., & Eelen, P. (2001). A time course analysis of the affective priming effect. *Cognition and Emotion, 15*(143-165).

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition, 92*(1-2), 67-99.

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 353*(1373), 1257-1270.

Hohlfeld, A., Mierke, K., & Sommer, W. (2004). Is word perception in a second language more vulnerable than in one's native language? Evidence from brain potentials in a dual task setting. *Brain and Language, 89*(3), 569-579.

Hopfinger, J. B., & Mangun, G. R. (1998). Reflexive attention modulates processing of visual stimuli in human extrastriate cortex. *Psychological Science, 9*, 441– 447.

Jolicoeur, P. (1998). Modulation of the attentional blink by on-line response selection: evidence from speeded and unspeeded task decisions. *Memory and Cognition, 26*(5), 1014-1032.

Junghöfer, M., Bradley, M. M., Elbert, T. R., & Lang, P. J. (2001). Fleeting images: a new look at early emotion discrimination. *Psychophysiology, 38*(2), 175-178.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: object-specific integration of information. *Cognitive Psychology, 24*, 175–219.

Keil, A. (2006). Macroscopic brain dynamics during verbal and pictorial processing of affective stimuli. *Progress in Brain Research, 156*, 217-232.

Keil, A., Bradley, M. M., Hauk, O., Rockstroh, B., Elbert, T., & Lang, P. J. (2002). Large-scale neural correlates of affective picture processing. *Psychophysiology, 39*(5), 641-649.

Keil, A., & Ihssen, N. (2004). Identification facilitation for emotionally arousing verbs during the attentional blink. *Emotion, 4*(1), 23-35.

Keil, A., Ihssen, N., & Heim, S. (2006). Early cortical facilitation for emotionally arousing targets during the attentional blink. *BMC Biology, 4*(1), 23.

Keil, A., Moratti, S., Sabatinelli, D., Bradley, M. M., & Lang, P. J. (2005). Additive effects of emotional content and spatial selective attention on electrocortical facilitation. *Cerebral Cortex, 15*(8), 1187-1197.

Keil, A., Müller, M. M., Gruber, T., Wienbruch, C., Stolarova, M., & Elbert, T. (2001). Effects of emotional arousal in the cerebral hemispheres: a study of oscillatory brain activity and event-related potentials. *Clinical Neurophysiology, 112*(11), 2057-2068.

Keil, A., Müller, M. M., Ray, W. J., Gruber, T., & Elbert, T. (1999). Human gamma band activity and perception of a gestalt. *Journal of Neuroscience, 19*(16), 7152-7161.

Keil, A., Stolarova, M., Heim, S., Gruber, T., & Müller, M. M. (2003). Temporal stability of high-frequency brain oscillations in the human EEG. *Brain Topography, 16*(2), 101-110.

Keysers, C., & Perrett, D. I. (2002). Visual masking and RSVP reveal neural competition. *Trends in Cognitive Science, 6*(3), 120-125.

Kissler, J., Assadollahi, R., & Herbert, C. (2006). Emotional and semantic networks in visual word processing: insights from ERP studies. *Progress in Brain Research, 156*, 147-183.

Kitayama, S. (1990). Interaction between affect and cognition in word perception. *Journal of Personality and Social Psychology, 58*(2), 209-217.

Klauer, K. C., & Musch, J. (2003). Affective priming: Findings and theories. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 7 - 49): Mahwah, NJ, US: Lawrence Erlbaum Associates, Publishers.

Klein, G. S. (1964). Semantic power measured through the interference of words with color-naming. *American Journal of Psychology, 77*, 576-588.

Knight, D. C., Smith, C. N., Stein, E. A., & Helmstetter, F. J. (1999). Functional MRI of human Pavlovian fear conditioning: Patterns of activation as a function of learning. *Neuroreport, 10*, 3665–3670.

Konorski, J. (1967). *Integrative activity of the brain: an interdisciplinary approach.* Chicago: The University of Chicago Press.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Science, 4*(12), 463-470.

Lachs, L., & Pisoni, D. B. (2004). Cross-modal source information and spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 30*(2), 378-396.

Lane, R. D., Reiman, E. M., Bradley, M. M., Lang, P. J., Ahern, G. L., Davidson, R. J., & Schwartz, G. E. (1997). Neuroanatomical correlates of pleasant and unpleasant emotion. *Neuropsychologia, 35*(11), 1437-1444.

Lang, P. J. (1979). A bio-informational theory of emotional imagery. *Psychophysiology, 16*(6), 495-512.

Lang, P. J. (1993). The network model of emotion: motivational concerns. In R. S. Wyer & T. K. Scrull (Eds.), *Advances in social cognition* (pp. 109-133). Hillsdale, NJ: Erlbaum.

Lang, P. J. (1995). The emotion probe. Studies of motivation and attention. *American Psychologist, 50*(5), 372-385.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review, 97*(3), 377-395.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). Motivated Attention: Affect, Activation, and Action. In P. J. Lang, R. F. Simons & M. T. Balaban (Eds.), *Attention and Orienting: Sensory and Motivational Processes* (pp. 97-135). Hillsdale, N.J: Lawrence Erlbaum Associates.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biological Psychiatry, 44*, 1248-1263.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. Technical Report A-6. University of Florida, Gainesville, FL.

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*(2), 199-210.

Lang, P. J., & Davis, M. (2006). Emotion, motivation, and the brain: reflex foundations in animal and human research. *Progress in Brain Research, 156*, 3-29.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience, 23*, 155-184.

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.

Linden, D. E. (2005). The p300: where in the brain is it produced and what does it tell us? *Neuroscientist, 11*(6), 563-576.

Mack, A., Pappas, Z., Silverman, M., & Gay, R. (2002). What we see: inattention and the capture of attention by meaning. *Consciousness and Cognition, 11*, 488–506.

MacKay, D. G., & Ahmetzanov, M. V. (2005). Emotion, memory, and attention in the taboo Stroop paradigm. *Psychological Science, 16*(1), 25-32.

MacKay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., & Dyer, J. R. (2004). Relations between emotion, memory, and attention: evidence from taboo stroop, lexical decision, and immediate memory tasks. *Memory and Cognition, 32*(3), 474-488.

Maki, W. S., Frigen, K., & Paulson, K. (1997). Associative priming by targets and distractors during rapid serial visual presentation: does word meaning survive the attentional blink? *Journal of Experimental Psychology: Human Perception and Performance, 23*(4), 1014-1034.

Martens, S., & Johnson, A. (2005). Timing attention: cuing target onset interval attenuates the attentional blink. *Memory and Cognition, 33*(2), 234-240.

Martin-Loeches, M., Hinojosa, J. A., Gomez-Jarabo, G., & Rubia, F. J. (2001). An early electrophysiological sign of semantic processing in basal extrastriate areas. *Psychophysiology, 38*(1), 114-124.

Massaro, D. W. (1998). *Perceiving talking Faces: From Speech Perception to a Behavioral Principle*. Cambridge, MA: MIT Press.

Matthews, G., Pitcaithly, D., & Mann, R. L. E. (1995). Mood, neuroticism, and the encoding of affective words. *Cognitive Therapy and Research, 19*(5), 563-587.

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in Cognitive Science, 7*(7), 293-299.

McCarthy, G., & Donchin, E. (1981). A metric for thought: a comparison of P300 latency and reaction time. *Science, 211*(4477), 77-80.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746-748.

McKenna, F. P., & Sharma, D. (1995). Intrusive cognitions: An investigation of the emotional stroop task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(6), 1595-1607.

McKenna, F. P., & Sharma, D. (2004). Reversing the emotional stroop effect reveals that it is not what it seems: the role of fast and slow components. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(2), 382-392.

McLaughlin, E. N., Shore, D. I., & Klein, R. M. (2001). The attentional blink is immune to masking-induced data limits. *Quarterly Journal of Experimental Psychology A, 54*(1), 169-196.

Micrografx. (1997). Graphics Suite 2: Micrografx.

Milders, M., Sahraie, A., Logan, S., & Donnellon, N. (2006). Awareness of faces is modulated by their emotional meaning. *Emotion, 6*(1), 10-17.

Miller, J., Patterson, T., & Ulrich, R. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology, 35*(1), 99-115.

Miller, J., & Ulrich, R. (2003). Simple reaction time and statistical facilitation: a parallel grains model. *Cognitive Psychology, 46*(2), 101-151.

Mineka, S. (1992). Evolutionary memories, emotional processing, and the emotional disorders. In D. L. Medin (Ed.), *The psychology of learning and motivation* (pp. 161-206). San Diego, CA: Academic Press.

Mogg, K., Bradley, B. P., de Bono, J., & Painter, M. (1997). Time course of attentional bias for threat information in non-clinical anxiety. *Behaviour Research and Therapy, 35*(4), 297-303.

Morris, J. S., de Gelder, B., Weiskrantz, L., & Dolan, R. J. (2001). Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. *Brain, 124*(Pt 6), 1241-1252.

Morris, J. S., Öhman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature, 393*(6684), 467-470.

Morris, J. S., Öhman, A., & Dolan, R. J. (1999). A subcortical pathway to the right amygdala mediating "unseen" fear. *Proceedings of the National Academy of Science USA, 96*(4), 1680-1685.

Most, S. B., Chun, M. M., Widders, D. M., & Zald, D. H. (2005). Attentional rubbernecking: cognitive control and personality in emotion-induced blindness. *Psychonomic Bulletin & Review, 12*(4), 654-661.

Mullennix, J. W., Bihon, T., Bricklemyer, J., Gaston, J., & Keener, J. M. (2002). Effects of variation in emotional tone of voice on speech perception. *Language and Speech, 45*(Pt 3), 255-283.

Näätänen, R. (1992). *Attention and brain function*. Hillsdale, NJ: Erlbaum.

Nieuwenstein, M. R., & Potter, M. C. (2006). Temporal limits of selection and memory encoding. *Psychological Science, 17*(6), 471-475.

Nobre, A. C., Allison, T., & McCarthy, G. (1994). Word recognition in the human inferior temporal lobe. *Nature, 372*(6503), 260-263.

Ogawa, T., & Suzuki, N. (2004). On the saliency of negative stimuli: Evidence from attentional blink. *Japanese Psychological Research, 46*(1), 20-30.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: detecting the snake in the grass. *Journal of Experimental Psychology: General, 130*(3), 466-478.

Öhman, A., Flykt, A., & Lundqvist, D. (2000). Unconscious emotion: Evolutionary perspectives, psychophysiological data and neuropsychological mechanisms. In R. Lane, L. Nadel & A. Kaszniak (Eds.), *The cognitive neuroscience of emotion* (pp. 296-327). New York: Oxford University Press.

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*(3), 381-396.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh Inventory. *Neuropsychologia, 9*, 97-113.

Olivers, C. N., van der Stigchel, S., & Hulleman, J. (2005). Spreading the sparing: against a limited-capacity account of the attentional blink. *Psychological Research*, 1-14.

Olson, I. R., Chun, M. M., & Anderson, A. K. (2001). Effects of phonological length on the attentional blink for words. *Journal of Experimental Psychology: Human Perception and Performance, 27*(5), 1116-1123.

Osgood, C., Suci, G., & Tannenbaum, P. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois.

Panksepp, J. (1992). A critical role for "affective neuroscience" in resolving what is basic about basic emotions. *Psychological Review, 99*, 554-560.

Pegna, A. J., Khateb, A., Lazeyras, F., & Seghier, M. L. (2005). Discriminating emotional faces without primary visual cortices involves the right amygdala. *Nature Neuroscience, 8*(1), 24–25.

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology, 72*(2), 184-187.

Petersen, S. E., Fox, P. T., Snyder, A. Z., & Raichle, M. E. (1990). Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. *Science, 249*(4972), 1041-1044.

Phelps, E. A. (2006). Emotion and cognition: insights from studies of the human amygdala. *Annual Review of Psychology, 57*, 27-53.

Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological Science, 17*(4), 292-299.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3-25.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information Processing and Cognition* (pp. 55–85). Hillsdale, NJ: Erlbaum.

Potter, M. C. (1975). Meaning in visual search. *Science, 187*(4180), 965-966.

Potter, M. C., Dell'Acqua, R., Pesciarelli, F., Job, R., Peressotti, F., & O'Connor, D. H. (2005). Bidirectional semantic priming in the attentional blink. *Psychonomic Bulletin & Review, 12*(3), 460-465.

Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B., & Crommelinck, M. (2000). The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport, 11*(6), 1329-1333.

Pourtois, G., Grandjean, D., Sander, D., & Vuilleumier, P. (2004). Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cerebral Cortex, 14*(6), 619-633.

Pratto, F., & John, O. P. (1991). Automatic vigilance: the attention-grabbing power of negative social information. *Journal of Personality and Social Psychology, 61*(3), 380-391.

Price, C. J., & Devlin, J. T. (2003). The myth of the visual word form area. *Neuroimage, 19*(3), 473-481.

Pritchard, W. S., Houlihan, M. E., & Robinson, J. H. (1999). P300 and response selection: a new look using independent-components analysis. *Brain Topography, 12*(1), 31-37.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences, 22*(2), 253-279.

Putman, P., Hermans, E., & van Honk, J. (2004). Emotional stroop performance for masked angry faces: it's BAS, not BIS. *Emotion, 4*(3), 305-311.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114*(3), 510-532.

Raymond, J. E. (2003). New objects, not new features, trigger the attentional blink. *Psychological Science, 14*(1), 54-59.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: an attentional blink? *Journal of Experimental Psychology: Human Perception and Performance, 18*(3), 849-860.

Richards, A., French, C. C., Johnson, W., Naparstek, J., & Williams, J. (1992). Effects of mood manipulation and anxiety on performance of an emotional Stroop task. *British Journal of Psychology, 83*, 479-491.

Ritter, W., Simson, R., & Vaughan, H. G., Jr. (1983). Event-related potential correlates of two stages of information processing in physical and semantic discrimination tasks. *Psychophysiology, 20*(2), 168-179.

Roberts, M., & Summerfield, Q. (1981). Audiovisual presentation demonstrates that selective adaptation in speech perception is purely auditory. *Perception and Psychophysics, 30*(4), 309-314.

Rolls, E. T. (1990). A theory of emotion, and its application to understanding the neural basis of emotion. *Cognition and Emotion, 4*, 161-190.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review, 5*(4), 296-320.

Rudell, A. P. (1992). Rapid stream stimulation and the recognition potential. *Electroencephalography and Clinical Neurophysiology, 83*(1), 77-82.

Rutherford, E. M., MacLeod, C., & Campbell, L. W. (2004). Negative selectivity effects and emotional selectivity effects in anxiety: differential attentional correlates of state and trait variables. *Cognition & Emotion, 18*(5), 711-720.

Sabatinelli, D., Bradley, M. M., Fitzsimmons, J. R., & Lang, P. J. (2005). Parallel amygdala and inferotemporal activation reflect emotional intensity and fear relevance. *Neuroimage, 24*(4), 1265-1270.

Sabatinelli, D., Flaisch, T., Bradley, M. M., Fitzsimmons, J. R., & Lang, P. J. (2004). Affective picture perception: Gender differences in visual cortex? *Neuroreport: For Rapid Communication of Neuroscience Research, 15*(7), 1109-1112.

Sabatinelli, D., Lang, P. J., Keil, A., & Bradley, M. M. (2007). Emotional perception: correlation of functional MRI and event-related potentials. *Cerebral Cortex, 17*(5), 1085-1091.

Saldana, H. M., & Rosenblum, L. D. (1994). Selective adaptation in speech perception using a compelling audiovisual adaptor. *Journal of the Acoustical Society of America, 95*(6), 3658-3661.

Scherer, K. R. (1984). Emotion serves to decouple stimulus and response. In P. Ekman & R. J. Davidson (Eds.), *The Nature of Emotion: Fundamental Questions* (pp. 127-130). New York: Oxford University Press.

Scherg, M., Ille, N., Bornfleth, H., & Berg, P. (2002). Advanced tools for digital EEG review: virtual source montages, whole-head mapping, correlation, and phase analysis. *Journal of Clinical Neurophysiology, 19*(2), 91-112.

Schimmack, U., & Derryberry, D. (2005). Attentional interference effects of emotional pictures: threat, negativity, or arousal? *Emotion, 5*(1), 55-66.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*, 1-66.

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Hillman, H. H., Hamm, A. O., & Lang, P. J. (2004). Brain processes in emotional perception: Motivated attention. *Cognition and Emotion, 18*, 593-611.

Schupp, H. T., Flaisch, T., Stockburger, J., & Junghöfer, M. (2006). Emotion and attention: event-related brain potential studies. *Progress in Brain Research, 156*, 31-51.

Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2003). Emotional facilitation of sensory processing in the visual cortex. *Psychological Science, 14*(1), 7-13.

Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2004). The selective processing of briefly presented affective pictures: an ERP analysis. *Psychophysiology, 41*(3), 441-449.

Schupp, H. T., Öhman, A., Junghofer, M., Weike, A. I., Stockburger, J., & Hamm, A. O. (2004). The facilitated processing of threatening faces: an ERP analysis. *Emotion, 4*(2), 189-200.

Schupp, H. T., Stockburger, J., Bublatzky, F., Junghoefer, M., Weike, A. I., & Hamm, A. O. (2007). Explicit attention interferes with selective emotion processing in human extrastriate cortex. *BMC Neuroscience, 8*(1), 16.

Schupp, H. T., Stockburger, J., Codispoti, M., Junghofer, M., Weike, A. I., & Hamm, A. O. (2007). Selective visual attention to emotion. *Journal of Neuroscience, 27*(5), 1082-1089.

Schupp, H. T., Stockburger, J., Codispoti, M., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2006). Stimulus novelty and emotion perception: the near absence of habituation in the visual cortex. *Neuroreport, 17*(4), 365-369.

Shapiro, K., Schmitz, F., Martens, S., Hommel, B., & Schnitzler, A. (2006). Resource sharing in the attentional blink. *Neuroreport, 17*(2), 163-166.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127 - 190.

Shigeno, S. (2000). Influence of vowel context on the audio-visual speech perception of voiced stop consonants. *Japanese Psychological Research, 42*(3), 155-167.

Shigeno, S. (2002). Anchoring effects in audiovisual speech perception. *Journal of the Acoustical Society of America, 111*(6), 2853-2861.

Shore, D. I., McLaughlin, E. N., & Klein, R. M. (2001). Modulation of the attentional blink by differential resource allocation. *Canadian Journal of Experimental Psychology, 55*(4), 318-324.

Sprengelmeyer, R., Rausch, M., Eysel, U. T., & Przuntek, H. (1998). Neural structures associated with recognition of facial expressions of basic emotions. *Proceedings of the Royal Society of London: Biological Sciences, 265*, 1927–1931.

Stalans, L., & Wedding, D. (1985). Superiority of the left hemisphere in the recognition of emotional faces. *International Journal of Neuroscience, 25*, 219–233.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donder's method. In W. G. Roster (Ed.), *Attention and performance II* (pp. 276-315). Amsterdam: North-Holland.

Stolarova, M., Keil, A., & Moratti, S. (2006). Modulation of the C1 visual event-related component by conditioned stimuli: evidence for sensory plasticity in early affective perception. *Cerebral Cortex, 16*(6), 876-887.

Stormark, K. M., Hugdahl K., Posner M.I. (1999). Emotional modulation of attention orienting: A classical conditioning study. *Scandinavian Journal of Psychology, 40*(2), 91-99.

Stormark, K. M., Nordby, H., & Hugdahl, K. (1995). Attentional shifts to emotionally charged cues: Behavioral and ERP data. *Cognition and Emotion, 9*, 507-523.

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: the mobilization-minimization hypothesis. *Psychological Bulletin, 110*(1), 67-85.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*(6582), 520-522.

Tippana, K., Andersen, T. S., & M., S. (2004). Visual attention modulates audiovisual speech perception. *European Journal of cognitive psychology, 16*(3), 457-472.

Trippe, R. H., Hewig, J., Heydel, C., Hecht, H., & Miltner, W. H. (2007). Attentional Blink to emotional and threatening pictures in spider phobics: Electrophysiology and behavior. *Brain Research*.

Ulrich, R., Rinkenauer, G., & Miller, J. (1998). Effects of stimulus duration and intensity on simple reaction time and response force. *Journal of Experimental Psychology: Human Perception & Performance, 24*(3), 915-928.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–586). Cambridge, MA: MIT Press.

van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences of the USA, 102*(4), 1181-1186.

Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature, 383*(6597), 254-256.

VanRullen, R., & Thorpe, S. J. (2001). The time course of visual processing: from early perception to decision-making. *Journal of Cognitive Neuroscience, 13*(4), 454-461.

Verleger, R. (1997). On the utility of P3 latency as an index of mental chronometry. *Psychophysiology, 34*(2), 131-156.

Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance, 24*(6), 1656-1674.

von der Malsburg, C. (1995). Binding in models of perception and brain function. *Current Opinion in Neurobiology, 5*(4), 520-526.

Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends in Cognitive Science, 9*(12), 585-594.

Vuilleumier, P., Arrmony, J., Dolan, R. (2003). Reciprocal links between emotion and attention. In R. Frackowiak, Mazziotta, J. (Ed.), *Human Brain Function* (2 ed., pp. 419-444). San Diego: Academic Press.

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*(11), 1271-1278.

Watts, F. N., McKenna, F. P., Sharrock, R., & Trezise, L. (1986). Colour naming of phobia-related words. *British Journal of Psychology, 77 (Pt 1)*, 97-108.

Wentura, D., Rothermund, K., & Bak, P. (2000). Automatic vigilance: the attention-grabbing power of approach- and avoidance-related social information. *Journal of Personality and Social Psychology, 78*(6), 1024-1037.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. *Psychologische Forschung, 4*, 301-350.

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience, 18*(1), 411-418.

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin, 120*, 3-24.

Williamson, S., Harpur, T. J., & Hare, R. D. (1991). Abnormal processing of affective words by psychopaths. *Psychophysiology, 28*(3), 260-273.

Wingenfeld, K., Bullig, R., Mensebach, C., Hartje, W., Driessen, M., & Beblo, T. (2006). Attention bias towards personally relevant stimuli: the individual emotional Stroop task. *Psychological Reports, 99*(3), 781-793.

Zald, D. H. (2003). The human amygdala and the emotional evaluation of sensory stimuli. *Brain Research Reviews, 41*(1), 88-123.

Zeelenberg, R., Wagenmakers, E. J., & Rotteveel, M. (2006). The impact of emotion on perception: bias or enhanced processing? *Psychological Science, 17*(4), 287-291.

SUMMARY

A common perspective in contemporary emotion research states that the human brain processes emotionally arousing information in a near-optimal manner. This view has been derived from brain imaging work showing that viewing (or hearing) of emotionally charged stimuli enhances cortical responses at various stages of processing. However, empirical findings are less conclusive as to how affective stimulus perception translates into behavioral responses, especially when mental operations are carried out on the stimulus. The present thesis centers on the questions whether and when task performance is facilitated or impaired by emotional content of a stimulus. Behavioral and electrophysiological data of four experimental series are reported and discussed capitalizing on emotion-performance interactions in different perceptual and cognitive tasks. Aversive picture content was found to delay visual discrimination of natural scenes. Conversely, pleasant content accelerated categorization responses, especially when pictures were judged as mildly and not highly arousing. The behavioral pattern was mirrored by event-related brain potential measures (late positive potential latencies) indicative of cognitive processing time. Early electrocortical activation patterns pointed to the operation of three independent mechanisms underlying the response modulation: i) early sensitivity to appetitive features (between 90 - 130 ms), ii) selection of arousing events (200 - 250 ms), and, iii) attending to threat (300 - 400 ms). Detrimental effects of stimulus' emotional content were found for subsequent processing of neutral information. Arousing pleasant and unpleasant pictures prolonged lexical decision times to neutral word/pseudowords irrespective of stimulus-onset asynchrony (80 ms, 200 ms, 440 ms) and salient semantic category differences (e.g., erotica versus mutilation pictures). On an electrocortical level, interference was reflected in reduced N1 (204 - 264 ms) and late positivity (LP, 404 - 704 ms) responses. Regional source modeling indicated that reduction effects originated from inhibited cortical activity in posterior areas of the left inferior temporal cortex associated with orthographic processing. Modeling of later reduction effects argued for interference in distributed semantic networks. Sustained interference with subsequent processing was also demonstrated for emotional words and pictures preceding word targets during Rapid Serial Visual Presentation. Finally, using a McGurk paradigm, facial and vocal expressions of emotion (pleasant and unpleasant) were associated with enhanced probability of audiovisual fusion. A hierarchical model is delineated, attempting to account for the outlined results. Task performance is predicted based on three dimensions: i) Connections of the target stimulus to an emotion network, ii) the assessment of speed versus accuracy measures, ii) the degree of emotional arousal.

ZUSAMMENFASSUNG

Hirnphysiologische Studien konnten wiederholt zeigen, dass emotionale Reize auf verschiedenen perzeptuellen Verarbeitungsstufen die neuronale Aktivität erhöhen. Demgegenüber weisen Leistungsmaße, die im Rahmen einfacher kognitiver Aufgaben erhoben wurden, eine große Variabilität hinsichtlich der leistungsfördernden oder -hemmenden Wirkung von emotionalen Reizeigenschaften auf. Mittels behavioraler (Reaktionszeit und Identifikationsgenauigkeit) und elektrophysiologischer Maße versucht die vorliegende Arbeit Bedingungen zu spezifizieren, die zu Einbußen oder Nutzen für die Aufgabenleistung führen, wenn ein Aufgabenreiz emotional gefärbt ist.

Es wird aufgezeigt, dass aversiver Bildinhalt zu einer Verlangsamung der visuellen Informationsverarbeitung führen kann. Umgekehrt wurden angenehme Bilder in einer einfachen Kategorisierungsaufgabe schneller erkannt, insbesondere wenn sie geringe emotionale Erregung auslösten. Anhand von Amplituden- und Latenzmaßen des ereigniskorrelierten Hirnpotentials ließen sich verschiedene elektrophysiologische Korrelate dieser Verhaltenseffekte aufzeigen. Einen hemmenden Einfluss emotionalen Bildmaterials konnte auf die Verarbeitung nachfolgender Information nachgewiesen werden. So führten hoch erregende Bildreize zu einer nachhaltigen Verschlechterung bei einer lexikalischen Entscheidungsaufgabe. Dies ging einher mit einer Amplitudenreduktion spezifischer Komponenten des ereigniskorrelierten Potentials, die im Zusammenhang mit lexiko-semantischen Prozessen stehen. Eine Projektion der Oberflächenpotentiale auf regionale, kortikale Quellen deutete auf eine Unterdrückung neuronaler Aktivität in posterioren und anterioren Bereichen des ventralen Verarbeitungsstromes hin. Eine Beeinträchtigung der Verarbeitung nachfolgender Information lösten emotionale Reize (Bilder oder Wörter) auch aus, wenn sie innerhalb eines Rapid Serial Visual Presentation (RSVP) Paradigmas dargeboten werden. In der vierten Versuchsreihe schließlich wurde der Zusammenhang zwischen emotionalem Gehalt und Aufgabenleistung im Hinblick auf multimodale Wahrnehmungsaspekte beleuchtet. Mit Hilfe der McGurk-Illusion konnte gezeigt werden, dass affektive Mimik und Prosodie audiovisuelle Integrationsprozesse erleichtern kann.

Ein Modell wird skizziert, welches die Leistung in perzeptuellen/kognitiven Aufgaben mit affektivem Reizmaterial ausgehend von drei grundlegenden Dimensionen vorhersagt: i) Der Zielreiz ist emotionales neuronales Netzwerk eingebunden versus aktiviert ein eigenes Netwerk, ii) die Aufgabe erfasst die Zeit versus die Genauigkeit der Verarbeitung als abhängige Variable, iii) der Zielreiz löst geringe versus hohe emotionale Erregung aus.