Biological Psychology 114 (2016) 23–32

Contents lists available at ScienceDirect

Biological Psychology

journal homepage: www.elsevier.com/locate/biopsycho

Attentional enhancement for positive and negative tones at an early stage of auditory processing

Tímea Folyi[a,*], Heinrich René Liesefeld[b], Dirk Wentura[a]

[a] Department of Psychology, Saarland University, Saarbrücken, Germany
[b] Department of Psychology, Ludwig-Maximilian-University of Munich, Munich, Germany

ARTICLE INFO

Article history:
Received 23 September 2014
Received in revised form 3 December 2015
Accepted 4 December 2015
Available online 8 December 2015

Keywords:
Attention
Valence
ERP
Auditory N1

ABSTRACT

We report an event-related potential (ERP) study based on the hypothesis that valenced (i.e., positive and/or negative) tones are prioritized over neutral ones at an early, perceptual stage of auditory processing. In order to avoid perceptual confounds, we induced valence experimentally during a learning phase by assigning positive, negative, and neutral valences to tone-frequencies in a balanced design. In a subsequent test phase, EEG was recorded while these tones were entirely task-irrelevant. The amplitude of the auditory N1 was increased for valenced compared with neutral tones, indicating enhanced attention. While behavioral results of the learning phase, and both implicit and explicit measures of tone evaluation indicated differentiation between positive and negative valence, there was no such differentiation on the N1 amplitude. Our results suggest that it is the general relevance of the valenced tones that governs early attentional processes.

1. Introduction

Some stimuli are highly relevant for our survival and well-being, including signals warning of potential danger or, on the opposite end of the spectrum, signals indicating potential benefits. Efficient selection of these valenced stimuli from the rapidly changing, multisensory environment is essential for fast and adaptive reactions. In the visual modality, numerous studies have demonstrated that valenced stimuli are processed more efficiently than neutral ones (for a review, see, e.g., Olofsson, Nordin, Sequeira, & Polich, 2008). Converging evidence suggests that this benefit can stem not only from the higher-order, post-perceptual stages of stimulus processing, but even from the stages of sensory encoding (Junghöfer et al., 2006; Keil et al., 2003; Olofsson et al., 2008; Schupp, Junghöfer, Weike, & Hamm, 2003; Vuilleumier, 2005; Vuilleumier & Huang, 2009), giving rise to a rapid allocation of attention (Keil et al., 2001; Pourtois, Schettino, & Vuilleumier, 2013; Vuilleumier, 2005; Vuilleumier & Huang, 2009; Yiend, 2010). A rapid activation of subcortical regions, in particular the amygdala, and the connected sensory cortices is supposed to account for early affective effects (e.g., LeDoux, 2003).

Despite a growing research interest toward affective processing, remarkably, some aspects have been neglected so far. Most importantly, we live and act in a multisensory environment. Compared to the large body of research investigating the processing of evaluative stimuli in the visual domain, similar attempts are relatively sparse in other stimulus modalities such as the auditory modality (for exceptions, see, e.g., Bröckelmann et al., 2011, 2013; Kluge et al., 2011; and Thönnessen et al., 2010, for affect-related modulation of auditory evoked fields; Czigler, Cox, Gyimesi, & Horváth, 2007, for ERP-correlates of natural aversive sounds; Heim & Keil, 2006; Kryuchkova, Tucker, Wurm, & Baayen, 2012; Paulmann & Kotz, 2008a, 2008b, for electrophysiological correlates of affective speech stimuli; Grandjean et al., 2005; Mitchell, Elliott, Barry, Cruttenden, & Woodruff, 2003; Sander, Frome, & Scheich, 2007; Sander & Scheich, 2001, for functional magnetic resonance imaging studies employing complex emotional sounds). This fact is surprising given the vital importance of auditory perception in monitoring our environment: For instance, it allows for 360-degree monitoring in space and is characterized by extensive preattentive and predictive modeling of the acoustic environment (e.g., Bendixen, SanMiguel, & Schröger, 2012; Middlebrooks & Green, 1991; Näätänen, Paavilainen, Rinne, & Alho, 2007). In the present study, we raised the following questions: Are valenced tones processed preferentially compared with neutral ones? If so, at what

* Corresponding author at: Saarland University, Department of Psychology, Campus Building A2 4, 66123 Saarbrücken, Germany. Fax: +49 681 302 4049.
E-mail addresses: t.folyi@mx.uni-saarland.de (T. Folyi), Heinrich.Liesefeld@psy.lmu.de (H.R. Liesefeld), wentura@mx.uni-saarland.de (D. Wentura).

http://dx.doi.org/10.1016/j.biopsycho.2015.12.001
0301-0511/

stage(s) of the auditory processing chain does this preferential processing occur? Is there an early (i.e., within 100–150 ms following tone onset) attentional enhancement for valenced tones?

Targeting these questions ties in with important characteristics of affective and basic cognitive research. In affective studies, the stimuli that are typically used are perceptually complex and carry strong, well-defined intrinsic valence (e.g., human emotional expressions, scene pictures). While such rather "coarse-grained" materials possess the advantage of ecological validity, in turn, they suffer from a lack of tight control over the physical stimulus attributes. This approach is particularly susceptible to stimulus confounds when early, perceptual stages of affective processing are investigated. By contrast, basic cognitive research into sensory processing and attention typically employs perceptually simple, physically well-controlled stimulus materials. In the present study, we strove for a meaningful synthesis of these two approaches: on the one hand, we exerted strict control over physical stimulus attributes by assigning positive and negative valence to simple tones in a learning phase (for a similar approach in the visual domain, see Müller, Rothermund, & Wentura, 2015; Wentura, Müller, & Rothermund, 2014; for investigating differences of simple versus complex affective stimuli, see Bradley, Hamby, Löw, & Lang, 2007). On the other hand, we wanted to induce valence in an ecologically valid context. To this end, we created a game-like situation, in which different tone-frequencies signaled the danger of losing and the chance of gaining money, respectively. Note, that we used direct, tangible resources (i.e., money that could be obtained immediately during the experiment), in contrast to paradigms employing more indirect rewards, such as presenting only the representations of aversive or desirable objects (e.g., viewing the picture of a desirable food that cannot be obtained during the experiment).

In the visual domain, several studies have applied the approach of associating initially neutral stimuli with evaluative content during a learning phase. However, these studies have mostly associated threat with simple stimuli via conditioning, without any comparison with positive valence (e.g., Batty, Cave, & Pauli, 2005; Hintze, Junghöfer, & Bruchmann, 2014; Notebaert, Crombez, Van Damme, De Houwer, & Theeuwes, 2011; Smith, Most, Newsome, & Zald, 2006; for exceptions, see, e.g., Müller et al., 2015; Wentura et al., 2014). In the present study, importantly, both negative and positive valence was attached to a priori neutral stimuli.

Given this background, we addressed a further question arising from a lingering debate in the field of evaluative picture processing, that is, whether attentional prioritization of valenced stimuli is driven (a) by a specific valence category or (b) by evaluative content in general (regardless of the direction of valence). An influential theory supporting the former assumption (a) is the threat-detector hypothesis. It claims that our attentional system is tuned to prioritize negative information over positive information, allowing for the rapid detection of negative—or more specifically, threatening—stimuli (Öhman, Flykt, & Esteves, 2001; Öhman & Mineka, 2001). We contrasted this assumption with a general relevance bias hypothesis (b), namely, that valenced stimuli in general trigger attentional resources based on their higher goal-relevance compared with neutral stimuli (Brosch, Sander, Pourtois, & Scherer, 2008; Müller et al., 2015; Rothermund, Voss, & Wentura, 2008; Wentura et al., 2014). To the best of our knowledge, our study is the first to contrast these two assumptions systematically in the auditory domain.

The excellent temporal resolution of the ERP method allows us to identify the point in time at which evaluative features influence auditory processing. In line with previous findings in the visual domain (for reviews, see Olofsson et al., 2008; Schupp, Flaisch, Stockburger, & Junghöfer, 2006) and auditory domain (Bröckelmann et al., 2011, 2013; Kluge et al., 2011), we expected that valenced tones evoke an early attentional enhancement. In the auditory domain, the N1 component is most often reported to be modulated by the participant's attentional state in the relatively early time range (i.e., within 100–150 ms following tone onset; e.g., Folyi, Fehér, & Horváth, 2012; Herrmann & Knight, 2001; Woldorff & Hillyard, 1991). The auditory N1 is a negative-going waveform that peaks maximally at frontocentral electrodes approximately 80–120 ms following a tone onset. It results from the activity of various neural sources, presumably including primary and secondary auditory cortices and non-modality specific brain regions (Giard et al., 1994; Näätänen & Picton, 1987; Woods, 1995). The N1 is considered as an example of "exogenous" auditory ERP components, as it reacts sensitively to changes in acoustic features and stimulus presentation characteristics (Barry, Cocker, Anderson, Gordon, & Rennie, 1992; Budd, Barry, Gordon, Rennie, & Michie, 1998; Crottaz-Herbette & Ragot, 2000; Dimitrijevic, Michalewski, Zeng, Pratt, & Starr, 2008; Jacobson, Lombardi, Gibbens, Ahmad, & Newman, 1992; Weise, Schröger, Fehér, Folyi, & Horváth, 2012). ERP studies on auditory attention typically report an enlarged N1 amplitude for attended tones compared with identical but unattended tones (e.g., Hillyard, Hink, Schwent, & Picton, 1973; Woldorff & Hillyard, 1991). If the sensory encoding of valenced tones is facilitated by rapid attention, one should expect an early attentional enhancement for valenced tones—reflected in an increased amplitude of the N1—as compared with neutral tones.

As most of our knowledge about affective processing originates from visual studies, unique features of sensory modalities—which can be potentially relevant concerning affective processing—were scarcely taken into consideration so far (for exception, see e.g., Czigler et al., 2007). For instance, while the full evaluative content of emotional pictures is available at stimulus onset, auditory stimuli are characterized by complex temporal structure and temporal unfolding of sound information. In the present experiment, we associated valence to sinusoidal tones (i.e., tones consisting of a single sinusoidal wave) with different tone-frequencies. We employed tone-frequency as the critical stimulus feature to convey valence information because of its high importance in auditory object formation (e.g., Bregman, Liao, & Levitan, 1990; Griffiths & Warren, 2004), and as it is processed quickly, within 100 ms following tone onset (e.g., Roberts, Ferrari, Stufflebeam, & Poeppel, 2000). Thus, we can assume that attaching valence to tone-frequency allows for earlier extraction of valence information compared with a situation where the evaluative content is defined by a combination of several stimulus features and a complex temporal structure, thus the evaluative information unfolds gradually in time (e.g., sounds of an attack scene; Bradley & Lang, 2000).

Fig. 1 gives an overview on the main phases of the experiment. The experiment consisted of three main parts: (1) a valence induction phase, (2) a test phase, and (3) a manipulation check phase. First, in the valence induction phase (Fig. 1A), half of the tones were presented with constant intensity and half of the tones with increasing intensity. The participants' task was to detect the tones with increasing intensity. Depending on their performance, participants could win or lose money in a game-like situation. Three tone-frequencies were associated with a negative (i.e., danger of losing money in the case of insufficient performance), positive (i.e., chance to win money in the case of sufficient performance), and neutral (i.e., no substantial loss or gain) connotation, respectively. The assignment of tone-frequencies to valence conditions was counterbalanced between participants. In a subsequent test phase (Fig. 1B), a selective listening task was administered, that is, participants were instructed to attend the auditory stimulation presented to one ear (task-relevant channel), while ignoring stimulation presented to the other ear (task-irrelevant channel). On the task-relevant channel, participants had to accomplish a perceptual detection task on a series of noise bursts. On the task-irrelevant channel, negative, positive, and neutral tones of the previous

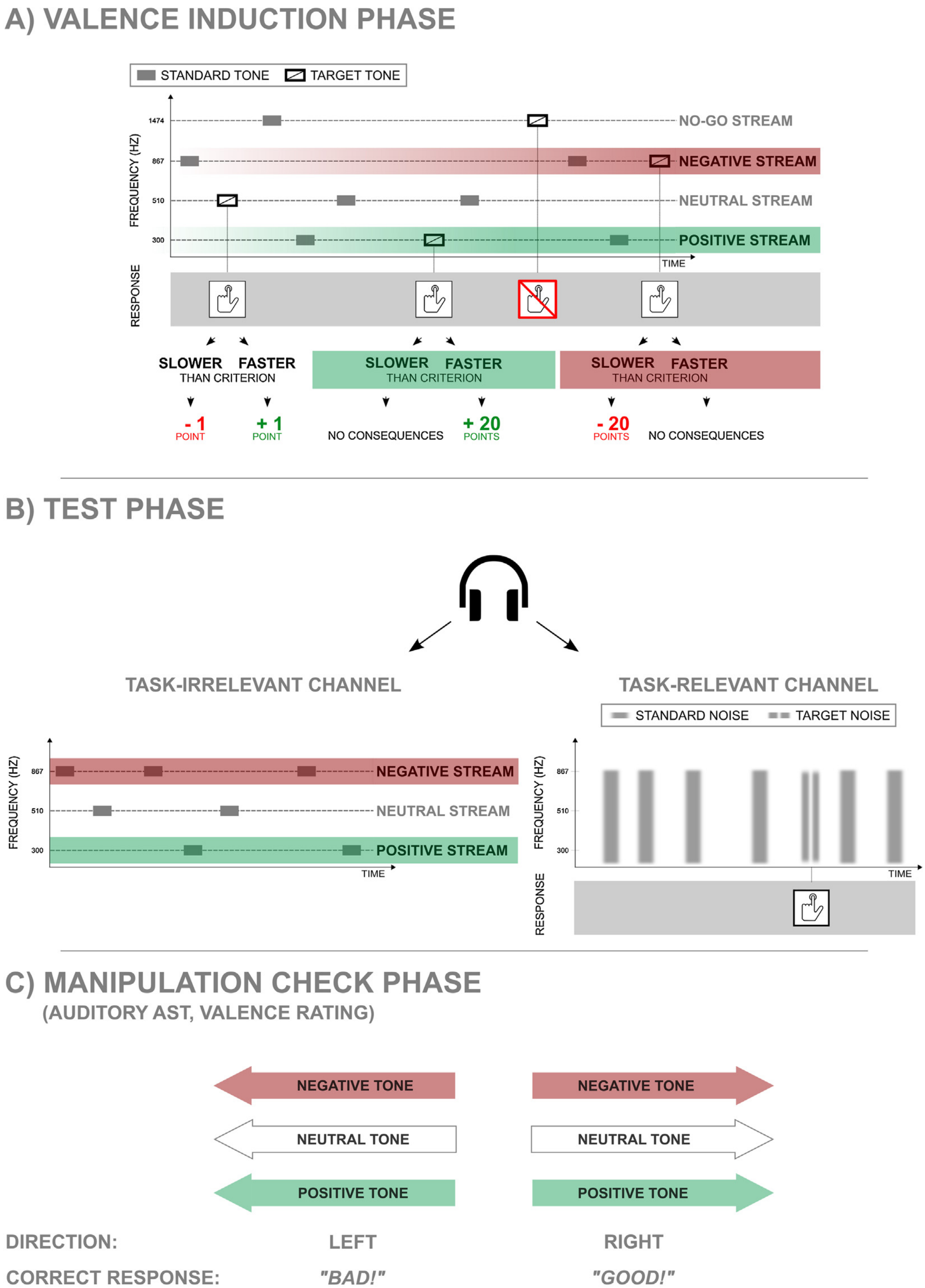

Fig. 1. Schematic illustration of the experimental design. Note that only one instance of the counterbalancing scheme is depicted. (A) In the valence induction phase, valence was induced experimentally in a game paradigm by assigning positive (chance to gain money), negative (danger to lose money), and neutral (without substantial gain or loss) meaning to tone-frequency. (B) In the test phase, positive, neutral and negative tones were presented in a task-irrelevant channel while a perceptual detection task was administered in the task-relevant channel. Participants were instructed to direct their attention to the task-relevant channel, while ignoring the task-irrelevant channel. (C) In the manipulation check phase, the effectiveness of the valence induction was tested in an auditory version of the affective Simon task (AST) and by collecting explicit valence ratings.

valence induction phase were presented, and ERPs elicited by these tones were analyzed after the experiment. Finally, in the manipulation check phase (Fig. 1C), we tested whether tones had acquired positive and negative valence during the valence induction phase. Therefore, we applied an auditory version of the affective Simon task (AST) as an indirect method to assess stimulus valence. In a nutshell, in the AST participants are presented with valenced stimuli (such as positive and negative words) and give intrinsically valent responses (e.g., saying “good” or “bad”) to a valence-neutral stimulus feature (e.g., the word’s color) that is orthogonally varied to valence. Here, the to-be-categorized valence-neutral feature of the tones was the direction of an illusory movement from a central position to one side of the perceiver. If the valence of the tones is processed implicitly, we can expect shorter reaction times (RTs) for congruent (stimulus valence and response valence match) compared with incongruent (stimulus valence and response valence mismatch) trials (for an auditory variant, see Folyi & Wentura, 2015; for the original version, see De Houwer & Eelen, 1998). At the end of the experiment, participants rated the pleasantness of the tones explicitly.

2. Methods

2.1. Participants

Twenty-four students from Saarland University (13 females; aged 17–28 years, *Mdn* = 22 years; two left-handers) participated in the experiment for monetary compensation. All participants were German native speakers, had normal or corrected-to-normal vision, and self-reported normal hearing. Before the experiment, all participants gave written informed consent. On average, participants received €34 (compensation was partly dependent on performance; see Procedure for details). Six additional individuals were excluded before data analyses because of technical problems resulting in incomplete data or because they failed to understand instructions. Data from three further participants were discarded during data analysis: two due to extensive EEG-artifacts, and one due to the absence for an observable N1 waveform in the averaged ERPs.[1] The assignment of the tone-frequencies to valence conditions was counterbalanced between participants according to a Latin square scheme (Winer, 1962), resulting in four counterbalancing groups (Table A1 in the Appendix A depicts the exact counterbalancing scheme). In the final sample, two counterbalancing groups had six, one seven and one five participants. In the affective Simon task twenty-two students participated, as two participants had to be excluded because of technical failure of the voice key apparatus.

2.2. Materials

Sinusoidal tones with four different tone-frequencies (300 Hz, 510 Hz, 867 Hz, and 1473.9Hz)[2] were presented via headphones (HD-600, Sennheiser, Wedemark, Germany). The maximal intensity level was 45 dB sensation level (SL, above individual hearing level referred to a 1000 Hz sinusoidal tone). Note that the four tones with different tone-frequencies can have slightly different perceived loudness. All auditory stimuli faded in and out with 10 ms linear rise and fall times.

2.2.1. Valence induction phase

In the valence induction task, tone duration was 1400 ms and all four tones were presented in two versions. Standard tones were presented with constant intensity (0.5 × maximal intensity; ~39 dB SL). Target tones started with the same constant intensity, but 1000 ms post onset their intensity rose to the maximal level (i.e., 45 dB SL) with a linear rise time of 390 ms.

2.2.2. Test phase

In the selective listening task, tones presented on the task-irrelevant channel were identical to the standard tones of the valence induction task, except for the following changes: tone duration was 800 ms and intensity was 0.4 × maximal intensity (~37 dB SL). On the task-relevant channel, white noise bursts were presented with an overall duration of 500 ms and intensity level of 32 dB SL. We used two versions of the noise bursts: while standard noise burst was continuous, target noise burst was interrupted by a 4 ms silent period ("gap") starting at 200 ms post onset.

2.2.3. Manipulation check phase

The amplitude of the four tones was modulated in the following way: from 500 ms after stimulus onset, intensity in the left or right auditory signal channel of the stereo sound was reduced linearly over a 1000 ms interval by a total of 75%. Thereby we created two versions of each tone: with an illusory movement from a central position toward the right ("moving to the right" tones) or left ("moving to the left" tones) side of the perceiver, respectively (see, e.g., Rosenblum, Carello, & Pastore, 1987). Overall tone duration was 6000 ms, but tones were terminated when a response was recorded. Additionally, 30 natural sounds from the IADS battery (Bradley & Lang, 2007) were presented as filler trials. We selected 10 positive (e.g., baby laugh), 10 negative (e.g., women crying), and 10 neutral (e.g., office noise) sounds and created a "moving to the left" and "moving to the right" versions of them, in the same way as described above. Averaged normative valence ratings on a 9-point scale ranging from most unpleasant (1) to most pleasant (9) were: $M = 7.20$ ($SD = 0.45$) for positive, $M = 2.62$ ($SD = 0.53$) for negative, and $M = 4.62$ ($SD = 0.49$) for neutral sounds, respectively.

2.3. Procedure

Participants initially received €10 as a payment for the first 1.5 h of the experiment, but were obliged to risk the money as the stakes in a subsequent "game". Participants received further €4.50 for each additional started half-hour. The experimental session lasted about 3.5–4 h, including individual breaks and the time for electrode application and removal.

Participants were comfortably seated in a sound-attenuated room. Hearing thresholds were individually determined by using a continuous, 1000 Hz sinusoidal tone at the beginning of the experiment. This level (0 dB SL) was used as a reference during the experiment. Before the actual experiment started, participants had to accomplish two discrimination tasks: the first required participants to discriminate the four tone-frequencies, the second required them to learn the associations between tone-frequencies and the gain/loss odds. Before the second discrimination task, the possible outcomes related to each tone were shown to the participants. Depending on their outcome odds, three tones were referred in the instructions as "*danger*", "*chance*", and "*neutral*" tones (there was a forth tone during the game termed as "*no-go*" tone, see Valence induction phase for details). The second discrimination task tested for the association between the four tone-frequencies and the four "meanings" during the game (i.e., "*danger*", "*chance*", "*neutral*", "*no-go*"). The discrimination tasks terminated when the accuracy level exceeded 95% (first task) or 90% (second task).

The valence induction phase (1) consisted of one practice and three experimental blocks of the valence induction task. The test phase (2) comprised 10 blocks of the selective listening task, with an additional block of the valence induction task interspersed after every two blocks of selective listening (i.e., the two phases together consisted of 7 valence induction blocks and 10 selective listening blocks). Blocks were separated by short, participant-terminated breaks. Finally, in the manipulation check phase (3), an auditory version of the affective Simon task was administered. Additionally, explicit valence ratings were collected.

[1] Including this data set to the final sample does not change the pattern of results reported below.

[2] Frequency differences between tones were calculated by the following formula: $\text{frequency}_n = 1.7 \times \text{frequency}_{n-1}$; e.g., 510 Hz = 1.7 × 300 Hz. The resulting tones are perceived as roughly equidistant in pitch, but there was no harmonic or musical relation between them.

2.3.1. Valence induction phase

In the valence induction phase (Fig. 1A), participants could win or lose game scores depending on their performance. Every valence induction block started with a score of zero. Participants were informed that they would immediately win €1 if the final score was zero or above. If the score was below zero, they immediately lost €1. Thus, there was a chance of winning or losing up to €7 throughout the experiment.

A valence induction block comprised 56 auditory trials presented in random order. Each trial started with the presentation of a black fixation cross with a randomized duration of 500–1000 ms. Thereafter, one of the tones was played binaurally. Half the trials featured standard tones (i.e., tones with constant intensity) and half the trials featured target tones (i.e., tones with increasing intensity). Participants were instructed to respond only to the target tones by pressing the space bar as quickly and accurately as possible. A fast and correct response was considered a success trial; a slow or incorrect response (i.e., a missed target or a false alarm) was considered a failure trial. The response-speed criterion for being successful on a given trial was defined by the moving median of the preceding six trials (see, e.g., Rothermund et al., 2008).[3] The tone-frequency determined the consequences of success or failure. One tone-frequency was associated with a gain of 20 points in case of a success, but no negative consequences in case of a failure (positive tone). Another tone-frequency was associated with a loss of 20 points in case of a failure, but no positive consequences in case of a success (negative tone). A third tone-frequency was associated with a negligible gain or loss of one point in case of either success or failure (neutral tone). Note that the possible outcomes related to each tone were presented explicitly to the participants in the beginning of the experiment (see above). A fourth tone-frequency served as a no-go signal, requiring participants to withholding their response even when the tone was presented in its target form (i.e., with increasing loudness). The no-go tone was introduced to make the tone-frequency a task-relevant feature, thereby supporting the learning of the frequency-valence associations. This aspect of the design is in accordance with our previous experience concerning a comparable valence induction method in the visual domain (e.g., Wentura et al., 2014, here the color feature was employed in an analogous way as tone-frequency in the present design). Responding to the no-go tone resulted in a loss of 20 points. The no-go tone was presented only in the valence induction phase. Visual feedback was given after each target trial and in case of a false alarm, with the feedback indicating (1) the type of the tone that was presented (e.g., "danger" target tone), (2) the consequences of the recent response (e.g., "−20 points"; on the break-even trials—for example in case of successful performance on a negative trial—the feedback stated that "There are no consequences"), and (3) the current total score.[4] In a valence induction block, 14 positive, 14 negative, 14 neutral and 14 no-go tones were presented (half of them in their standard, and half of them in their target version, respectively). In each valence induction block, seven additional visual detection trials were introduced to ensure that participants keep their visual attention on the screen and thus encode the visually presented feedback. A visual detection trial started with a 500 ms presentation of a black fixation cross, which then turned red; participants had to press the space bar as quickly as possible when the color changed. If the participant did not respond within 2000 ms of the color change, error feedback was presented visually.

2.3.2. Test phase

In the test phase (Fig. 1B), participants were instructed to direct their attention to one ear (task-relevant channel), while ignoring the sequence of tones presented to the other ear (task-irrelevant channel). In the task-relevant channel, in each block 18 target noise bursts and 54 standard noise bursts were presented monaurally in a randomized order, with a random inter-stimulus interval (ISI) of 2500–3500 ms. Participants had to respond to the target noise bursts as quickly and accurately as possible by pressing the space bar. Visual feedback about the hit rate was given after each block to motivate participants to maintain a high level of performance. In the task-irrelevant channel, the standard version of the positive, neutral, and negative tones of the valence induction phase were presented monaurally with a randomized ISI of 1000–1333 ms. The two series of auditory stimuli presented to the two ears were allowed to overlap at random temporal positions. Task channel and ear assignment was counterbalanced between participants.

2.3.3. Manipulation check phase

In the manipulation check phase (Fig. 1C), the effectiveness of the valence induction manipulation was tested in an auditory version of the affective Simon task (AST). The AST comprised 12 practice and 30 experimental trials. Positive, negative, and neutral tones were presented 10 times each in random order (five times in "moving to the left" and five times "moving to the right" version). The sequence was randomly interspersed by filler trials. That is, 10 positive, 10 negative, and 10 neutral natural sounds were presented once during the task (with half of them as "moving to the left" and half of them as "moving to the right" sounds, respectively). Participants had to categorize the direction of the "movement" by saying "good" for the "moving to the right" sounds and "bad" for "moving to the left" sounds as quickly and accurately as possible ("gut" and "böse" in German, respectively). The assignment of response (saying "good" or "bad") to illusory movement direction (right or left) was not counterbalanced between participants, as our previous experiments did not show differences between counterbalancing groups (Folyi & Wentura, 2015).

A trial started with a 1000 ms long visual presentation of a fixation cross without auditory stimuli. The fixation cross remained on the screen until the end of the trial. After 1000 ms, a positive, negative, or neutral tone (experimental trials) or a natural sound (filler trials) was played. In the first 500 ms of the stimulus presentation, the intensity was equal in the two auditory signal channels of the stereo stimulus. From 500 ms after stimulus onset, intensity in one auditory signal channel declined over a 1000 ms interval, thereby creating a movement illusion toward the right or the left side of the perceiver. While a voice key apparatus recorded the RT, the experimenter registered the response category online. After the vocal response was detected by the voice key, the auditory stimulus was terminated. Finally, error feedback was presented visually in the case of an incorrect response.

At the end of the experiment, pleasantness of the auditory stimuli presented in the AST was rated in a 9-point scale from most unpleasant (1) to most pleasant (9). During the pleasantness rating, auditory stimuli were presented with constant intensity.

2.4. EEG recording and analysis

The EEG was recorded only during the test phase from 64 scalp locations (following the international 10–10 system, including left and right mastoids). The common reference electrode was placed

[3] The actual response-speed criterion (median') was dependent on a participant's current game score. If the game score exceeded 50 points: median' = 1.02 × median; if it was below −50 points: median' = 0.98 × median; if it was between −50 and 50 points: median' = median.

[4] 20-point gains and losses additionally elicited feedback sounds (a fanfare-like trumpet sound in case of a gain and a guitar sound with decreasing pitch in case of a loss; both sounds were provided by the FreeSound Project, http://www.freesound.org). Note that these additional sounds were presented only in the valence induction phase.

on the tip of the nose.[5] The continuous EEG was amplified from DC to 100 Hz at a 500-Hz sampling rate. On-line 70-Hz low-pass filtering was applied and the signal was band-pass filtered offline (0.5–30 Hz). Horizontal eye movements were monitored with a bipolar setup, with electrodes placed laterally to the outer canthi of both eyes; vertical eye movements were monitored with electrodes placed above and below the right eye.

ERPs elicited by the tones of the task-irrelevant channel were calculated during offline analysis. To this end, the continuous EEG was segmented into 1000 ms long epochs, each including a 200 ms long baseline preceding tone onset. Epochs contaminated with severe artifacts were rejected (rejection criteria: signal range exceeding 200 μV, or voltage step exceeding 50 μV/ms, or a voltage difference exceeding 150 μV on any channel). Additionally, epochs corresponding to the first three trials of each block were discarded, because auditory N1 is reported to show an enhanced amplitude after a longer silent period that is decreasing rapidly during the first few trials of a stimulus block (e.g., Näätänen & Picton, 1987). We thus excluded the first few trials of the blocks to avoid this error variance and, consequently, to increase the signal-to-noise ratio. Epochs containing task-relevant sounds were also discarded from the analysis (rejection criteria: onset of a task-relevant standard sound within a −800 ms to +800 ms time-window relative to the onset of the task-irrelevant tone, or onset of a task-relevant target sound within −1000 ms to +800 ms of the task-irrelevant tone onset). We applied these rejection criteria on the one hand to avoid overlapping of neural responses elicited by the task-irrelevant and the task-relevant stimuli, and on the other hand, to maximize trial numbers in order to increase signal-to-noise ratio. On average, ERPs were based on 165, 163, and 159 trials per participant in the positive, neutral, and negative valence conditions, respectively. Epochs were baseline corrected using the 200-ms pre-stimulus interval and averaged separately for the different valence conditions.

We formed a region of interest (ROI) from frontocentral electrode sites (F1, Fz, F2, FC1, FCz, FC2, C1, Cz, C2) according to the frontocentral scalp distribution of the grand average N1 (group average collapsed across experimental conditions) and according to the auditory N1-literature (for a review, see Näätänen & Picton, 1987). The N1 was identified in the individual ERPs as the strongest negative local peak within a 60 ms long time window centered around the latency of the group-average peak (114 ms; experimental conditions collapsed). Baseline-to-peak amplitudes were derived for each participant as the mean value in an 8 ms long time window around the individual peak. For testing the valence effect on N1 amplitudes, we applied a 3 × 4 mixed design with valence (positive, neutral, negative) as within-participants factor and tone-frequency-to-valence assignment (four counterbalancing groups) as a between-participants factor. We added counterbalancing group as a between-participants factor to use the correct error term (see Pollatsek & Well, 1995) and to make sure that our results are not affected by slightly different group sizes.[6] We used the multivariate approach to repeated measures analysis, which means that the tripartite factor of valence was transformed into a vector of two orthogonal contrast variables (see, e.g., O'Brien & Kaiser, 1985; Petrova & Wentura, 2012; Rohr, Degner, & Wentura, 2012). We a priori chose the contrasts in a way that they represent the specific hypotheses outlined above. That is, for the first contrast, amplitude values were averaged across positive and negative stimuli and contrasted with the neutral stimuli. This contrast represents the hypothesis that valenced tones (in general) produce larger attentional effects compared with neutral tones. The second contrast was the contrast between positive and negative tones, representing the hypothesis of larger attentional effects for negative compared with positive tones.

3. Results

3.1. Behavioral performance

Behavioral performance was adequate in the valence induction and selective listening tasks, suggesting that participants were engaged in both tasks. RTs below 100 ms were discarded when calculating averaged RTs. In the valence induction task, RTs were calculated from the beginning of the loudness increase in the ongoing tone. The average RT across participants was 327 ms (SD = 32 ms); 323 ms (SD = 39 ms) for positive, 326 ms (SD = 31 ms) for neutral, and 332 ms (SD = 38 ms) for negative tones, and did not show significant valence differences, $F < 1$. The average accuracy across participants was 91.3% (SD = 10.6%); 88.4% (SD = 13.5%) for positive, 93.7% (SD = 10.5%) for neutral, and 92.0% (SD = 9.5%) for negative tones, and showed a valence main effect, $F(2,22) = 6.82$, $p = .005$, $\eta_p^2 = .383$. This valence difference in accuracy is more understandable in the light of the false alarm rates: Average false alarm rate were 17.3% (SD = 26.3%) in positive, 10.9% (SD = 20.8%) in neutral and 7.2% (SD = 12.0%) in negative conditions, respectively. Note, that high false alarm rate pays off only in the positive condition, as false alarm had no negative consequence only in this condition (see Procedure for details). Accordingly, false alarm rates showed a tendency for a valence effect, $F(2,22) = 2.81$, $p = .082$, $\eta_p^2 = .203$; and the contrast of positive vs. negative valence was significant, $F(1,23) = 5.40$, $p = .029$, $\eta_p^2 = .190$. The contrast of valenced (positive and negative) vs. neutral conditions was not significant, $F(1,23) = 0.45$, $p = .511$, $\eta_p^2 = .019$. In the selective listening task of the test phase, the average RT across participants was 590 ms (SD = 74 ms) with an average accuracy of 93.2% (SD = 5.1%).

3.2. ERP data

Prototypical auditory ERPs (P1, N1, P2) were clearly observable in all conditions in the group average waveforms (Fig. 2A). A 3 × 4 repeated measures MANOVA with valence (positive, neutral, negative) as within-participants factor and balancing group as a between-participants factor[7] of the N1 amplitudes yielded a valence main effect, $F(2,19) = 3.64$, $p = .046$, $\eta_p^2 = .277$. Balancing group did not show any significant main effect or interaction with valence, all Fs < 1.43. The a priori planned contrast of valenced (positive and negative) vs. neutral condition was significant, $F(1,20) = 7.01$, $p = .015$, $\eta_p^2 = .260$, indicating higher N1 amplitudes in the valenced conditions compared with the neutral condition ($M = -3.03$ μV, SD = 1.54 μV in the valenced condition; $M = -2.39$ μV, SD = 1.59 μV in the neutral condition). The contrast of the two valenced conditions (positive vs. negative) did not show any differences, $F(1,20) = 0.33$, $p = .572$, $\eta_p^2 = .016$.

Inspection of Fig. 2 indicates an apparent valence-effect in the P1 time window (approximately 50 ms following tone onset) and in a later time range (350–500 ms following tone onset). Although our hypotheses relate only to the N1 component (approximately

[5] An anonymous reviewer pointed out that applying nose reference may decrease N1 effects observed on the frontocentral regions as an electrode attached to the nose can pick up activity from frontal generators. Although this choice might therefore decrease possible effects (and might be reconsidered in further studies), it did not bias our main analyses of finding a valence-dependent moderation.

[6] Note, (a) an analysis without group factor, (b) an analysis with dependent values adjusted for the main effect of tone-frequency, and (c) completely balanced final samples achieved by random extraction of participants (i.e., final samples with five participants in each balancing group) yielded essentially the same patterns of results as the method presented above.

[7] We used the Pillai–Barlett trace criterion and its associated F-test for any test involving the group factor (see Olson, 1976).

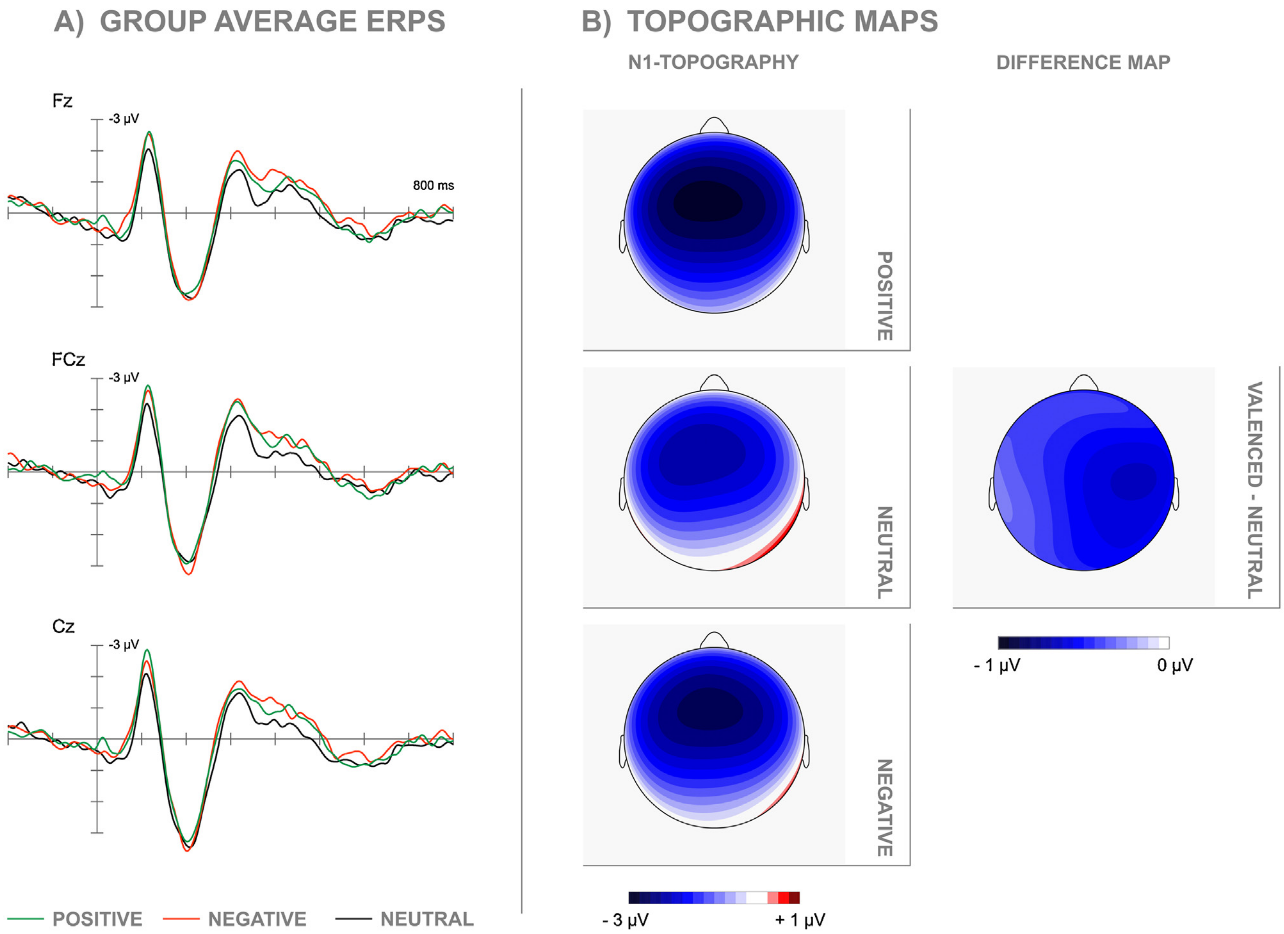

Fig. 2. ERP results. (A) Group average ERP waveforms to the positive, negative and neutral tones on the representative Fz, FCz and Cz electrode sites. The physical onset of the tones is at the crossing of the axes (0 ms). Negative polarity is plotted upwards. (B) Group average topography maps in the N1 time window (110–118 ms) in positive, neutral and negative conditions; and valenced-minus-neutral difference map.

100–150 ms following tone onset), we performed post-hoc analyses on activity in these time ranges. P1 peaks were identified in the individual ERPs as the strongest positive local peak at the frontocentral ROI within a 40 ms long time window centered around the latency of the group-average peak (46 ms; experimental conditions collapsed), and baseline-to-peak amplitudes were derived for each participant as the mean value in an 8 ms long time window around the individual peak. A 3 × 4 repeated measures MANOVA with valence (positive, neutral, negative) as within-participants factor and balancing group as a between-participants factor, however, showed no valence effect on the P1 amplitudes, $F(2,19) = 1.72$, $p = .206$, $\eta_p^2 = .153$; all other $Fs < 2.19$, $ps > .064$.

The second post-hoc analysis targeted the apparent enhancement of a frontocentrally distributed negativity for valenced compared with the neutral tones in a later time range. A 3 × 4 repeated measures MANOVA with valence (positive, neutral, negative) as within-participants factor and balancing group as a between-participants factor on the mean activity measured in the 350–500 ms time range at the frontocentral ROI showed a tendency for a valence main effect, $F(2,19) = 2.76$, $p = .089$, $\eta_p^2 = .225$. Balancing group did not show any significant main effect or interaction with valence, $Fs < 1.47$. The contrast of valenced (positive and negative) vs. neutral condition was significant, $F(1,20) = 5.38$, $p = .031$, $\eta_p^2 = .212$, indicating more negative activity in the valenced conditions compared with the neutral condition ($M = -0.83$ µV, $SD = 0.68$ µV in the valenced condition; $M = -0.34$ µV, $SD = 1.12$ µV in the neutral condition). The contrast of the two valenced conditions (positive vs. negative) did not show any differences, $F(1,20) = 0.48$, $p = .497$, $\eta_p^2 = .023$.

3.3. *Manipulation check*

3.3.1. *Affective Simon task*

RTs were calculated from the beginning of the illusory movement in the ongoing tone. RT analyses were restricted to correct AST trials (6.3% of the trials was excluded because of erroneous response of the participant or erroneous or non-reaction of the voice key). As an a priori criterion, RTs below 200 ms and above 2000 ms were discarded from further analyses (2.4% of the trials). Additionally, RT outliers (1.8% of the trials) were excluded (for each participant an upper and a lower outlier criterion were calculated according to Tukey, 1977, using the distribution of all AST latencies of this person). Mean RT for congruent trials (i.e., positive tone/response “good”; negative tone/response “bad”) was $M = 812$ ms ($SD = 182$ ms); mean RT for incongruent trials (i.e., positive tone/response “bad”; negative tone/response “good”) was $M = 854$ ms ($SD = 181$ ms). Mean RT for neutral tones was $M = 809$ ms ($SD = 156$ ms). As expected, the incongruent-minus-congruent difference showed a significant AST effect, $t(21) = 2.19$, $p = .040$, $d = 0.47$. However, similar analyses on the filler trials (natural sounds) did not yield significant results, $t < 1$.

3.3.2. *Ratings*

Preliminary inspection of explicit pleasantness ratings of the three tones revealed a dominant pattern: Participants oriented

Table 1
Mean valence ratings in each valence condition for individually lowest, intermediate, and highest pitch. Valence ratings range from very unpleasant (1) to very pleasant (9); *SD* in parentheses.

Valence	Pitch			All
	Low	Intermediate	High	
Positive	6.00 (2.00)	4.83 (1.47)	3.10 (1.45)	4.64
Neutral	5.40 (1.34)	4.70 (1.77)	2.17 (1.47)	4.09
Negative	5.90 (1.73)	3.17 (1.47)	2.67 (0.52)	3.91
All	5.77	4.23	2.64	

themselves on the comparison of stimuli with regard to the most salient feature, that is, pitch. Thus, ratings were highest for the individually lowest pitch and lowest for the individually highest pitch with a medium rating for the third tone (for the mean ratings, see Table 1). As a clear linear pattern of pitch emerged for each participant (i.e., the individually lowest pitch was rated as individually most pleasant, the individually highest pitch was rated as individually most unpleasant, and the intermediate pitch received an individually intermediate pleasantness rating; with some exceptions when ratings were equal for the adjacent tones), we used this dominant pattern in the following analysis, i.e., we coded tones according to their individual, relative pitch instead of their objective tone-frequency. As this clear linear pattern of pitch was conspicuous in our data, we have chosen the approach of linear mixed models to assess whether valence had an effect above the linear effect of pitch. We used the *lme4* and *lmerTest* packages of R 3.1.3 (Bates, Maechler, Bolker, & Walker, 2014) with significance of predictors assessed using Satterthwaite's approximation for degrees of freedom (Kuznetsova, Brockhoff, & Christensen, 2014). First, we ran two random effects model with pitch (-1 = lowest, 0 = medium, +1 = highest) as predictor and rating as dependent variable. Model 1 included random intercepts and random slopes; Model 2 included only random intercepts. Model comparison yielded no significant difference, that is, the removal of the parameter for by-participant random slopes for pitch is justified, $\chi^2(2) = 2.50$, $p = .287$ (see, e.g., Baayen, Davidson, & Bates, 2008). Matching expectations, means (and *SD*s) of the residuals of Model 2 are $M = 0.36$ ($SD = 1.55$) for positive stimuli, $M = -0.29$ ($SD = 1.46$) for negative stimuli, and $M = -0.07$ ($SD = 1.59$) for neutral stimuli. Thus, second, we ran two further models on the basis of Model 2: in Model 3 we added valence (-1 = negative, 0 = neutral, +1 = positive) as a predictor including random slopes; in Model 4, we removed the random slopes parameter. Model comparison yielded no significant difference between Models 3 and 4, $\chi^2(2) = 1.35$, $p = .510$, thus, Model 4 can be considered the final model. In this model, pitch as well as valence are significant predictors of the rating, $B = -1.61$, $t(41.39) = -10.33$, $p < .001$ for pitch and $B = 0.34$, $t(41.16) = 2.20$, $p = .033$ for valence.

4. Discussion

The main finding of the study is that valenced tones were processed preferentially compared with neutral ones. In particular, we found augmented N1 amplitude for valenced tones compared with neutral ones, thus suggesting enhanced attention for these tones. This attentional enhancement in the N1 time range indicates preferential processing of valenced tones already at a perceptual stage of auditory processing. This result is in line with the view that processing of evaluative stimuli can be facilitated by rapid attention (for a review, see Yiend, 2010). EEG was measured when the valenced tones were presented on an entirely task-irrelevant channel and participants were asked to direct their attention to the concurrent task-relevant channel to accomplish a demanding perceptual task. This setting favors the interpretation of the reported attentional enhancement as automatic in the sense of involuntariness, in line with the results of Bröckelmann et al. (2011, 2013). However, we also leave room for the explanation that voluntarily directed attention could contribute to the present results. Further research can elucidate the nature of early attentional enhancement for evaluative sounds concerning its automaticity.

To the best of our knowledge, the present study is the first to present evidence for such an early attention effect for valenced stimuli in auditory ERPs, demonstrating a fast interplay between attention and affective factors during sound processing. ERP studies of valenced stimuli show great variability, especially in an early latency range (within about 100–150 ms after stimulus onset; e.g., Olofsson et al., 2008). One possible source of this variability is that early ERPs are strongly influenced by the physical characteristics of the presented stimuli (e.g., Burkard, Don, & Eggermont, 2007), which makes experimental designs comparing physically complex stimuli susceptible to stimulus confounds. A strength of the present study is the strict control over arbitrary physical differences. This was achieved by assigning a priori neutral tones with positive, negative, and neutral meaning through a learning phase. It is important to note that our finding—a clear difference on N1 amplitude between valenced and neutral stimuli—cannot be explained by the variable physical feature of tone-frequency, as it was counterbalanced between participants. Additionally, we should highlight that our approach has the advantage of an immediate signal value: As we employed constant tone-frequency as the critical stimulus feature to convey valence information, evaluative information was available already at sound onset. By contrast, evaluative connotation of naturally valenced sounds becomes evident typically only after a delay due to temporal unfolding of the sound information. Thus, our approach and the approach to explore naturally valenced sounds complement each other.

Importantly, similarly to previous findings in the visual domain (e.g., Brosch et al., 2008; Rothermund et al., 2008; Wentura et al., 2014) we did not find a difference between positive and negative valence conditions in the N1-time range. This result is in accordance with a general relevance principle, that is, it is the goal-relevance of the stimulus that possesses attention-grabbing power at the early level of sound encoding rather than a specific valence category.

Besides our main ERP results on the relatively early N1, we found an increased sustained negativity for valenced tones in a later time range (350–500 ms following tone onset). This activity is obviously different from the enhanced positive-going waves for evaluative stimuli starting at about 200–300 ms after stimulus onset (P3 and late positive potentials) that typically occur for several versions of the oddball paradigm with visually presented complex stimulus material. These late positive-going waveforms have been associated with increased attentive processing of the motivationally significant information for subsequent memory storage (for a review, see Olofsson et al., 2008). As in our paradigm valenced tones were entirely task-irrelevant during the test phase, a possible interpretation is that the late negativity indexes inhibition for the valenced tones and/or reorientation to the main task (for a similar argumentation, see e.g., Roye, Jacobsen, & Schröger, 2007).

While the early attentional effect reflected in the enhanced N1 amplitude was comparable for positive and negative valence, both implicit and explicit behavioral measures differed between the acquired positive and negative valence. In the valence induction phase, behavioral results differentiated between valence conditions as reflected in higher false alarm rates for positive compared with negative tones. RTs, in contrast, did not differ between the valence conditions. RT-based paradigms employing monetary incentives often show reward-related speed-up (for reward-related RT-facilitation in monetary incentive delay tasks, see e.g., Knutson, Fong, Adams, Varner, & Hommer, 2001; Pizzagalli et al., 2009 but see also Knutson, Adams, Fong, & Hommer, 2001, for no valence differences on RTs but significant valence differences on self-report measures). The absence of this RT-facilitation in the present study can be a consequence of the task characteristics that were optimized for inducing tone-frequency to valance associations: (1) contrary to monetary incentive delay tasks, no pretarget cue signaling possible outcomes was applied in our paradigm but this information was indicated by the task-relevant stimulus itself, thereby possibly giving rise to less pronounced behavioral effects; and (2) our neutral condition was defined by equivocal and relatively insignificant monetary contingency—instead of complete absence of monetary contingency—thereby creating an "active" neutral condition. Furthermore, in our task design speeded responses were necessary for successful task-performance according to a strict RT-criterion in each valence condition. The pattern of RTs can thus be interpreted as an indicator that participants were engaged in the task with a constant effort in line with the task instructions. In the manipulation check phase, using an auditory variant of the affective Simon paradigm for accessing implicit evaluations, we found evidence that the simple tones had indeed acquired positive and negative affective valence during the previous learning phase. It might appear somewhat surprising that the natural sounds (i.e., the filler trials) did not reveal an AS effect. Note, that the manipulation check was administered at the end of a long and demanding experimental session. For precisely that reason we used only few trials; each natural sound was presented once and was not balanced for movement direction. Importantly, in independent experiments using only natural sounds we did find the expected AS effect (Folyi & Wentura, 2015). Besides the evidence for acquired implicit valence, we also found support for valence differentiation in explicit self-reports: although the explicit pleasantness ratings reflected dominantly the most salient feature of the stimuli, that is, pitch, they also revealed the expected differences in acquired affective valence.

In summary, our results support the view that sensory encoding of valenced stimuli is facilitated by an enhancement of "natural" attention. Moreover, besides the great number of previous studies investigating evaluative picture processing, the present study found evidence for rapid preferential processing of valenced tones. While explicit and implicit behavioral results differentiated between positive and negative valence, our ERP results suggest that at the early level of sound encoding the general relevance of the valenced tones governs attentional processes.

Acknowledgements

The present research was conducted within the International Research Training Group "Adaptive Minds" supported by the German Research Foundation [GRK 1457]. The authors thank Gabrielle Cooper for her help with style and grammar on an earlier version of the manuscript. We also thank Ullrich Ecker and Erich Schröger for their helpful comments. The authors wish to thank Thorid Römer and Claudinéia Tirciane Flach for assistance in data collection.

Appendix A.

Table A1
Assignment of tone-frequency to valence conditions according to a Latin square design. No-go tones were presented only during the valence induction phase.

Tone-frequency	Counterbalancing group			
	1	2	3	4
Low	Neutral	Positive	No-go	Negative
Middle low	Positive	No-go	Negative	Neutral
Middle high	No-go	Negative	Neutral	Positive
High	Negative	Neutral	Positive	No-go

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.
Barry, R. J., Cocker, K. I., Anderson, J. W., Gordon, E., & Rennie, C. (1992). Does the N100 evoked potential really habituate? Evidence from a paradigm appropriate to a clinical setting. *International Journal of Psychophysiology*, *13*, 9–16.
Bates D., Maechler, M., Bolker, B., Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1. 1–7, <URL: http://CRAN.R-project.org/package=lme4>.
Batty, M. J., Cave, K. R., & Pauli, P. (2005). Abstract stimuli associated with threat through conditioning cannot be detected preattentively. *Emotion*, *5*, 418–430.
Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: a review. *International Journal of Psychophysiology*, *83*, 120–131.
Bradley, M. M., Hamby, S., Löw, A., & Lang, P. J. (2007). Brain potentials in perception: picture complexity and emotional arousal. *Psychophysiology*, *44*, 364–373.
Bradley, M. M., & Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology*, *37*, 204–215.
Bradley, M. M., & Lang, P. J. (2007). *The International Affective Digitized Sounds: affective ratings of sounds and instruction manual* (2nd Edition). Gainesville, FL: University of Florida. Tech. Rep. B-3 (IADS-2)
Bregman, A. S., Liao, C., & Levitan, R. (1990). Auditory grouping based on fundamental frequency and formant peak frequency. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *44*, 400–413.
Bröckelmann, A.-K., Steinberg, C., Dobel, C., Elling, L., Zwanzger, P., Pantev, C., & Junghöfer, M. (2013). Affect-specific modulation of the N1m to shock-conditioned tones: magnetoencephalographic correlates. *European Journal of Neuroscience*, *37*, 303–315.
Bröckelmann, A.-K., Steinberg, C., Elling, L., Zwanzger, P., Pantev, C., & Junghöfer, M. (2011). Emotion-associated tones attract enhanced attention at early auditory processing: magnetoencephalographic correlates. *The Journal of Neuroscience*, *31*, 7801–7810.
Brosch, T., Sander, D., Pourtois, G., & Scherer, K. R. (2008). Beyond fear: rapid spatial orienting toward positive emotional stimuli. *Psychological Science*, *19*, 362–370.
Budd, T. W., Barry, R. J., Gordon, E., Rennie, C., & Michie, P. T. (1998). Decrement of the N1 auditory event-related potential with stimulus repetition: habituation vs. refractoriness. *International Journal of Psychophysiology*, *31*, 51–68.
Burkard, R. F., Don, M., & Eggermont, J. J. (2007). *Auditory evoked potentials: basic principles and clinical application*. Baltimore, USA: Lippincott Williams & Wilkins.
Crottaz-Herbette, S., & Ragot, R. (2000). Perception of complex sounds: N1 latency codes pitch and topography codes spectra. *Clinical Neurophysiology*, *111*, 1759–1766.
Czigler, I., Cox, T. J., Gyimesi, K., & Horváth, J. (2007). Event-related potential study to aversive auditory stimuli. *Neuroscience Letters*, *420*, 251–256.
De Houwer, J., & Eelen, P. (1998). An affective variant of the Simon paradigm. *Cognition and Emotion*, *12*, 45–62.
Dimitrijevic, A., Michalewski, H. J., Zeng, F. G., Pratt, H., & Starr, A. (2008). Frequency changes in a continuous tone: auditory cortical potentials. *Clinical Neurophysiology*, *119*, 2111–2124.
Folyi, T., Fehér, B., & Horváth, J. (2012). Stimulus-focused attention speeds up auditory processing. *International Journal of Psychophysiology*, *84*, 155–163.
Folyi, T., & Wentura, D. (2015). Fast and automatic evaluation of emotional sounds: Evidence from a brief segment rating task and an auditory variant of the affective Simon task. *Cognition & Emotion*, http://dx.doi.org/10.1080/02699931.2015.1110514. Advance online publication
Giard, M. H., Perrin, F., Echallier, J. F., Thévenet, M., Froment, J. C., & Pernier, J. (1994). Dissociation of temporal and frontal components in the human auditory N1 wave: a scalp current density and dipole model analysis. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *92*, 238–252.
Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nature Neuroscience*, *8*, 145–146.
Griffiths, T. D., & Warren, J. D. (2004). What is an auditory object? *Nature Reviews Neuroscience*, *5*, 887–892.
Heim, S., & Keil, A. (2006). Effects of classical conditioning on identification and cortical processing of speech syllables. *Experimental Brain Research*, *175*, 411–424.
Herrmann, C. S., & Knight, R. T. (2001). Mechanisms of human attention: event-related potentials and oscillations. *Neuroscience and Biobehavioral Reviews*, *25*, 465–476.
Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*, 177–180.
Hintze, P., Junghöfer, M., & Bruchmann, M. (2014). Evidence for rapid prefrontal emotional evaluation from visual evoked responses to conditioned gratings. *Biological Psychology*, *99*, 125–136.
Jacobson, G. P., Lombardi, D. M., Gibbens, N. D., Ahmad, B. K., & Newman, C. W. (1992). The effects of stimulus frequency and recording site on the amplitude and latency of multichannel cortical auditory evoked potential (CAEP) component Nl. *Ear and Hearing*, *13*, 300–306.

Junghöfer, M., Sabatinelli, D., Bradley, M. M., Schupp, H. T., Elbert, T. R., & Lang, P. J. (2006). Fleeting images: rapid affect discrimination in the visual cortex. *Neuroreport*, *17*, 225–229.

Keil, A., Gruber, T., Müller, M. M., Moratti, S., Stolarova, M., Bradley, M. M., & Lang, P. J. (2003). Early modulation of visual perception by emotional arousal: evidence from steady-state visual evoked brain potentials. *Cognitive, Affective, & Behavioral Neuroscience*, *3*, 195–206.

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*, 2057–2068.

Kluge, C., Bauer, M., Leff, A. P., Heinze, H.-J., Dolan, R. J., & Driver, J. (2011). Plasticity of human auditory-evoked fields induced by shock conditioning and contingency reversal. *PNAS*, *108*, 12545–12550.

Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L., & Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport*, *12*, 3683–3687.

Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *The Journal of Neuroscience*, *21*, RC159.

Kryuchkova, T., Tucker, B. V., Wurm, L. H., & Baayen, R. H. (2012). Danger and usefulness are detected early in auditory lexical processing: evidence from electroencephalography. *Brain and Language*, *122*, 81–91.

Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. (2014). lmerTest: Tests in linear mixed effects models. R package version 2. 0–20, <URL: http://CRAN.R-project.org/package=lmerTest>.

LeDoux, J. (2003). The emotional brain, fear, and the amygdala. *Cellular and Molecular Neurobiology*, *23*, 727–738.

Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology*, *42*, 135–159.

Mitchell, R. L., Elliott, R., Barry, M., Cruttenden, A., & Woodruff, P. W. (2003). The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia*, *41*, 1410–1421.

Müller, S., Rothermund, K., & Wentura, D. (2015). Relevance drives attention: attentional bias for gain- and loss-related stimuli is driven by delayed disengagement. *Quarterly Journal of Experimental Psychology*, http://dx.doi.org/10.1080/17470218.2015.1049624. Advance online publication

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology*, *118*, 2544–2590.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, *24*, 375–425.

Notebaert, L., Crombez, G., Van Damme, S., De Houwer, J., & Theeuwes, J. (2011). Signals of threat do not capture, but prioritize, attention: a conditioning approach. *Emotion*, *11*, 81–89.

O'Brien, R. G., & Kaiser, M. K. (1985). MANOVA method for analyzing repeated measures designs: an extensive primer. *Psychological Bulletin*, *97*, 316–333.

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

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological Review*, *108*, 483–522.

Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: an integrative review of ERP findings. *Biological Psychology*, *77*, 247–265.

Olson, C. L. (1976). On choosing a test statistic in multivariate analysis of variance. *Psychological Bulletin*, *83*, 579–586.

Paulmann, S., & Kotz, S. A. (2008a). Early emotional prosody perception based on different speaker voices. *Neuroreport*, *19*, 209–213.

Paulmann, S., & Kotz, S. A. (2008b). An ERP investigation on the temporal dynamics of emotional prosody and emotional semantics in pseudo- and lexical-sentence context. *Brain and Language*, *105*, 59–69.

Petrova, K., & Wentura, D. (2012). Upper–lower visual field asymmetries in oculomotor inhibition of emotional distractors. *Vision Research*, *62*, 209–219.

Pizzagalli, D. A., Holmes, A. J., Dillon, D. G., Goetz, E. L., Birk, J. L., Bogdan, R., & Fava, M. (2009). Reduced caudate and nucleus accumbens response to rewards in unmedicated individuals with major depressive disorder. *The American Journal of Psychiatry*, *166*, 702–710. http://dx.doi.org/10.1176/appi.ajp.2008.08081201

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: a suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 785–794.

Pourtois, G., Schettino, A., & Vuilleumier, P. (2013). Brain mechanisms for emotional influences on perception and attention: what is magic and what is not. *Biological Psychology*, *92*, 492–512.

Roberts, T. P., Ferrari, P., Stufflebeam, S. M., & Poeppel, D. (2000). Latency of the auditory evoked neuromagnetic field components: stimulus dependence and insights toward perception. *Journal of Clinical Neurophysiology*, *17*, 114–129.

Rohr, M., Degner, J., & Wentura, D. (2012). Masked emotional priming beyond global valence activations. *Cognition and Emotion*, *26*, 224–244.

Rosenblum, L. D., Carello, C., & Pastore, R. E. (1987). Relative effectiveness of three stimulus variables for locating a moving sound source. *Perception*, *16*, 175–186.

Rothermund, K., Voss, A., & Wentura, D. (2008). Counter-regulation in affective attentional biases: a basic mechanism that warrants flexibility in emotion and motivation. *Emotion*, *8*, 34–46.

Roye, A., Jacobsen, T., & Schröger, E. (2007). Personal significance is encoded automatically by the human brain: an event-related potential study with ringtones. *European Journal of Neuroscience*, *26*, 784–790.

Sander, K., Frome, Y., & Scheich, H. (2007). FMRI activations of amygdala, cingulate cortex, and auditory cortex by infant laughing and crying. *Human Brain Mapping*, *28*, 1007–1022.

Sander, K., & Scheich, H. (2001). Auditory perception of laughing and crying activates human amygdala regardless of attentional state. *Cognitive Brain Research*, *12*, 181–198.

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*, 7–13.

Smith, S. D., Most, S. B., Newsome, L. A., & Zald, D. H. (2006). An emotion-induced attentional blink elicited by aversively conditioned stimuli. *Emotion*, *6*, 523–527.

Thönnessen, H., Boers, F., Dammers, J., Chen, Y.-H., Norra, C., & Mathiak, K. (2010). Early sensory encoding of affective prosody: neuromagnetic tomography of emotional category changes. *NeuroImage*, *50*, 250–259.

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

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

Vuilleumier, P., & Huang, Y. M. (2009). Emotional attention: uncovering the mechanisms of affective biases in perception. *Current Directions in Psychological Science*, *18*, 148–152.

Weise, A., Schröger, E., Fehér, B., Folyi, T., & Horváth, J. (2012). Auditory event-related potentials reflect dedicated change detection activity for higher-order acoustic transitions. *Biological Psychology*, *91*, 142–149.

Wentura, D., Müller, P., & Rothermund, K. (2014). Attentional capture by evaluative stimuli: gain- and loss-connoting colors boost the additional-singleton effect. *Psychonomic Bulletin & Review*, *21*, 701–707.

Winer, B. J. (1962). *Statistical principles in experimental design*. New York, NY: McGraw Hill.

Woldorff, M. G., & Hillyard, S. A. (1991). Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography and Clinical Neurophysiology*, *79*, 170–191.

Woods, D. L. (1995). The component structure of the N1 wave of the human auditory evoked potential. *Electroencephalography and Clinical Neurophysiology*, *44*(Suppl), 102–109.

Yiend, J. (2010). The effects of emotion on attention: a review of attentional processing of emotional information. *Cognition and Emotion*, *24*, 3–47.