Psychophysiology, 45 (2008), 130–140. Blackwell Publishing Inc. Printed in the USA.

DOI: 10.1111/j.1469-8986.2007.00606.x

The coupling of emotion and cognition in the eye: Introducing the pupil old/new effect

MELISSA L.-H. VÕ,[a] ARTHUR M. JACOBS,[b] LARS KUCHINKE,[b] MARKUS HOFMANN,[b]
MARKUS CONRAD,[b] ANNEKATHRIN SCHACHT,[c] AND FLORIAN HUTZLER[d]

[a]Department of Psychology, Ludwig-Maximilians-Universität München, Munich, Germany
[b]Department of Psychology, Freie Universität Berlin, Berlin, Germany
[c]Department of Psychology, Humboldt Universität Berlin, Berlin, Germany
[d]Faculty of Psychology, University of Vienna, Vienna, Austria

Abstract

The study presented here investigated the effects of emotional valence on the memory for words by assessing both memory performance and pupillary responses during a recognition memory task. Participants had to make speeded judgments on whether a word presented in the test phase of the experiment had already been presented ("old") or not ("new"). An emotion-induced recognition bias was observed: Words with emotional content not only produced a higher amount of hits, but also elicited more false alarms than neutral words. Further, we found a distinct pupil old/new effect characterized as an elevated pupillary response to hits as opposed to correct rejections. Interestingly, this pupil old/new effect was clearly diminished for emotional words. We therefore argue that the pupil old/new effect is not only able to mirror memory retrieval processes, but also reflects modulation by an emotion-induced recognition bias.

Descriptors: Emotion, Cognition, Memory, Old/new effect, Pupillary response

The current article was inspired by studies showing a twofold influence of emotional processing on memory task performance: Although it has been shown that memory for emotional items is enhanced as compared to items that lack emotional content, the same emotional content seems to elicit more false memories as well (e.g., Maratos, Allan, & Rugg, 2000; McNeely, Dywan, & Segalowitz, 2004; Windmann & Kutas, 2001). The study reported here went a step further by investigating whether or not the influence of emotional content on memory processes is also mirrored in pupillary responses. This seems to be a rather self-evident approach considering the long history of pupillary response research on both cognitive and emotional processing. However, the *interaction* of both processes has been largely neglected. To bridge this gap, we conducted a recognition memory task using words of varying emotional content and recorded pupillary responses throughout the entire experiment. In the following, we will give an overview of the relevant findings regarding the influence of emotional content on subsequent memory task performance, followed by a summary of reported modulations of the pupillary response due to either emotional or cognitive processing.

Numerous studies have shown a strong influence of emotion on memory processes (e.g., Dietrich et al., 2000; Maratos et al., 2000; Ochsner, 2000; Richardson, Strange, & Dolan, 2004; Taylor et al., 1998; Windmann, Daum, & Güntürkün, 2002; Windmann & Kutas, 2001; for a review, see Phelps, 2006). A common finding is memory enhancement for those stimuli presented during the study phase of a recognition paradigm that are characterized by an emotional content (e.g., Bradley, Greenwald, Petry, & Lang, 1992; Kensinger & Corkin, 2003). This can be due to modulatory effects of emotion on encoding processes, on subsequent consolidation, or on later retrieval and even post-retrieval processes (for a review, see Hamann, 2001). Activation of the amygdala and the medial temporal lobe memory regions during the encoding of emotional stimuli have shown to be predictive of later memory performance for these stimuli (for a review, see LaBar & Cabeza, 2006). A commonly used approach to the study of memory encoding—the so-called subsequent memory paradigm—provides measures of neural activity correlated with later remembering by contrasting neural responses to stimuli later remembered to those later forgotten. This $\underline{d}$ifferential neural activity based on $\underline{m}$emory is also referred to as *Dm* (e.g., Paller & Wagner, 2002). Event-related fMRI studies have provided support for the close coupling of amygdala activity during encoding and delayed retention accuracy for emotional pictures (e.g., Dolcos, LaBar, & Cabeza, 2004a, 2004b; Kensinger & Corkin, 2004). There is additional evidence that—regardless of any effects on encoding or consolida-

---

This research was partly supported by a grant of the Deutsche Forschungsgemeinschaft (Ja 823/4-1/Jacobs "Kognitiv-affektive Konfliktkontrolle beim impliziten und expliziten Wiedererkennen") as a subproject of the DFG-Forschergruppe "Zwischen Interferenz und Optimierung: Konflikte als Signale in kognitiven Systemen". We the students of the Katholische Universität Eichstätt-Ingolstadt and the Freie Universität Berlin for participating in this study.

Address reprint requests to: Melissa Le-Hoa Võ, Ludwig-Maximilians-Universität München, Department of Psychology, Leopoldstraße 13, 80802 Munich, Germany. E-mail: lehoa@psy.uni-muenchen.de

tion—emotion also influences retrieval processes (e.g., Dolan, Lane, Chua, & Fletcher, 2000; Dolcos, LaBar, & Cabeza, 2005; Maratos, Dolan, Morris, Henson, & Rugg, 2001). Maratos and colleagues (2001) as well as Dolcos and colleagues (2005) found evidence that retrieval of emotional as opposed to neutral information was associated with enhanced activity in brain regions, which are known to be specialized in the processing of emotion (e.g., the amygdala). However, this enhanced activity for emotional stimuli during retrieval can lead to impaired memory performance when the previously encountered stimuli are joined by new, equally emotional stimuli during test.

For example, in a recognition memory task carried out by Maratos and colleagues (2000), the behavioral data showed that discrimination between words that had been presented before (i.e., "old" words) and words presented for the first time (i.e., "new" words) was significantly higher for neutral items than for negative ones. Further, the response bias also differed across neutral and negative items, such that participants showed a tendency to respond "old" to negative words more often than to neutral ones. Both hit rates (a hit refers to correctly judging an old item as old) and false alarm rates (a false alarm refers to falsely judging a new item as old) were significantly increased for negative words. This finding is consistent with other studies reporting that, although emotional stimuli are often better remembered, they seem to increase the number of false alarms as well (e.g., Maratos et al., 2000; McNeely et al., 2004; Windmann & Kutas, 2001). Windmann and colleagues have repeatedly referred to this finding as an *emotion-induced recognition bias* (Windmann & Krüger, 1998; Windmann & Kutas, 2001; Windmann, Sakhavat, & Kutas, 2002) arguing that even though this response pattern does not improve participants' accuracy scores, it does ensure that memories for emotional events are not erroneously considered irrelevant.

Strongly related to these findings are indices of a close coupling of emotional and cognitive processes during a memory task as observed in the modulation of so-called ERP old/new effects. ERP old/new effects refer to the comparison of event-related potentials (ERPs) elicited by stimuli correctly judged as old (hits) and ERPs elicited by stimuli correctly judged as new (correct rejections). It has repeatedly been shown that ERPs elicited by words correctly judged as old are characterized by a greater positivity than ERPs elicited by words correctly judged as new (e.g., Allan & Rugg, 1997; Inaba, Nomura, & Ohira, 2005; Johnson, 1995; for reviews, see Rugg, 1995; Rugg & Curran, 2007).

Dual-process models propose that recognition memory is supported by distinct retrieval processes known as familiarity and recollection (e.g., Yonelinas, 2001, for a review, see Yonelinas, 2002). Accordingly, ERP studies have identified two topographically distinct correlates of recognition memory—the "partietal" and the "mid-frontal" old/new effects. Whereas the parietal ERP old/new effect signifies the engagement of neural systems supporting episodic retrieval, the frontal effects reflect neural activity that comes into play when retrieved information has to be monitored (e.g., Maratos et al., 2000). These ERP old/new effects are dissociated by variables that selectively modulate recollection and familiarity, respectively (e.g., Curran, 2000; for a review, see Rugg & Curran, 2007). Although familiarity-based recognition is supposed to be supported by an undifferentiated, strengthlike form of information related to a signal-detection process, recollection relies on a second, functionally distinct memory signal that results from the retrieval of qualitative information about the study phase reflecting a thresholdlike retrieval process. Thus, in a recognition paradigm decisions on new items that have not been presented during study should mainly be based on familiarity processes, whereas decisions on old items should be additionally accompanied by recollection processes.

The ERP old/new effect has proven to be a suitable platform for investigating the influence of emotional content on memory processes by studying the relationship between neural activity and memory performance during retrieval. Maratos and colleagues (2000), for example, supported their behavioral data by finding that the emotional valence of words had a significant influence on the ERP old/new effect. They were able to show that the old/new effect was greater for neutral than for negative items. This difference was largely due to the enhanced positivity of the ERPs for negative new items, which diminished the old/new effects for this valence category. One explanation for the enhanced positivity to negative new items is that these negative items may elicit illusory episodic memory traces due to an enhanced but misleading feeling of familiarity. Because these illusory memory traces are not caused by true recollection, correct decision is hampered, as marked by higher false alarm rates. Similarly, Windmann, Sakhavat, and colleagues (2002) found significantly smaller early frontal ERP old/new effects for negative words in controls starting at around 300 ms after stimulus onset, which resembled the participants' tendency to classify negative words as old more often than neutral ones. Timing of the early ERP differences suggests that they are related to automatic memory and familiarity processes more than to consciously controlled memory. Further, Windmann and colleagues argue that the emotion-induced recognition bias is acting early on memory retrieval processes rather than on postretrieval response verification processes (Windmann, Urbach, & Kutas, 2002).

To summarize the existing results, there is ample evidence that the emotional content of stimuli affects memory by means of encoding and retrieval processes. Regarding the latter, especially early, familiarity based retrieval processes might be sensitive to emotional modulation. Obviously, most of the aforementioned studies have shown the close coupling of emotional and cognitive processes by using brain imaging techniques. The findings of an interaction between cognitive and emotional processes as seen in the modulations of ERPs old/new effects have led us to investigate whether or not similar patterns can be mirrored in the pupillary response as well.

The pupillary response has a long history of research dedicated to either emotional or cognitive processing. Ever since the reintroduction of the pupillary response into psychological research (Hess & Polt, 1960), two main interpretations of the pupillary response have been proposed: After the pupillary response was first related to emotionality (Hess & Polt, 1960), it was later suggested that it might instead reflect cognitive activity (Hess & Polt, 1964; Kahneman & Beatty, 1966). More recent studies have resumed emphasizing the use of pupillometric measures in order to assess psychological processes due to the enhanced spatial and temporal resolution provided by the latest eye trackers as well as the relatively easy assessment of pupillary responses as compared to ERPs (for a review, see Granholm & Steinhauer, 2004). The pupil starts to dilate within the first few hundred milliseconds after the onset of a cognitive demand. Although it is regarded as a key finding that pupils dilate with increasing cognitive load (e.g., Beatty & Kahneman, 1966; Karatekin, 2004; Verney, Granholm, & Marshall, 2004), the contribution of emotional processing to the pupillary response has been a matter of debate.

Early works by Hess (1965) reported pupil dilations when subjects looked at pictures of positive valence and pupil constrictions when they looked at pictures of negative valence (see also Mudd, Conway, & Schindler, 1990). However, other studies found that the pupil dilated to emotional stimuli regardless of their actual valence compared to neutral stimuli (e.g., Janisse, 1974; Partala & Surakka, 2003; Steinhauer, Boller, Zubin, & Pearlman, 1983). These contradicting results might have been due to varying degrees of experimental control applied to the studies. There was not only a great difference in the quality of stimulus presentation and data recording, but also in the quality of the selected stimulus material. Hess, for example, who presented pictures, was later criticized for not taking enough precautions to control for stimulus luminance, and Janisse, using word material, had controlled for luminance but not for word recognition variables such as word frequency, number of letters and so on.

Regarding emotional processing in general, there is an ongoing debate on the specific contributions of mainly two variables: emotional valence—from pleasure to displeasure—and arousal—from calm to aroused—spanning a valence/arousal space. Bradley and Lang proposed a psychological model of emotion in which emotional stimuli are organized along the bipolar dimensions of valence and arousal (for a review, see Bradley & Lang, 2000). The relative contributions to the pupillary response of valence and arousal on the one hand and cognitive load on the other are not yet fully understood.

Janisse (1974), for example, suggested that the relationship between pupillary response and emotional valence seems to be curvilinear; that is, pupils dilate most to negative and positive items and least to neutral ones, whereas the relationship between pupillary response and arousal was linear, that is, the pupillary response increases with increasing arousal levels. However, note that in Janisse's study, arousal ratings were not separately collected but were inferred from the set of valence ratings; that is, valence ratings close to neutral were considered to indicate low arousal, whereas more extreme valence ratings were classified as highly arousing. Partala and Surakka (2003) investigated pupil size variation during and after auditory emotional stimulation. In this more recent study, sounds were selected from an already existing database (International Affective Digitized Sounds, IADS) and matched for valence and arousal separately. Results showed that pupils dilated more to highly arousing negative and positive auditory stimulation as opposed to neutral stimulation.

Earlier, Stanners, Coulter, Sweet, and Murphy (1979) had investigated whether the pupillary response was more an indicator of arousal or cognition. In their study, they used tasks that employed either both arousal (threat of shock) and cognitive factors (mental arithmetic) or only arousal manipulations with no explicit cognitive demands. Interestingly, when the cognitive demands of the task were held constant, the pupillary response showed no influence of the arousal manipulations. However, when the cognitive demands of the task were manipulated, this was clearly indicated by the pupillary response. The pupillary response only showed an effect of arousal when cognitive demands were minimal. According to Stanners and colleagues, the control system of the pupil is such that if the situation requires a substantial level of cognitive activity, predominantly cognitive processing is indicated by the pupillary response.

Investigating the influence of emotional valence on word perception, Kuchinke, Võ, Hofmann, and Jacobs (2007) used a lexical decision task, where participants had to decide whether a letter string formed a word or a nonword. Although response times (RTs) and error rates showed clear emotional valence effects, effects of emotional valence on pupillary responses were not observed. Only word frequency significantly modulated the pupillary response. Although RTs and error rates reflect speed and accuracy, the pupillary response seems to be more a measure of cognitive resources required by the task (Nuthmann & van der Meer, 2005). Emotional valence might not influence the processing demands in a lexical decision task, but it possibly influences the temporal component of lexical access. A key difference between lexical decision and recognition memory tasks might be that, although both tasks involve memory processes, these are more implicit for lexical decision than the explicit memory processes demanded during a recognition memory task.

Siegle (1999) developed a computational neural network model of affective information processing that specifically accounts for affective information processing of words. In this model, word representations are fully connected to and feed activation forward to nodes representing the nonaffective features of stimuli and to nodes representing the affective content of stimuli, in parallel. Both affective and nonaffective features are connected through a feedback mechanism in order to account for the interaction of emotional and cognitive processing. According to Siegle's model, emotional word representations are characterized by an elevated activation pattern due to the input of affective nodes. Words of neutral content lack this extra activation. Thus, the presentation of an emotional word should not only lead to emotional processing, but the emotional content should influence cognitive processing as well. As we will see, there are indices in the pupillary response that mirror the close relationship of these processes.

The aim of this study was to find correlates of the coupling of emotional and cognitive processes in the pupillary response. Therefore, we recorded pupillary responses during a recognition memory task in which participants were first presented with a series of words and, after a short delay, were tested with lists that included these old words randomly intermixed with new words. Additionally, we manipulated emotional processing by using words of varying emotional content while not explicitly instructing participants to pay attention to the emotional content of the presented words. A great number of highly controlled word-based stimulus material was taken from the Berlin Affective Word List (BAWL; Võ, Jacobs, & Conrad, 2006), a corpus of more than 2000 words that contains not only values for the most influential word processing variables, but also a listing of emotional valence and imageability ratings.

First of all, we expected higher hit and false alarm rates for negative and positive words as predicted by the finding of an emotion-induced recognition bias, which characterizes the tendency of participants to classify emotional words as old more often than is the case for neutral words (e.g., Windmann & Kutas, 2001; Windmann, Sakhavat, et al., 2002). Further, correctly classifying emotional new words should yield longer RTs than correctly classifying neutral new words, because the emotion-induced recognition bias has to be overcome to produce correct responses to emotional new words. Accordingly, correctly classifying emotional old words should yield shorter RTs than correctly classifying neutral old words, because the emotion-induced recognition bias should facilitate the correct classification of emotional old words.

Regarding the investigation of the pupillary response, our methods were derived from a series of functional imaging and ERP studies reporting a modulation of memory processes by the emotional content of stimuli during encoding (e.g., Kensinger & Corkin, 2004) or retrieval (e.g., Windmann, Sakhavat, et al., 2002) or both (e.g., Dolcos et al., 2004a, 2004b, 2005). To investigate effects of emotional valence during *encoding*, we compared pupillary responses during the study phase of the experiment as a function of whether the words were later correctly remembered (hits) or not (misses), as has been done in subsequent memory paradigms. These studies have reported greater neural responses to later remembered than forgotten items (see Paller & Wagner, 2002). Translating these findings to the pupillary response, we expected to find a differential memory effect characterized as elevated pupillary responses to words that were later remembered than to words that were forgotten due to the deployment of more cognitive effort for successful elaboration during encoding. Further, we expected to find a modulation of the differential memory effect as a function of emotional valence: Later correctly remembered negative and positive words should exhibit smaller pupillary responses than correctly remembered neutral words, because emotional stimuli have a processing advantage compared to neutral stimuli; that is, less cognitive load should be necessary to successfully encode emotional words during study.

Additionally, we expected to find effects of emotional valence on *retrieval* processes. Similar to our hypotheses regarding the pupillary response during encoding processes, our hypotheses regarding retrieval processes were also deduced from studies that have investigated the coupling of emotional and cognitive processes using fMRI or ERP paradigms because, to our knowledge, there are no other studies that have tried to investigate this close coupling using pupillometric measures. Thus, we expected to find an old/new effect in pupillary responses during the test phase characterized as a greater pupillary response to correctly classified old words as opposed to correctly classified new words. According to dual-process models, items that have been encountered during the study phase can be retrieved by a combination of both familiarity and recollection processes, whereas items presented for the first time cannot be recollected. Because recollection is conceived as a slower, more demanding process that gives rise to consciously accessible information about prior occurrence of the test item (see Rugg & Curran, 2007), pupils should dilate to a greater degree to correctly classified old words than to new words. We further expected to find a diminished old/new effect for negative and positive words, similar to the findings reported for the ERP old/new effect (Maratos et al., 2000; Windmann, Sakhavat, et al., 2002) mirroring the twofold influence of emotional processing on memory performance.

Method

Participants

Nineteen psychology students (11 female) of the Freie Universität Berlin ranging in age from 21 to 30 years ($M = 22.94$, $SD = 2.48$) participated in the study to partially fulfill course requirements. All reported normal or corrected-to-normal vision, were native German speakers, and reported no history of neurologic and affective disorders. The participants were not familiar with the stimulus material used in this study and were not told that emotional valence was the focus of investigation.

Stimulus Material

The Berlin Affective Word List. The BAWL (Võ et al., 2006) is a corpus of over 2000 words characterized by the following variables taken from the CELEX data base (Baayen, Piepenbrock, & van Rijn, 1993)[1]: The number of letters, number of syllables, the number of phonemes, the total frequency of appearance per one million words, the number of orthographic neighbors, and the number of higher frequency orthographic neighbors. Additionally, for every word, emotional valence was rated on a 7-point scale ranging from -3 (*very negative*) over 0 (*neutral*) to $+3$ (*very positive*) by a total of 48 students (42 female) at Katholische Universität Eichstätt-Ingolstadt and Freie Universität Berlin. Another 40 students at these institutions rated imageability of the words on a 7-point scale ranging from 1 (*low imageability*) to 7 (*high imageability*). These ratings are the core of the BAWL.

Old words. From the BAWL, a subset of 180 words was selected to construct three distinct emotional valence categories: negative, neutral, and positive. These words were chosen according to the following criteria: (1) a mean rating of emotional valence, which had to fall into one of the three emotional valence categories: negative (mean emotional valence rating < -1.3), neutral ($-0.8 <$ mean emotional valence rating < 0.8), or positive (mean emotional valence rating > 1.3); (2) small standard deviations of ratings; (3) same number of verbs and nouns within each valence category; and (4) ambiguous words were excluded or controlled for. After careful matching, 60 negative, 60 neutral, and 60 positive words were finally chosen as experimental items not differing significantly across the three emotional valence categories according to their mean frequency, number of letters, number of syllables, the number of orthographic neighbors, the number of higher frequency orthographic neighbors, and mean imageability rating (see Table 1). Effort was also taken to make negative words as "negative" as positive words were "positive"; that is, the mean ratings of negative words and positive words were equidistant from the neutral value 0, -1.84 and $+1.84$, respectively, whereas neutral words had a mean valence rating of 0.02. Finally, words with extreme valence ratings were excluded from the stimulus set, because extremely rated words were regarded as also being potentially arousing.

New words. The new words were chosen from the BAWL in the same way as the old words. Again, neither mean frequency, number of letters, number of syllables, number of orthographic neighbors, number of higher frequency orthographic neighbors, nor mean imageability ratings differed significantly between negative, neutral, or positive new words (see Table 1). Further,

---

[1] In the BAWL, word length ranges from 3 to 10 letters ($M = 6.42$, $SD = 1.58$) and 1 to 4 syllables ($M = 2.18$, $SD = 0.67$), the number of phonemes varies from 2 to 10 ($M = 5.58$, $SD = 1.50$), and the total frequency of appearance per one million words ranges from low-frequency words to high-frequency words ($M = 62.99$, $SD = 164.26$). The list also contains information on the number and frequency of orthographic neighbors ($M = 1.68$, $SD = 2.27$, and $M = 191.92$, $SD = 1205.30$, respectively), as well as on the number and frequency of higher frequency orthographic neighbors ($M = 0.50$, $SD = 1.08$, and $M = 175.78$, $SD = 1199.86$, respectively).

Table 1. *Mean Values (Standard Deviations) of Variables across Valence Categories for Old and New Words*

	Negative	Neutral	Positive	F	p
Ftot/1Mio					
Old words	30.49 (29.35)	34.07 (34.85)	34.32 (33.82)	0.26	.77
New words	30.08 (29.32)	32.96 (26.34)	36.59 (28.95)	0.80	.45
Syllables					
Old words	2.15 (0.55)	2.25 (0.57)	2.17 (0.46)	0.61	.54
New words	2.27 (0.71)	2.37 (0.61)	2.27 (0.66)	0.46	.64
Letters					
Old words	6.98 (1.36)	6.92 (1.23)	6.75 (1.36)	0.50	.61
New words	7.07 (1.33)	7.03 (1.47)	6.85 (1.46)	0.40	.67
Phonemes					
Old words	5.84 (1.38)	5.76 (1.21)	5.88 (1.26)	0.09	.91
New words	6.07 (1.69)	5.95 (1.70)	5.98 (1.65)	0.05	94
N					
Old words	1.30 (1.36)	1.33 (1.57)	1.35 (1.39)	0.02	.98
New words	1.48 (1.61)	1.45 (1.70)	1.65 (1.88)	0.03	.80
HFN					
Old words	0.35 (0.71)	0.33 (0.71)	0.45 (0.70)	0.48	.61
New words	0.35 (0.63)	0.32 (0.75)	0.33 (0.68)	0.04	.97
Valence rating					
Old words	– 1.84 (0.33)	0.02 (0.44)	1.84 (0.38)	1356.70	.00
New words	1.83 (0.38)	0.01 (0.45)	1.82 (0.38)	2592.34	.00
Arousal rating					
Old words	3.51 (0.36)	2.48 (0.46)	2.93 (0.49)	133.93	.00
New words	3.73 (0.45)	2.47 (0.47)	2.37 (0.51)	152.36	.00
Imageability rating					
Old words	4.30 (1.04)	4.31 (1.40)	4.45 (1.58)	0.22	.81
New words	4.28 (1.04)	4.21 (1.44)	4.59 (1.29)	1.53	.22

Note. Old and new words were matched across valence categories for frequency per one million words (Ftot/1Mio), number of syllables, number of letters, number of phonemes, number of orthographic neighbors (N), number of higher frequency orthographic neighbors (HFN), emotional valence ratings, and imageability ratings. Imageablity, valence, and arousal ratings were all taken from different samples.

the means of all these variables did not differ between old and new words (with F ranging from 0 to 3.12).

Arousal ratings. Because the BAWL in its current version does not contain arousal ratings, these had to be collected for the 360 words used in our study by subsequently conducting an arousal rating study. For this purpose, 30 students (27 female) of the Katholische Universität Eichstätt-Ingolstadt rated the relevant words using a paper–pencil version of the Self-Assessment Manikin (SAM) rating system (Lang, 1980) on a 5-point scale with arousal represented graphically by changes in a cartoon figure. The arousal ratings showed higher arousal values for the negative valence category, whereas positive and neutral categories did not significantly differ in their mean arousal values (see Table 1).

Apparatus

Pupillometric measures were recorded with a video-based iView X Hi-Speed eye tracker (SensoMotoric Instruments, Germany) with a sampling rate of 250 Hz connected to a Pentium IV IBM compatible computer. An infrared sensitive camera recorded pupil diameters from the left eye. Participants seated themselves in an adjustable, comfortable chair in a quiet room. Their heads were stabilized in a chin rest. The experimental sessions were carried out on an IBM compatible computer running on OS Windows 2000. All stimuli were presented in Courier 24 point type font on a 17-in. computer screen (resolution 1024 × 768 pixel, 85 Hz) subtending a vertical visual angle of 0.92° (0.8 cm in height). Because words were 3 to 10 letters long, they subtended a horizontal visual angle ranging from 1.72° to 5.72° (0.5 cm letter width). Black words were printed on gray background (RGB: 150, 150, 150) to minimize differences in luminance during stimulus presentation. Stimulus presentation and reaction recording was controlled by Presentation 9.0 Software (Neurobehavioral Systems, Inc., Albany, Canada).

Procedure

The experimental sessions were conducted in a moderately lit room (background luminance about 500 lx), in which the illumination was held constant. Each participant received written instructions before being seated in front of the presentation screen: Participants were informed that they would be presented with a series of words during the study phase of the experiment with the instruction to remember as many of them as possible. They would then be presented with a second series of words during the test phase having to indicate—as fast and as accurately as possible—whether a word had been presented during study ("old") or if it was presented for the first time ("new"). However, the participants were not informed about the emotional valence manipulation. Participants wearing glasses or contact lenses had no difficulties in adjusting to the iView X Highspeed System. To avoid artifacts due to misunderstanding the instructions, head movements, or blinking, participants were given five training trials to get accustomed to the apparatus and especially to the structure of the experimental trials, which contained marked periods during which blinking was allowed. Both the study and the test phase were preceded by a 5-point calibration to ensure that the participant's eye was correctly tracked by the iView X Highspeed System. The recording of pupillary responses started with successfully completing the first 5-point calibration and lasted throughout the experiment. Stimulus-locked segments of 2000 ms in length were marked for further analysis starting 200 ms before stimulus onset and ending 1800 ms after stimulus onset. The study phase consisted of a randomized word-by-word presentation of the 180 old words. To compensate for primacy and recency effects, five buffer words were added at the beginning and the end of the study phase. These words were subsequently excluded from further analysis. Each trial was initiated by the appearance of a fixation cross "+" centered on a screen (1000 ms). This fixation cross was immediately replaced by a word (1000 ms). Following a blank screen (800 ms), a row of asterisks "**********" indicated the blinking period during which the participants were allowed to blink. This procedure minimized blinking during experimental trials, keeping the number of excluded trials as low as possible. Participants initiated the next trial by pressing the space bar of the keyboard with one hand while giving "old"/"new" responses with the other using a mouse, thus preventing a bias toward either the left or the right mouse button used for collecting reaction time data. The study phase lasted for about 15 min. Before starting the test phase, nonwords—pronounceable strings of letters—were presented, which had to be rated on their word likeness. This intermission served as a disturbance of rehearsal by engaging participants in a speech-based task. The test phase consisted of the 180 old words randomly intermixed with 180 new words. Again, each trial was initiated by displaying a fixation cross (1000 ms), followed by the test stimulus (500 ms). Participants were instructed to decide as quickly and accurately as possible whether the test stimulus had been seen during the study phase (old words) or not (new words) by pressing the left or the right mouse button. After the presentation of the stimulus the screen went blank (1300 ms) followed by the appearance of the asterisks to indicate that blinking was allowed. RTs were measured from stimulus presentation until button press. The test phase lasted approximately 30 min.

Data Reduction and Statistical Analysis

Reaction times were calculated from stimulus onset until button press. For the preparation of reaction time data, outliers were defined as incorrect responses and responses with a latency deviating more than two standard deviations from within-participant mean. Outliers were subsequently excluded from the reaction time analysis (4%).

Further, measures of discriminability (d') and decision criterion (C) were calculated according to Snodgrass and Corwin (1988) as $d' = z(\mathrm{FA}) - z(\mathrm{HIT})$ and $C = 0.5^* (z(\mathrm{FA})+z(\mathrm{HIT}))$. Accordingly, discriminability increases with increasing d'. Negative C values indicate more liberal decision criterion, whereas positive C values hint toward a more conservative one.

Pupil data were prepared and preanalyzed using Vision Analyzer 5.1. software (Brain Products GmbH, Germany) normally used for EEG data preparation. Stimulus-locked segments of 2000 ms in length (from 200 ms pre- to 1800 ms poststimulus onset) were analyzed for the 180 old words presented during study as well as for the 180 old and 180 new words presented during the test phase. Artifacts, including blinks, were identified as large changes in pupillary responses occurring too rapidly to signify actual dilation or contraction. Trials containing such artifacts were excluded from analysis (4.6%; ranging from 0% to 12.9% per subject). The baseline of the pupillary response (*SD*)—measured as the average dilation over a time window of 200 ms preceding stimulus onset—averaged to 3.95 mm (0.69) and was subtracted from pupillary responses after stimulus onset to gain comparable pupillary response indices. Peak horizontal dilations in a time windows of 1800 ms after stimulus onset were determined and submitted separately to ANOVAs with emotional valence (negative, neutral, positive) as the within-subject factor.

To focus comparisons and to prevent the need for multiple contrasts, we confined post hoc tests solely to comparisons of valence (negative vs. positive words) and of arousal (negative and positive vs. neutral words) for both behavioral and pupillary data. *P* values are reported as $p < .05$ or $p < .01$, when probabilities are below .05 or .01, respectively, and as $p > .05$ when probabilities failed to reach significance. *P* values are reported with exact probabilities when a strong tendency is observed but fails to reach statistical significance or when *p* values are presented in tables. Further, for effects with multiple degrees of freedom, *p* values were Greenhouse–Geisser adjusted.

Results

Behavioral Data

Reaction time data. Data was submitted to an ANOVA with emotional valence (negative, neutral, and positive) and old/new (old vs. new words) as within-subject factors. As can be seen in Table 2, there was a main effect of emotional valence on RT data, $F(2,17) = 4.53$, $MSE = 1505.99$, $p < .05$. Although RTs to negative words did not differ from RTs to positive words, $t(17) = 1.24$, $p > .05$, the RTs to negative and positive words were significantly longer than the RTs to neutral words, $t(17) = 2.47$, $p < .05$. Further, there was neither a main effect of old/new, $F(1,17) = 3.01$, $MSE = 6487.38$, $p > .05$, nor an interaction of emotional valence and old/new, $F(2,17) = 0.21$, $MSE = 4732.93$, $p > .05$. Thus, it took longer to correctly classify emotional than neutral words.

Overall error rates. The overall error rates did not differ across emotional valence categories, $F(2,17) = 0.61$, $MSE = 0.002$, $p > .05$ (see Table 2).

False alarm rates. As can be seen in Table 2, false alarm rates differed significantly across emotional valence categories, $F(2,17) = 9.85$, $MSE = 0.001$, $p < .01$. In accordance with our hypotheses, negative and positive new words produced significantly more false alarms than neutral new words, $t(17) = 3.81$, $p < .01$, whereas there was no significant difference between positive and negative new words, $t(17) = 1.57$, $p > .05$.

Hit rates. There was a main effect of emotional valence for hits, $F(2,17) = 17.49$, $MSE = 0.001$, $p < .01$ (see Table 2). Negative and positive words elicited more hits than neutral words, $t(17) = 5.68$, $p < .01$. Additionally, negative old words produced significantly more hits than positive old words, $t(17) = 2.73$, $p < .05$.

Thus, the consistent error rates across all valence categories was the result of both elevated hit and false alarm rates for negative and positive words as opposed to words that lack emotional content.

Discriminability d' and decision criterion C. There was no main effect of emotional valence for the discriminability d', $F(1,17) = 0.74$, $MSE = 0.06$, $p > .05$ (see Table 2). However, the decision criterion C differed significantly across emotional valence categories, $F(1,17) = 18.34$, $MSE = 0.02$, $p < .01$ (see Table 2). Taken together, negative and positive words elicited a signifi-

Table 2. *Means (Standard Errors) for Behavioral Data for Each Emotional Valence Category*

	Negative	Neutral	Positive	F	p
RTs (ms)				4.53	.02
Old words	859.56 (28.62)	842.68 (34.34)	857.20 (36.78)		
New words	898.02 (32.34)	860.68 (31.73)	881.46 (31.26)		
Overall errors (%)	0.34 (0.02)	0.35 (0.02)	0.35 (0.02)	0.61	.56
Hits rates (%)	0.35 (0.02)	0.28 (0.02)	031 (0.02)	17.49	.00
False alarm rates (%)	0.18 (0.02)	0.13 (0.01)	0.16 (0.02]	9.85	.00
Decision criterion C	– 0.06 (0.09)	0.25 (0.07)	0.07 (0.09)	18.34	.00
Discriminability d'	0.94 (0.11)	0.86 (0.13)	0.83 (0.11)	0.74	.48
Pupillary response during study (mm)	0.045 (0.01)	0.055 (0.01)	0.061 (0.01)	3.25	.06
Pupil old/new effect during test (mm)	0.027 (0.01)	0.062 (0.01)	0.041 (0.01)	54.89	.01

cantly less conservative decision criterion than neutral words, $t(1,17) = 5.91$, $p < .01$. Within the emotional valence categories, negative words showed a more liberal decision criterion than positive words, $t(1,17) = 2.34$, $p < .05$. Thus, although the discriminability d' across all valence categories did not differ, the data imply an emotion-induced decision bias toward a more liberal decision criterion for emotionally valent words as opposed to neutral words.

Pupil Data

Differential memory (Dm) effect. During the study phase, participants were instructed to read and memorize as many of the presented words as possible in order to make old/new recognition judgments in a subsequent test phase. Recorded pupillary responses were first categorized according to whether the corresponding word was later correctly recognized (hit) or not (miss) and then submitted to an ANOVA with emotional valence (negative, neutral, and positive) and differential memory (hits vs. misses) as within-subject factors.

There was neither an identifiable differential memory effect in pupillary responses, $F(1,17) = 0.53$, $MSE = 0.001$, $p > .05$, nor an interaction, $F(2,17) = 0.24$, $MSE < 0.001$, $p > .05$. However, pupillary responses showed a strong tendency to vary as a function of emotional valence categories, $F(2,17) = 3.25$, $MSE = 0.001$, $p = .06$. Although pupillary responses to negative and positive words showed no significant differences as opposed to neutral words, $t(17) = 0.43$, $p > .05$, pupillary responses to negative words were smaller than to positive words, $t(17) = 2.70$, $p < .05$.

Thus, possible differences between successful and erroneous encoding of the stimulus material in the study phase were not mirrored in the pupillary response, whereas the pupillary response was sensitive to specific emotional valence of negative and positive words.

Pupil old/new effects. During the test phase of the experiment, participants made decisions on whether a presented word had already been presented during the study phase (old) or not (new). We performed an ANOVA with emotional valence (negative, neutral, and positive) and old/new (hits vs. correct rejections) as within-subject factors. There was no main effect of emotional valence on pupillary responses during test, $F(2,17) = 0.45$, $MSE < 0.001$, $p > .05$. However, pupillary responses showed a distinct old/new effect, $F(1,17) = 54.89$, $MSE < 0.001$, $p < .01$, in that pupils dilated more to hits than to correct rejections. Further, the pupil old/new effect varied as a function of emotional valence categories, $F(2,17) = 5.23$, $MSE = 0.001$, $p < .05$ (see Table 2).

More specifically, the pupil old/new effect for the two emotional categories taken together was significantly diminished compared to the old/new effect elicited by neutral words, $t(17) = 3.41$, $p < .01$., whereas the size of the old/new effect did not differ between negative and positive words, $t(17) = 1.17$, $p > .05$. Figure 1 nicely depicts the modulation of the pupil old/new effect as a function of emotional valence revealing graded old/new effects across valence categories. Pupil dilations to neutral, positive, and negative content are inversed depending on whether they refer to old or new words.

To our knowledge, these findings of diminished old/new effects in pupil data are the first of their kind.

Discussion

The starting point of this study was to investigate the influence of emotional valence on explicit memory for words while aiming to show that this interaction can also be seen in the pupillary response. Thus, we used word-based stimulus material from the BAWL (Võ et al., 2006) in a recognition memory task throughout which pupillary responses were recorded. This led to three key findings which will be discussed subsequently in greater detail: First, memory performance was modulated by a more liberal decision criterion for emotional words as seen in both higher hit and higher false alarm rates for emotional words as opposed to neutral words, a result that has also been reported in other studies (e.g., Maratos et al., 2000; Mathews & Barch, 2004; McNeely

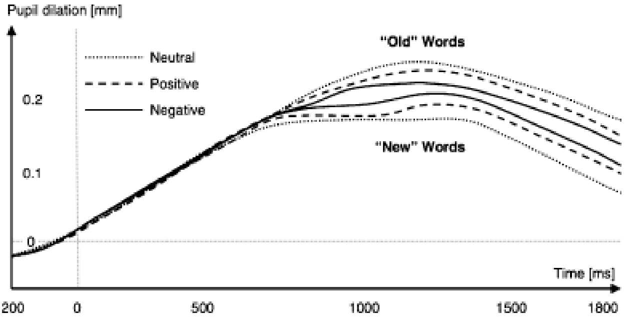

Figure 1. Mean pupil dilation curves to old and new words subdivided for valence categories (negative, neutral, positive) during the test phase of a recognition memory task. The three upper lines indicate pupil dilation curves for old words; the lower three indicate pupil dilation curves for new words. Note that although pupil dilations to neutral old words are greatest, followed by pupil dilations to positive and negative old words, this effect is reversed for new words.

et al., 2004; Windmann & Kutas, 2001). Second, although we did not find a memory effect in the pupillary response during encoding, we observed a distinct memory effect in the pupillary response during retrieval characterized by an elevated pupillary response when correctly classifying old words as old compared to a smaller pupillary response when correctly classifying new words as new. Thus, we were able to observe the basic pattern of the old/new effect in the pupillary response. We therefore term this effect the "pupil old/new effect," referring to findings in the ERP literature, where ERPs elicited by words correctly judged as old have repeatedly been observed to show a greater positive activation than ERPs elicited by words correctly judged as new (e.g., Allan & Rugg, 1997; Inaba et al., 2005; Johnson, 1995). Finally, we found that this pupil old/new effect was modulated by the emotional content of a presented word. The pupil old/new effect was greatest for neutral words and significantly diminished for emotional words.

Before discussing the data in greater detail, we first want to address the possible influence of arousal on the outcome of the results obtained in this study. Because the BAWL (Võ et al., 2006) does not yet contain arousal ratings for the listed words, we were not able to match our stimulus material according to the potential arousal levels of the words. However, we obtained the missing arousal ratings for the 360 words used in our experiment by subsequently conducting an arousal rating study. The arousal ratings showed higher arousal values for the negative valence category, whereas positive and neutral categories did not significantly differ in their mean arousal values. Despite the lack of a validated corpus of arousal ratings for emotional words, we will—with reservations—take the collected arousal data into account for the discussion of our data and hope that this particular topic will be addressed in future research.

In the following, we first discuss effects of emotional processing on memory performance and subsequently discuss possible relationships between these behavioral findings and the pupillary response.

Behavioral data showed a clear effect of emotional content on both hit and false alarm rates. Although negative and positive words were better remembered than neutral words, they also elicited more false alarms. Thus, our results indicate that the emotional content of words does not only increase correct classification of old words, but also increases the false classification of emotional new words. These opposing processes leveled out the number of errors across valence categories.

According to the computational neural network model of affective information processing by Siegle (1999), emotional word representations should be characterized by higher activation levels than neutral word representations due to extra activation from affective nodes. This additional activation could, on the one hand, support correct classification of emotional old words leading to more hits. On the other hand, the higher activation of representations for new words could be misinterpreted as activation stemming from old words leading to the false classification of emotional new words, that is, observed in increased false alarm rates.

Signal detection measures revealed that emotional words evoked more liberal decision criteria than neutral words, whereas the ability to discriminate between old and new items remained unaffected by the words' emotional content. Similarly, Windmann and colleagues argue that an emotion-induced recognition bias causes the elevated hit and false alarm rates for emotional words, whereas the discriminability is not affected by emotional valence (Windmann & Kutas, 2001; Windmann, Sakhavat, et al., 2002). Overcoming the emotion-induced recognition bias in order to correctly reject emotional new words requires more cognitive processing, resulting in a higher cognitive load.

Accordingly, we had predicted prolonged latencies for correct rejection of emotional words and shorter latencies for hits. Reaction time data only partly confirmed our predictions. Although correctly classifying emotional words in general resulted in prolonged response latencies, the expected interaction between old/new responses and valence manipulation was not visible. This could be due to the high amount of errors, which we had aimed for in order to calculate signal detection measures and in turn could have attenuated the effect.

Interestingly, we observed an effect of emotional valence on the decision criterion not just between emotional and neutral words, but in addition, negative words elicited a significantly more liberal decision criterion than positive words. One could argue that this valence effect is simply an effect of emotional arousal, because negative words were rated as more arousing than neutral or positive words. However, negative and positive words produced the same level of hit and false alarm rates even though negative words had been rated as more arousing. Thus, the effect observed cannot solely be attributable to the arousal levels of emotional words. Adding to this assumption, positive and neutral words produced significantly different hit and false alarm rates while not differing in their rated arousal levels. A possible explanation for an even more liberal decision criterion for negative words as compared to positive words could be that along the lines of Windmann, Sakhavat, and colleagues (2002), the high survival value of negative information might be greater than that of positive information. Therefore, the emotion-induced recognition bias seems to set the criterion even lower for negative words to make sure that no relevant, possibly harming information is missed.

In terms of the dual-process framework, our data suggest that at least familiarity processes could be susceptible to the emotional valence of the presented words. According to Yonelinas (2001, 2002) familiarity, not recollection, is influenced by shifts of the decision criterion. Thus, the increased false alarm rates for emotional words should be the result of a misleading feeling of familiarity for emotional new words. The same feeling of familiarity for emotional old words, however, correctly leads to increased hit rates.

Apart from investigating the effects of emotional valence on memory performance, we aimed at finding indices of a coupling of emotion and cognition in the pupillary response as well. By using a recognition memory paradigm, we were able to look at possible modulations of the pupillary response as a function of emotional valence effects on either encoding or retrieval.

For the investigation of interacting processes during *encoding*, we analyzed the pupillary data according to the subsequent memory paradigm (see Paller & Wagner, 2002) by first categorizing pupillary responses to words that were either later remembered or not. We had expected to find greater pupillary responses for words later remembered than for words later forgotten. However, contrary to findings using ERPs or event related fMRI (e.g., Dolcos et al. 2004a, 2004b; Kensinger & Corkin, 2004), there was neither a main effect of subsequent memory nor an interaction with emotional valence. A reason for the lack of a differential memory effect on the pupillary response might be that although memory processes were generally employing cognitive load during study, the differences in cognitive demands between the encoding of later remembered and later forgotten words were not sizable enough to modulate the pupil accordingly. As a result, emotional valence did not have a suitable platform to exhibit its modulatory effects.

Interestingly, there was a main effect of emotional valence with pupillary responses to negative words being significantly smaller than to positive words. Thus, contrary to studies that observed greatest pupil dilations to emotional stimuli (e.g., Janisse, 1974; Partala & Surakka, 2003), we observed a linear increase of the pupillary response across the valence categories from negative to positive.

A reason for this deviance of our data from other data could be that we employed a task that involved explicit memory processes. For example, in the studies of both Janisse (1974) and Partala and Surakka (2003), participants did not have to complete a cognitively demanding task. Rather, participants in these studies were either instructed to simply read emotional words out loud or to listen to the emotional tones that were presented, whereas participants in our study phase had to memorize a list of words for later recognition.

Assuming that the pupillary response is a summative index of the brain activity associated with cognitive and emotional processes (Beatty, 1982), the smaller pupillary response to negative words suggests that negative words were memorized with less effort than positive words during encoding irrespective of whether a word was later remembered or not. Windmann and Krüger (1998) found an emotion-induced recognition bias in lexical decisions as well arguing that words with negative content are processed preferentially due to subconscious threat detection influences. This could explain the relative ease with which negative words were encoded. Interestingly, the only valence-specific effects were observed for the decision bias, the pupillary response during encoding, and the amount of hits. All these valence effects point in the same direction; that is, negative words seem to be preferentially processed, resulting in smaller pupillary responses during encoding and a greater number of hits in the test phase. Siegle's (1999) model is generally able to account for valence effects by the implementation of two separate affective nodes for negative and positive valence, respectively. Our data imply that activation levels of negative affective nodes should be even higher than activation levels of positive affective nodes, enabling especially efficient cognitive processing of negative stimuli. However, the lack of differences between the pupillary response to neutral versus emotional words contradicts predictions of Siegle's (1999) model, according to which the processing of negative and positive words should differ from the processing of neutral words. This misfit could be due to the different cognitive demands posed on a participant during the study phase of a recognition memory task, in which no overt decisions have to be made like on the valence, the lexicality, or prior presentation of a stimulus.

Regarding *retrieval*, we had first of all expected to find a memory effect in the pupillary response between correctly classified old words versus correctly classified new words similar to the basic pattern of the ERP old/new effect. So far there has neither been any evidence for old/new effects in the pupillary response nor are we aware of reported modulatory effects of emotional valence on such pupil old/new effects. Therefore, our hypotheses were closely related to findings in the literature on ERP old/new effects. In analogy with findings of a greater positive activation in ERPs to correctly identified old words (hits) as compared to correctly identified new words (correct rejections) (e.g., Dietrich et al., 2000; Johansson, Mecklinger, & Treese, 2004; Maratos et al., 2000; Windmann & Kutas, 2001; Windmann, Sakhavat, et al., 2002; for reviews, see Rugg & Allan, 2000; Rugg & Curran, 2007), we had expected greater activation for hits and less activation for correct rejections.

Indeed, we found a significant old/new effect in pupil size characterized by a greater pupillary response to hits, whereas pupils dilated less to correct rejections. Maratos and colleagues (2000) have pointed out that the mere retrieval of episodic information was not a reliable guide to distinguish between old and new items. Rather, it is necessary to evaluate the content of the retrieved information in order to ascertain that it represented a veridical episode from the study phase. For the pupillary response, the greater pupil dilation to old words could therefore be due to the need of additional retrieval or postretrieval processes for the correct classification of old as opposed to new words. In terms of dual-process models (e.g., Yonelinas, 2001, 2002), items that are presented for the first time during the test phase cannot be recollected due to the lack of qualitative information about the study phase. Because recollection is conceived as a slower, more demanding process that gives rise to consciously accessible information about prior occurrence of the test item (see Rugg & Curran, 2007), this could have caused the pupils to dilate to a greater degree to correctly classified old words than to new words.

Although we are aware of the fact that ERP effects cannot be translated one to one into effects in the pupillary response, we do propose that this pupil old/new effect might be a valid marker for the successful discrimination between old and new words. This view is further supported by the finding of diminished pupil old/new effects for emotional words, which is in line with the twofold effect of the emotion-induced recognition bias on memory performance.

In studies examining ERP old/new effects, findings regarding the influence of emotional content have been equivocal: Although there is evidence that emotional valence is accompanied by increased old/new effects (e.g., Dietrich et al., 2000; Johansson et al., 2004), diminished old/new effects for emotional words are reported as well (e.g., Maratos et al., 2000; Windmann, Sakhavat, et al., 2002). Either way, the ERP old/new effect always corresponded to the behavioral outcome in that diminished old/new effects were accompanied by a decreased performance in the memory task. Our data clearly show diminished old/new effects in pupillary responses for words with emotional content, which coincides with the elevated hit and false alarm rates we found in behavioral data caused by more liberal decision criteria when subjects responded to emotional words. The diminished pupil old/new effect was based on graded pupillary responses to correctly classified old words such that pupils dilated most to neutral, less to positive, and least to negative words. This graded effect was exactly inverted for new words; that is, pupils dilated most to negative, less to positive, and least to neutral new words. As we have seen from behavioral data, the emotional content seems not only to facilitate generating hits, but also tempts participants to produce more false alarms. Pupillary responses mirror this twofold influence of emotion on memory processes by reflecting the cognitive demands necessary for responding correctly: Although the correct recognition of emotional old words seems to pose fewer cognitive demands on retrieval—as seen in a smaller pupillary response—the effort to correctly classify an emotional new word as new and to overcome the bias to respond "old" is much greater as compared to neutral new words—as seen in the inverted pupillary response pattern. Siegle (1999) has accounted for the close coupling of emotional and cognitive processing in his model by implementing strong feedback loops between affective and nonaffective nodes. Because their summed activity is supposed to directly correspond to the pupillary response, the interaction of affective and nonaffective nodes not only predicts differences in the processing of emotional and neutral information to be evident in performance, but predicts a modulation of the pupillary response as well.

The extent to which the pupil old/new effect can be compared to the ERP old/new effect and the extent to which their modulations by emotional processes share common ground will require further examination. However, the strong relationship between memory performance on the one hand and variations of the pupil old/new effect on the other seems to imply that the pupillary response is sensitive enough to index the modulatory effect of emotion on memory processes.

The study presented here raises a number of questions that should be addressed in future work in this field. As we have pointed out in the introduction, the relative influences of arousal, emotional valence, and cognitive load on the pupillary response are not yet clarified.

For example, it is not clear whether the emotional valence effect observed in the pupillary response during encoding was simply due to increased arousal levels of negative words or due to valence-specific biases in the processing of negative versus positive information. Partala and Surakka (2003), for example, showed that pupils dilated more to highly arousing negative and positive auditory stimulation as opposed to neutral nonarousing stimulation. Thus, the data provide evidence that the pupil dilates to highly arousing stimuli. However, the lack of effects of emotional valence in their study could be due to the selection of highly arousing stimuli, which might have attenuated possible valence effects. Earlier, Stanners and colleagues (1979) had investigated whether the pupillary response was more an indicator of arousal or cognitive processing. The pupillary response only showed an effect of arousal when cognitive demands were minimal. Because in our study the cognitive demands were quite high, effects of arousal might have been weaker than in other studies where arousal was explicitly investigated and manipulated to be either very high or low.

Another possible factor that could have led to an emotional valence effect is semantic cohesion; that is, emotional words could be more tightly connected in semantic networks than neutral words. In our study, we did not control for semantic cohesion, and even though we doubt that it could sufficiently explain the strong effect of emotional valence, we cannot fully exclude the possibility that semantic cohesion played a role in causing the effects we report here. In an ERP study, McNeely and colleagues (2004) tried to disentangle the influence of emotionality and semantic association by including a highly associated but emotionally neutral category (animals). It could be shown that emotionality and not semantic cohesion elicited greater positivity in the ERPs and that emotionality could not be dismissed on the basis of emotional items being more semantically related. Further, in a study by Windmann and Kutas (2001), strong effects of emotional valence were observed even though stimulus material had been controlled for semantic cohesion. This adds evidence to the view that the influence of emotional content on the memory for words cannot solely be attributed to the semantic cohesion of emotional words, either.

In sum, we have replicated findings of a twofold influence of emotional processing on memory performance caused by an emotion-induced recognition bias. Additionally, we were able to replicate findings of an emotional modulation of the ERP old/new effect by using the pupillary response as a different, less invasive method, which has a long history in investigating effects of either cognitive or emotional processing. Clearly, further research is necessary and the results presented here have to be replicated before the pupil old/new effect can act as a reliable platform to investigate the coupling of emotional and cognitive processing. However, we were able to show that the modulatory effects on memory performance due to emotional processing are closely resembled by corresponding modulations of the pupillary response. We hope that the introduction of the pupil old/new effect will prompt others to follow this line of research.

REFERENCES

Allan, K., & Rugg, M. D. (1997). An event-related potential study of explicit memory on tests of cued recall and recognition. *Neuropsychologia*, *35*, 387–397.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX Lexical Database [CD-ROM]. *Performance*, *18*, 1174–1188.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*, 276–292.

Beatty, J., & Kahneman, D. (1966). Pupillary changes in two memory tasks. *Psychonomic Science*, *5*, 371–372.

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*, 379–390.

Bradley, M. M., & Lang, P. J. (2000). Measuring emotion: Behavior, feeling and physiology. In R. Lane & L. Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 242–276). New York: Oxford University Press.

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory and Cognition*, *28*, 923–938.

Dietrich, D. E., Emrich, H. M., Waller, C., Wieringa, B. M., Johannes, S., & Münte, T. F. (2000). Emotion/cognition-coupling in word recognition memory of depressive patients: An event-related potential study. *Psychiatry Research*, *96*, 15–29.

Dolan, R. J., Lane, R., Chua, P., & Fletcher, P. (2000). Dissociable temporal lobe activations during emotional episodic memory retrieval. *NeuroImage*, *11*, 203–209.

Dolcos, F., LaBar, K. S., & Cabeza, R. (2004a). Dissociable effects of arousal and valence on prefrontal activity indexing emotional evaluation and subsequent memory: An event-related fMRI study. *NeuroImage*, *23*, 64–74.

Dolcos, F., LaBar, K. S., & Cabeza, R. (2004b). Interaction between the amygdala and the emdial temporal lobe memory system predicts better memory for emotional events. *Neuron*, *42*, 855–863.

Dolcos, F., LaBar, K. S., & Cabeza, R. (2005). Remembering one year later: Role of the amygdala and the medial temporal lobe memory system in retrieving emotional memories. *Proceedings of the National Academy of Sciences, USA*, *102*, 2626–2631.

Granholm, E., & Steinhauer, S. R. (2004). Introduction: Pupillometric measures of cognitive and emotional processing. *International Journal of Psychophysiology*, *52*, 1–6.

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences*, *5*, 394–400.

Hess, E. H. (1965). Attitude and pupil size. *Scientific American*, *212*, 46–54.

Hess, E. H., & Polt, J. M. (1960). Pupil size as related to interest value of visual stimuli. *Science*, *132*, 349–350.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem solving. *Science*, *143*, 1190–1192.

Inaba, M., Nomura, M., & Ohira, H. (2005). Neural evidence of effects of emotional valence on word recognition. *International Journal of Psychophysiology*, *57*, 165–173.

Janisse, M. P. (1974). Pupil size, affect and exposure frequency. *Social Behavior and Personality*, *2*, 125–146.

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

Johnson, R. (1995). Event-related potential insights into the neurobiology of memory systems. In J. C. Baron & J. Grafman (Eds.), *The handbook of neuropsychology* (Vol. 10, pp. 135–164). Amsterdam: Elsevier.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *145*, 1583–1585.

Karatekin, C. (2004). Development of attentional allocation in the dual task paradigm. *International Journal of Psychophysiology*, *52*, 7–22.

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, *31*, 1169–1180.

Kensinger, E. A., & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences, USA*, *101*, 3310–3315.

Kuchinke, L., Võ, M. L.-H., Hofmann, M., & Jacobs, A. (2007). Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *International Journal of Psychophysiology*, *65*, 132–140.

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience*, *7*, 54–64.

Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, J. H. Johnson, & T. A. Williams (Eds.), *Technology in mental health care delivery systems* (pp. 119–137). Norwood, NJ: Ablex.

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

Maratos, E. J., Dolan, R. J., Morris, J. S., Henson, R. N. A., & Rugg, M. D. (2001). Neural activity associated with episodic memory for emotional context. *Neuropsychologia*, *39*, 910–920.

Mathews, J. R., & Barch, D. M. (2004). Episodic memory for emotional and nonemotional words in schizophrenia. *Cognition and Emotion*, *18*, 721–740.

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

Mudd, S., Conway, C. G., & Schindler, D. E. (1990). The eye as music critic: Pupil response and verbal preferences. *Studia Psychologica*, *32*, 23–30.

Nuthmann, A., & van der Meer, E. (2005). Time's arrow and pupillary response. *Psychophysiology*, *42*, 306–317.

Ochsner, K. N. (2000). Are effective events richly recollected or simply familiar? The experience and process of recognizing feelings past. *Journal of Experimental Psychology: General*, *129*, 242–261.

Paller, K. A., & Wagner, A. D. (2002). Observing the transformation of experience into memory. *Trends in Cognitive Sciences*, *6*, 93–102.

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies*, *99*, 185–198.

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

Richardson, M. P., Strange, B. A., & Dolan, R. J. (2004). Encoding of emotional memories depends on amygdala and hippocampus and their interactions. *Nature Neuroscience*, *7*, 278–285.

Rugg, M. D. (1995). Event-related potential studies of human memory. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 789–801). Cambridge, MA: MIT Press.

Rugg, M. D., & Allan, K. (2000). Memory retrieval: An electrophysical perspective. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 805–816). Cambridge, MA: MIT Press.

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*, 252–257.

Siegle, G. J. (1999). A neural network model of attention biases in depression. In J. J. A. Reggia, E. Ruppin, & D. L. Glanzman (Eds.), Disorders of brain, behavior, and cognition: The neurocomputational perspective (Progress in brain research). (pp. 407–433). Amsterdam: Elsevier Science BV.

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

Stanners, R. F., Coulter, M., Sweet, A. W., & Murphy, P. (1979). The pupillary response as an indicator of arousal and cognition. *Motivation and Emotion*, *3*, 319–340.

Steinhauer, S. R., Boller, F., Zubin, J., & Pearlman, S. (1983). Pupillary dilation to emotional visual stimuli revisited. *Psychophysiology*, *20*, 472.

Taylor, S. F., Liberzon, I., Fig, L. M., Decker, L. R., Minoshima, S., & Koeppe, R. A. (1998). The effect of emotional content on visual recognition memory: A PET activation study. *NeuroImage*, *8*, 188–197.

Verney, S. P., Granholm, E., & Marshall, S. P. (2004). Pupillary responses on the visual backward masking task reflect general cognitive ability. *International Journal of Psychophysiology*, *52*, 23–36.

Võ, M. L.-H., Jacobs, A. M., & Conrad, M. (2006). Crossvalidating the Berlin Affective Word List (BAWL). *Behavior Research Methods*, *38*, 606–609.

Windmann, S., Daum, I., & Güntürkün, O. (2002). Dissociating prelexical and postlexical processing of affective information in the two hemispheres: Effects of the stimulus presentation format. *Brain and Language*, *80*, 269–286.

Windmann, S., & Krüger, T. (1998). Subconscious detection of threat as reflected by an enhanced response bias. *Consciousness and Cognition*, *7*, 603–633.

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

Windmann, S., Sakhavat, Z., & Kutas, M. (2002). Electrophysiological evidence reveals affective evaluation deficits early in stimulus processing in patients with panic disorder. *Journal of Abnormal Psychology*, *111*, 357–369.

Windmann, S., Urbach, T. P., & Kutas, M. (2002). Cognitive and neural mechanisms of decision biases in recognition memory. *Cerebral Cortex*, *12*, 808–817.

Yonelinas, A. P. (2001). Components of episodic memory: The contribution of recollection and familiarity. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *356*, 1363–1374.

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

(RECEIVED February 15, 2007; ACCEPTED July 26, 2007)