International Journal of Psychophysiology 138 (2019) 11–26

Contents lists available at ScienceDirect

International Journal of Psychophysiology

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

Involuntary sensory enhancement of gain- and loss-associated tones: A general relevance principle☆

Timea Folyi*, Dirk Wentura

Department of Psychology, Saarland University, Germany

ARTICLE INFO

Keywords:
Involuntary attention
Affective valence
Value-based attention
Auditory processing
ERP
Auditory N1

ABSTRACT

In a recent event-related potential (ERP) study (Folyi et al., 2016), we have demonstrated that sensory processing of task-irrelevant tones is enhanced when they were previously associated with positive or negative (by the means of monetary gains and losses, respectively) affective meaning relative to tones with neutral meaning, as indexed by the enhancement of the auditory N1-amplitude. In the present study, (1) in line with the hypothesis of affective counter-regulation, we investigated whether positive versus negative tones can receive differential attentional enhancement, depending on motivational context (Experiment 1); and (2) whether the early facilitation of positive and negative tones can operate strictly outside of the focus of voluntary attention (Experiment 2). In Experiment 1, we replicated the basic N1 valence effect, but found no moderation by motivational context. In Experiment 2, we found a small valence effect on the N1. By combining data from the three experiments (i.e., our previous experiment and the present ones; $N = 72$), we found a clear enhancement of N1-amplitudes for valenced tones without moderation by experiment. This pattern of results suggests comparable early attentional enhancement of valenced tones in general: (a) despite different level of concurrent task-relevant attentional and motivational demands in these experiments; and (b) without prioritizing one valence category over another, supporting our claim that the general relevance of the tones with high motivational value that governs early attentional facilitation.

1. Introduction

The ability to detect and respond fast and efficiently to stimuli with high emotional or motivational value–such as signals of potential dangers or benefits–is certainly of crucial importance. Moreover, converging evidence indicates that a benefit for significant stimuli can emerge not only at later, post-perceptual stages, but already the perceptual representation of these stimuli can be enhanced by rapid attentional prioritization (for reviews, see e.g., Pourtois et al., 2013; Vuilleumier, 2015). So far, most of the research conducted on the field of prioritized processing of affective information has focused solely on the visual modality; and to date, results from this specific research domain dominate our thinking and understanding on affective attentional biases in general. However, in real life, different senses determine and influence our attentional processes, thus investigating affective attentional processes outside of vision is of special importance. In particular, the specific characteristics of auditory perception and attention make it a good candidate to understand affective attentional biases outside of the prevailing visual domain. For example, spatial characteristics of auditory perception differ crucially from vision: While vision has a limited spatial focus and it closely depends on the position of the head and eyes, auditory perception is less dependent on the spatial relation and distance of the sound source and the perceiver. Furthermore, hearing is more "obligatory" in the sense that we cannot control easily at the sensory periphery whether receiving a sensory input or not; and in turn, while audition receives omnidirectional, transient, and simultaneous stimulation in a more "obligatory" manner than vision, selecting important information cannot be achieved comparably effectively at the level of periphery as in vision where we can direct the fovea to the relevant information (see e.g., Spence and Driver, 1994). However, contributing to its "obligatory" nature, auditory information is processed and organized more extensively already before it reaches the cortex (e.g., King and Nelken, 2009); and surprisingly complex perceptual-cognitive processes take place pre-attentively including predictive modelling of the acoustic environment, that allows for extrapolation and prediction of the transient auditory input and

☆ This research was conducted within the International Research Training Group "Adaptive Minds" supported by the German Research Foundation [GRK 1457]. The reported research was conducted as part of TF's dissertation project.

* Corresponding author at: Saarland University, Department of Psychology, Campus A2.4, D-66123 Saarbruecken, Germany.
E-mail address: t.folyi@mx.uni-saarland.de (T. Folyi).

https://doi.org/10.1016/j.ijpsycho.2019.01.007
Received 3 August 2018; Received in revised form 21 January 2019; Accepted 22 January 2019
Available online 25 January 2019
0167-8760/

rapid detection of significant changes from this internal model (e.g., Bendixen et al., 2012; Horváth et al., 2001; Näätänen et al., 2010). In line with these characteristics, without doubt, audition has a great importance in monitoring our environment, and, above that, pre-attentive auditory processing is often considered as an "early warning system" that can register and prioritize initially unattended significant information without several constrains of vision (e.g., Murphy et al., 2013; Winkler et al., 2003).

In line with the assumption of a rapid "warning system" for significant signals in the auditory modality, in a previous study (Folyi et al., 2016), we have demonstrated that sensory encoding of task-irrelevant valenced tones are facilitated already at an early level of auditory processing as indicated by enhanced N1 ERP-response to this tones within about 100 ms following tone onset. ERP-responses to positive, negative, and neutral tones were recorded in a context where these tones were completely task-irrelevant and participants focused their attention to a concurrent auditory perceptual task. Notably, we associated positive and negative affective meaning to pure tones in a previous learning phase by the means of monetary rewards and losses in a balanced design, thus, across the sample, sensory representation of positive and negative tones was enhanced relative to the physically identical tones with neutral meaning. Thus, as we used both positive and negative valence, we could target the question whether both positive and negative tones are prioritized by rapid auditory attention given their high motivational relevance that was learned in a previous context (in the visual domain, see, e.g., Brosch et al., 2008; Müller et al., 2016; Wentura et al., 2014) or whether our auditory attentional system is tuned to selectively prioritize the negative valence category, in line with findings primarily from the visual domain, based on the assumption that detection of threat possesses arguably high survival value (e.g., Méndez-Bértolo et al., 2016; Öhman, 2005; Öhman et al., 2012; however, in the auditory modality, see Pinheiro et al., 2017, for preferential prioritization of positive vocal cues). While both implicit and explicit behavioral measures differed between valence categories, we did not find a difference between positive and negative valence conditions in the auditory N1-time range. This result is in accordance with a general relevance principle: it is the motivational relevance of the stimulus that possesses attention-grabbing power at the early level of sound encoding rather than a specific valence category.

Relating to the valence learning in our design, it is important to note that facilitated encoding of to-be-ignored valenced tones emerged after a relatively short learning phase by the means of associating admittedly mild rewards and losses to pure tones, indicating that already brief previous experiences can be sufficient to induce changes in the perceptual encoding of tones (as opposed to early preferential processing of stimuli with intrinsic physical or emotional salience). In this regard, findings from the visual domain are also of relevance that suggest that reward-associated stimuli (thus, typically only positive stimuli, but see e.g., Wentura et al., 2014) can capture attention even when they are irrelevant to or conflicting with the current goals (e.g., Anderson et al., 2011; Bucker and Theeuwes, 2016; Le Pelley et al., 2015; for a reviews, see Anderson, 2016a; Le Pelley et al., 2016). Note, that these studies typically used behavioral measures of attentional allocation, and in the auditory modality, our study was the first to demonstrate directly that prioritization of a priori neutral stimuli is possible based on their reward and loss-history already at a perceptual level of stimulus encoding within 100 ms following tone onset (for behavioral distraction effect by previously reward-associated sounds, see Anderson, 2016b; Asutay and Västfjäll, 2016). Although the topic of value-driven attention was mostly investigated in the visual domain, the issue of reward and loss-related "attentional capture" (thus, attentional prioritization independently of or against current goals) in the auditory modality is of utmost importance for real life. For example, considering the general characteristics of audition (e.g., its relative spatial-independence and "obligatory" nature), it is plausible that maladaptive attentional biases to stimuli with learned motivational value (e.g., sound of a slot machine or message ringtone in the case of gambling or social media addiction; or the buzzing of a dentist drill in the case of phobic fear) can be even more intrusive in the auditory world as compared with vision.

In the present study, we targeted two specific questions that test for the generality of the early auditory relevance effect: First, is this relevance effect non-modifiable by general motivational states (as found for comparable effects in the visual domain)? Second, can the early auditory relevance effect be found in the non-attended channel even if attention is more strictly focused on the attended channel (that is, if slips of attention are unlikely)?

1.1. Affective counter-regulation?

The counter-regulation principle of affective attentional biases (Rothermund et al., 2008) proposes that in order to prevent escalation or perseveration of current affective-motivational states, attentional biases to valent stimuli operate incongruently to the current motivational-emotional orientations. Thus, the counter-regulation principle predicts a bias for positive information in negative affective-motivational context and vice versa. This assumption has received numerous empirical support in the visual domain by using behavioral measures of attention allocation (e.g., Rothermund et al., 2011; Rothermund et al., 2008; Schwager and Rothermund, 2013, 2014; Wentura et al., 2018; Wentura et al., 2009).

Wentura et al. (2018), for example, conducted a study that has some resemblance with our earlier experiment and can help to elucidate the counter-regulation principle. In their study, colors were associated with valence in a learning phase. Subsequently, attentional capture/maintenance effects were tested in an additional singleton task. That is, in visual search displays, participants had to quickly categorize a target stimulus while ignoring distractors. If one of the distractors is colored, responses are slower (Theeuwes, 1992), an effect which is explained by attentional capture of the salient color. This effect was more pronounced for positive and negative colors (see also Wentura et al., 2014). In this recent study (Wentura et al., 2018), blocks of the additional singleton task were motivationally charged: If participants could win additional money by a good performance (but not lose any money; positive outcome focus), relatively the attentional effect of the negative color increased; if participants could lose additional money by a bad performance (but not win any money; negative outcome focus), relatively the attentional effect of the positive color increased.

It is an empirical question whether this principle can be found in our setting as well, since there are two important differences between the study by Wentura et al. (2018) and our earlier experiment (Folyi et al., 2016). First, it might be that the auditory modality is different from the visual modality, and in line with its general characteristics described above, it might support an "early warning" for all relevant stimuli instead of a situational flexibility. Second and potentially more important, Wentura et al. (2018) discussed their effect not in terms of early attentional enhancement of valent stimuli but in terms of increased attentional dwelling on valent stimuli. Thus, it might be that early enhancement effects are non-modifiable by top-down settings such as general motivational states whereas later processing stages are modifiable by top-down settings.

Hence, in the present Experiment 1, we introduced a new manipulation in order to promote unequivocal motivational focus: By the prospect of monetary reward and monetary loss, we promoted a motivational focus of anticipating positive and negative future outcomes, respectively. If flexible affective attentional biases operate in early auditory attention, based on the motivational counter-regulation principle we can expect that in salient positive (negative) motivational contexts negative (positive) tones receive differential attentional enhancement. Importantly, the present Experiment 1 is the first investigation of flexible affective attentional biases (a) outside of the visual domain, and (b) at level of very early attentional biases in general, thus, at the stage of sensory encoding of valenced stimuli.

1.2. Slips of attention?

The second test of the generality of the relevance effect relates to the involuntariness of the early attentional enhancement of valenced tones. As outlined above, the valent tones in our earlier study (Folyi et al., 2016) were task-irrelevant and were presented on the non-target channel. Thus, the task-irrelevant nature of the valenced tones and the early time course of the effect suggested that the attentional enhancement for valenced tones, at least partly, reflects involuntary attentional processes. However, given the general characteristics of the selective listening paradigm (i.e., non-continuous attentional load on the task-relevant channel), and more specific characteristic of our design (i.e., relatively long inter-stimulus-intervals), we cannot preclude the possibility that participants switched their voluntary attention between the task-relevant and the task-irrelevant channels and thus it is possible that some of the task-irrelevant tones were already in the focus of voluntary attention. That is, strictly speaking, the relevance effect might be partly due to slips of attention from the target channel to the non-target channel (Lachter et al., 2004).

Therefore, in the present Experiment 2, we increased the demands for continuous voluntary attentional selection of the task-relevant channel by employing continuous task-relevant stimulation. More specifically, we presented a continuous white noise mask in the task-relevant channel, and participants' task was to detect infrequent slightly louder target noise bursts in the white noise background that were perceived as slight, abrupt loudness increments. If early attentional enhancement for valenced tones emerges under the conditions when participants focus their attention continuously to the concurrent task-relevant stimulation, we can conclude that it reflects an involuntary “attentional capture”.

1.3. Overview

The buildup of the present experiments was highly similar to that of Folyi et al. (2016), with the exception of the specific additional manipulations of the present experiments. Fig. 1 gives an overview on the main phases of Experiment 1 and 2. Both experiments consisted of two main parts: a valence induction phase (Fig. 1A) and a test phase (Fig. 1B). In both experiments, during the valence induction task (Fig. 1A), we used a game paradigm to assign positive, negative, and neutral meaning to tone-frequencies in a balanced design by the means of monetary rewards and punishments. In a subsequent test phase (Fig. 1B), 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 performed a perceptual detection task, while positive, negative, and neutral tones were presented concurrently on the task-irrelevant channel. ERPs elicited by these task-irrelevant tones were of the most interest in the present experiment. More specifically, we expected to replicate an increased N1 for positive/negative tones in comparison to neutral tones.

In Experiment 1, we additionally introduced a motivational focus manipulation by assigning experimental blocks of this selective listening task with the chance of monetary reward (positive outcome focus condition), and with the danger of monetary loss (negative outcome focus condition). In Experiment 2, by using a continuous task-relevant stimulation (an additional noise mask on the task-relevant channel), we increased the demands for voluntary attentional selection of the task-relevant channel.

2. Experiment 1

In Experiment 1, we introduced positive and negative motivational foci concerning future outcomes by assigning experimental blocks of the selective listening task with the chance of monetary reward in the positive outcome condition, and with the danger of monetary loss in the negative outcome condition. In line with our previous results, we expected an early attentional enhancement for valenced tones, and we expected that this early attentional effect will be reflected in an enhancement of the N1-amplitude. Regarding the differentiation between positive and negative valence we had two specific hypotheses: If a counter-regulation principle (see e.g., Rothermund et al., 2008) operates on the attentional biases to valent information at the early stage of sound encoding, we can expect enhanced attention (reflected in enhanced N1) for positive compared with negative tones when anticipating negative future outcomes, and in turn enhanced attention (reflected in enhanced N1) for negative compared with positive tones when anticipating positive future outcomes. However, if positive and negative tones are facilitated by early attention in an undifferentiated way (i.e., in line with a rather fixed general relevance hypothesis), we can expect a similar pattern of results as in our former study, that is, enhanced attention to valenced tones in general (thus positive and negative tones together) compared with neutral ones without moderation by motivational outcome focus.

2.1. Method

2.1.1. Participants

Twenty-four students from Saarland University (11 females; aged 18–29 years, $Mdn = 23$ years; two left-handers) participated for monetary compensation. Given the sample size of $N = 24$, and an α-value of .05 (one-tailed), the effect size of $d_Z = 0.58$ (representing the most relevant valenced-minus-neutral difference – $M = -0.64\,\mu V$, $SD = 1.11\,\mu V$ – on the N1-amplitude in our previous study; Folyi et al., 2016) can be detected with a probability of $1 - \beta = .86$ (calculated with the aid of G*Power 3 software; Faul et al., 2007). All participants were native German 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 €30.5 (a fixed amount of €4 for each half-hour of the preparation preceding the experiment, i.e., electrode application; the rest of the compensation was partly dependent on performance; see below Materials and procedure).

2.1.2. Design

We applied a 2 × 3 design with outcome focus (positive, negative) and valence (positive, neutral, negative) as within-participants factors. Tone-frequency-to-valence assignment was counterbalanced between participants in a Latin square design.

2.1.3. Materials and procedure

Unless explicitly noted, everything was equivalent to Materials and procedure in Folyi et al. (2016). An experimental session lasted about 3.5–4 h, including electrode application and removal. During the experiment, participants were comfortably seated in a sound-attenuated room. Sinusoidal tones with three different tone-frequencies (300 Hz, 510 Hz, and 867 Hz) were presented via headphones (HD-600, Sennheiser, Wedemark, Germany) throughout the experiment. In the valence induction task, an additional fourth tone with tone-frequency of 1473.9 Hz was presented as no-go signal (see below). The maximal intensity level was 45 dB sensation level (SL, above individual hearing level referred to a 1000 Hz sinusoidal tone[1]).

The valence induction phase (Fig. 1A) consisted of one practice and three experimental blocks of the valence induction task. In the valence induction task, participants could win or lose money depending on their performance of detecting target tones. Therefore, at the beginning of the experiment, participants received €11 as an initial payment for the

[1] 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.

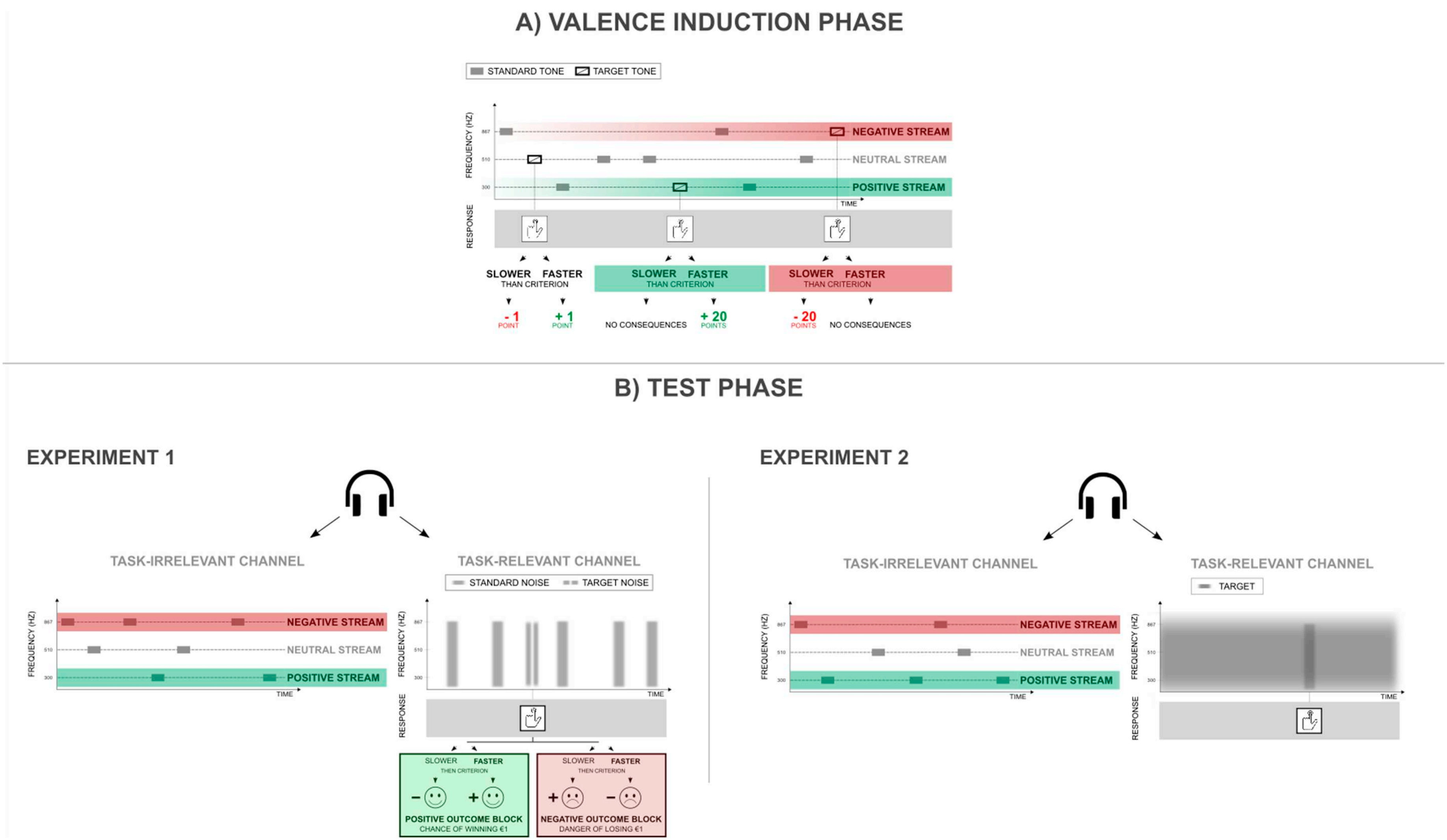

Fig. 1. Schematic illustration of the design of Experiment 1 and 2. Note that only one instance of the counterbalancing scheme is depicted. (A) Valence was induced experimentally in a game paradigm in the valence induction phase. (B) In the selective listening task of the test phase, participants had to detect target noise bursts in the task-relevant channel, while ignoring the positive, neutral and negative tones that were presented in the task-irrelevant channel. In Experiment 1, additionally, positive and negative outcome focus was introduced: Positive outcome blocks were associated with the chance of monetary reward, while negative outcome blocks were associated with the danger of monetary loss. In Experiment 2, a continuous white noise was presented in the task-relevant channel. Participants' task was to detect the target noise bursts that were perceived as slight and abrupt loudness increments in the white noise background.

valence induction phase, but they were obliged to risk the money as the stakes in a "game". Every valence induction block started with a score of zero. Participants were informed that they would immediately win €1 if the final total score of the block was zero or above, and they would immediately lose €1 if the score was below zero. At the end of each valence induction block, a feedback indicated the final total score and the actual outcome of the block. Together with the five valence induction blocks in the test phase (see below), there was a chance of winning or losing up to €8 in the valence induction task.

A valence induction trial started with the presentation of a black fixation cross with a randomized duration of 500–1000 ms, followed by one of the three tones. Tone duration was 1400 ms, and tones were presented binaurally in two versions: Half of the trials featured standard tones, that is, tones were presented with constant intensity (0.5 × maximal intensity; ~39 dB SL), and half of the trials featured target tones, that is, 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. Participants were instructed to respond only to the target tones by pressing the space bar as quickly as possible. A fast 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.[2] Critically, 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, termed as "chance" tone in the instructions), 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, termed as "danger" tone), while a further tone-frequency was associated with a negligible gain or loss of one point in case of either success or failure (neutral tone). Participants were explicitly informed about these possible outcomes before the experiment. To support contingency learning, visual feedback was given after each target trial and in case of a false alarm, with the feedback indicating the type of the tone, the consequences of the recent response, and the current total score.[3] A valence induction block comprised 42 valence induction trials featuring 14 positive, 14 negative, 14 neutral tones (half of them in their standard, and half of them in their target version, respectively) presented in a random order. Additionally, we presented 14 no-go tones (with a frequency of 1473.9 Hz, i.e., the highest tone) additionally on each valence induction block in a random order (half of them in standard, half of them in target version). The no-go tone required participants to withhold their response even when they were presented in their target form; and it was introduced to make the tone-frequency a task-relevant feature during valence induction. Responding to the no-go tone resulted in a loss of 20 points. Furthermore, seven additional visual detection trials were presented in each valence induction block 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; and 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.

The test phase (Fig. 1B) comprised one practice and twelve blocks of the selective listening task with an additional block of the valence induction task interspersed after every two blocks of selective listening (we increased the duration of the test phase relative to our previous study in order to provide a sufficient number of trials per condition for the averaged ERPs after introducing the outcome focus manipulation). During the selective listening task, participants received different stimulation to each ear: They were instructed to attend the auditory stimulation presented to one ear (task-relevant channel), while ignoring the tones presented to the other ear (task-irrelevant channel). Task channel-to-ear assignment was counterbalanced between participants. The task-irrelevant tones were identical to the standard version of the positive, negative, and neutral tones of the valence induction phase, except that their duration was reduced to 800 ms and their intensity was approximately 37 dB SL, and they were presented monaurally with a randomized ISI of 1000–1333 ms. ERPs elicited by these to-be-ignored tones were of the main interest in our study. In the task-relevant channel, white noise bursts were presented with an overall duration of 500 ms and intensity level of approximately 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. 18 target noise bursts and 54 standard noise bursts were presented monaurally in a randomized order in each block, with a random interstimulus interval (ISI) of 2500–3500 ms. The task-relevant and task-irrelevant series of auditory stimulation were allowed to overlap at random temporal positions. Participants had to respond to the target noise bursts as quickly and accurately as possible by pressing the space bar. During the selective listening task, a central fixation cross was presented visually. Participants were instructed to maintain a central fixation during the selective listening task in order to reduce eye-movement artifacts. After each block, visual feedback about the mean hit rate (HR) was presented in order to motivate participants to maintain a high level of performance.

Deviating from Folyi et al. (2016), we introduced an outcome focus manipulation in the selective listening task. The procedure was oriented on Wentura et al. (2018). Positive versus negative outcome focus was varied block-wise in the following way: Half of the selective listening blocks featured the motivational character of positive outcome focus, half of the selective listening blocks featured the motivational character of negative outcome focus. After two consecutive blocks of the same outcome focus, the opposite outcome focus was applied on the subsequent selective listening block. We counterbalanced between participants whether the test phase started with positive or negative outcome focus. Participants were instructed to respond only to the target noise bursts (i.e., noise bursts including a brief "gap") by pressing the space bar as quickly and accurately as possible, while ignoring the task-irrelevant stimulation. In the positive and negative outcome blocks, the motivational focus of anticipating positive and negative future outcomes was introduced by the prospect of substantial monetary reward and substantial monetary loss, respectively. Therefore, participants received €11 as an initial payment for the test phase, again with the obligation of risking the money as the stakes during the task. Depending on their performance, participants could collect "smileys" (depicted on the screen) in the positive outcome blocks, and get rid of "frownies" in the negative outcome blocks. Every selective listening block started with an initial score of twelve "smileys" or "frownies". Participants were informed that they would immediately win €1 if the final score of "smileys" was more than twelve at the end of a positive outcome block. If the final score of "smileys" was twelve or below, there were no negative consequences. However, participants immediately lost €1 if the final score of "frownies" was more than twelve in a negative outcome block. If the final number of "frownies" was twelve or below, there were no positive consequences. Thus, there was a chance of winning or losing up to €6 in the selective listening task. Fast detection of a target 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 five trials (see e.g., Rothermund et al.,

[2] To ensure that participants experience success and failure trials in a relatively balanced amount, the response-speed criterion on a given trial was defined by the moving median of the preceding six trials weighted by participant's current game score: median′ = median − 0.2 × current total game score (see also Folyi et al., 2016; Rothermund et al., 2008).

[3] 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). These additional sounds were presented only in the valence induction phase.

2008).[4] The outcome focus of the block determined the consequences of success or failure: In a positive outcome block, fast detection of a target noise resulted in increasing the number of "smileys" by one, thereby increasing participants' chance to win monetary reward at the end of the block. In case of a failure, the number of "smileys" was reduced by one. On the contrary, in a negative outcome block, fast detection of a target noise resulted in decreasing the number of "frownies" by one, thereby increasing the possibility to avoid monetary loss at the end of the block. In case of a failure, the number of "frownies" was increased by one, thereby increasing the possibility for monetary loss. Visual feedback was given after each target trial and in case of a false alarm, with the feedback indicating the outcome (e.g., "slow"; "missed target") and the consequences of the recent (non-)response (± 1 "smiley"/"frowny"). Additionally, the feedback included the current total (i.e., the number of smileys or frownies) in the following form: Possible positions of smileys or frownies were arranged into two 3 × 4 matrices that were separated by a blank line, thus, the critical value of twelve smileys or frownies was highlighted by the visual arrangement. After the feedback, participants could start the next trial by pressing the space bar. At the end of each selective listening block, a further feedback indicated the final total score of smileys or frownies and the actual outcome of the block (thus, plus or minus €1, or no consequences).

Before the actual experiment started, participants had to accomplish two preliminary tasks in order to control for their ability to discriminate the tone-frequencies and understanding and learning the associations between tone-frequencies and gain/loss odds. Accordingly, the first task required participants to discriminate the three tones based on their tone-frequencies, the second required them to learn the associations between tone-frequencies and the gain/loss odds (i.e., the three meanings during the valence induction: "danger", "chance", "neutral" and "no-go" tones). Before the second discrimination task, participants were informed about the possible outcomes related to each tone in the valence induction phase. The discrimination tasks terminated when the accuracy level exceeded 95% in the first task, and 90% in the second task.

2.1.4. EEG recording and analysis

Again, unless explicitly noted, everything was equivalent to Folyi et al. (2016). The EEG was recorded only during the test phase from 64 scalp locations (following the international 10–10 system, and including left and right mastoids). The common reference electrode was placed on the tip of the nose. 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. We segmented the continuous EEG into 800 ms long epochs (each including a 200 ms long pre-stimulus baseline).[5] Epochs contaminated with severe artifacts were rejected based on automatic artifact detection (rejection criteria: signal range exceeding 200 μV, or voltage step exceeding 50 μV/ms, or a voltage difference exceeding 150 μV on any channel including EOG; no further eye movement correction or artifact rejection based on manual data inspection was performed) and epochs containing task-relevant noise bursts were discarded (rejection criteria: onset of a task-relevant standard noise burst 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 noise burst within a −1000 ms to +800 ms time-window relative to the onset of the task-irrelevant tone). Epochs were baseline corrected using the 200-ms pre-stimulus interval and epochs were averaged separately for the different conditions. In the positive outcome condition, on average ERPs were based on 85 (SD = 14), 84 (SD = 16), and 79 (SD = 20) trials per participant in the positive, neutral, and negative valence conditions, respectively. In the negative outcome condition, on average ERPs were based on 86 (SD = 19), 84 (SD = 19), and 78 (SD = 19) trials per participant in the positive, neutral, and negative valence conditions, respectively.

We formed a region of interest (ROI) from frontocentral electrode sites (F1, Fz, F2, FC1, FCz, FC2, C1, Cz, C2) according to the auditory N1-literature (for a review, see Näätänen and Picton, 1987), and the frontocentral scalp distribution of the grand average N1. N1-amplitude was measured at the frontocentral ROI as mean voltage in a 20-ms time window centered at the latency of the group-average N1 peak with experimental conditions collapsed (108 ms; for comparable method see, e.g., Gilmore et al., 2009; Horváth et al., 2012; Jacobsen et al., 2003).

2.2. Results

In all analyses, we added counterbalancing group as a between-participants factor to use the correct error term (Pollatsek and Well, 1995; see also Folyi et al., 2016). 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., Folyi et al., 2016; O'Brien and Kaiser, 1985; Rohr et al., 2012). Similarly to our previous study, we applied a priori chosen contrasts that represented our specific hypotheses: (1) for the first contrast, amplitude values were averaged across positive and negative stimuli and contrasted with the neutral stimuli (i.e., testing for the general relevance bias hypothesis). (2) The second orthogonal contrast was the contrast between positive and negative tones (i.e., testing for the negativity bias hypothesis). Results for the behavioral performance in the valence induction and selective listening tasks can be found in Appendix B. A significance level of $\alpha = .05$ (two-tailed, unless otherwise noted) was adopted for all analyses.

The amplitude measure of the auditory N1 elicited by the task-irrelevant positive, negative, and neutral tones during the test phase was of most interest for the present study. Prototypical P1-N1-P2 waveform was clearly observable in the group average ERPs (see Fig. 2 for the group average ERP waveforms to the positive, neutral, and negative tones, and Fig. 1A of the Appendix A for the group average ERP waveforms to the positive, neutral, and negative tones in the positive and negative outcome focus conditions, respectively; for the mean amplitudes for the components of interests, see Table 1). A 2 × 3 × 3 MANOVA for repeated measures with outcome focus (positive, negative) and valence (positive, neutral, negative) as within-participants factors and counterbalancing group as a between-participants factor of the N1-amplitudes yielded no significant valence main effect, $F(2,20) = 1.56$, $p = .235$, $\eta_p^2 = .135$. Replicating Folyi et al. (2016), the a priori contrast of valenced (positive and negative) vs. neutral condition was significant, $F(1,21) = 3.20$, $p = .044$ (one-tailed),[6] $\eta_p^2 = .132$, indicating enhanced N1-amplitudes for valenced compared with neutral tones, whereas the contrast of positive vs. negative conditions was not significant, $F < 1$, *n.s.* Outcome focus did not show a significant main

[4] The actual response-speed criterion (median') was dependent on a participant's current score of "smileys" or "frownies" in the following way: In positive outcome focus blocks: median′ = median − (current total score of "smileys" − 12) ∗ 3, while in the negative outcome focus blocks: median′ = median + (current total score of "frownies" − 12) ∗ 3, thus, leading to current scores that tend to be around the critical value of 12.

[5] In our previous study (Folyi et al., 2016) we used 1000 ms long epoch duration. The reduction was done in Experiment 1 to increase the total number of trials following artifact rejection (within an acceptable duration of the experiment) since a further factor (i.e., outcome focus) was added.

[6] A one-tailed interpretation can be applied given the equivalence of an F-test with $df_N = 1$ with a two-tailed t-test and given our specific prediction; see Maxwell and Delaney (2004, p. 164).

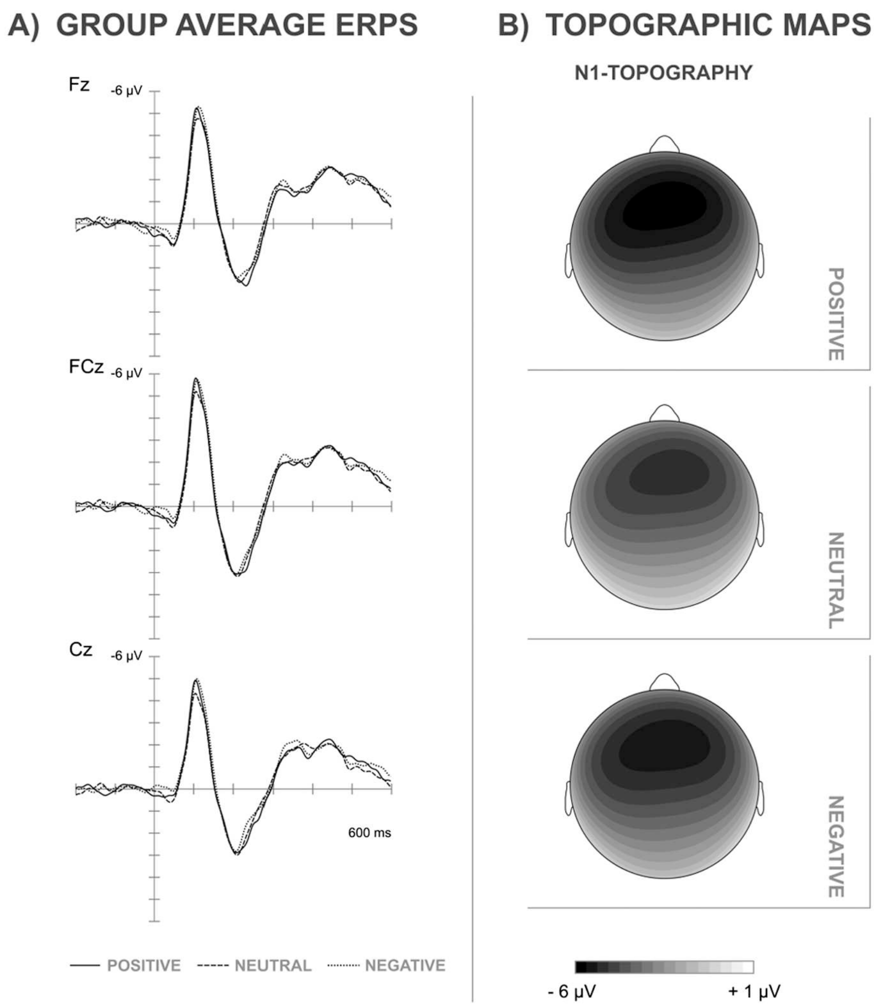

Fig. 2. ERP-results of Experiment 1. (A) Group average ERP waveforms to the positive, neutral, and negative tones on the representative Fz, FCz and Cz electrode sites. We present the ERP results collapsed across the outcome focus conditions, as our results did not show any indication for outcome focus modulation (see Fig. A1 for the complete 2 × 3 design). 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 (98–118 ms) in positive, neutral and negative conditions.

Table 1
ERP results of Experiment 1. Mean N1-amplitudes (in µV) in each valence and outcome focus condition; *SD* in parentheses.

Stimulus valence	Outcome focus		
	Positive	Negative	Mean
Positive	−4.80 (3.43)	−5.18 (3.05)	−4.99 (3.15)
Neutral	−4.40 (2.33)	−4.56 (3.15)	−4.48 (2.64)
Negative	−4.88 (3.54)	−5.10 (3.32)	−4.99 (3.28)

effect, $F(1,21) = 1.29$, $p = .269$, $\eta_p^2 = .058$. Of most interest, there was no indication of an interaction between valence and outcome focus, $F(2,20) = 0.11$, $p = .900$, $\eta_p^2 = .011$ for the overall test, $F(1,21) = 0.093$, $p = .763$, $\eta_p^2 = .004$; for the contrast of valenced vs. neutral condition; and importantly, $F(1,21) = 0.14$, $p = .715$, $\eta_p^2 = .006$; for the contrast of the positive vs. negative condition. The Bayes factor in favor of the null hypothesis corresponding to the latter analysis (i.e., testing the difference between the positive and negative outcome focus conditions on the positive-minus-negative valence difference representing our a priori valence contrast) is $BF_{01} = 4.36$.[7] This can be considered "substantial evidence" for H_0 according to Jeffreys (1961, p. 432; see also Wagenmakers et al., 2011).

2.3. Discussion

Based on behavioral findings from the visual domain (e.g., Rothermund et al., 2008; Schwager and Rothermund, 2014; Wentura et al., 2018, 2009), Experiment 1 tested whether the early attentional enhancement to valenced information follows a motivational counter-regulation principle in the auditory domain. The counter-regulation hypothesis on early auditory attentional biases would predict that an attentional enhancement to valences tones operates incongruently to the current motivational-emotional orientations. If early auditory attentional biases follow a motivational counter-regulation principle, we could expect enhanced attention to positive relative to negative tones in negative affective-motivational context, and enhanced attention to

[7] We used Bayesian one sample *t*-test procedure for testing the difference variables representing our a priori contrasts against zero, and Bayesian ANOVA

(footnote continued)
procedure when testing for moderation by Experiment factor (see below at Across-experiments analyses) in JASP (jasp-stats.org; JASP Team, 2018; Version 0.9.0.1). All Bayesian *t*-tests were performed with Cauchy prior width of 0.707 (see, Rouder et al., 2009); and Bayesian ANOVAs were performed with a scale parameter of r = 0.5 for the effect size prior (for a discussion, see Rouder et al., 2017).

negative relative to positive tones in positive affective-motivational context. Altogether, the results of Experiment 1 do not support the counter-regulation hypothesis of auditory affective attentional biases at the level of early sound encoding. There was no indication for modulation by outcome focus on the early attentional enhancement for valenced tones. Thus, the general pattern of results of Experiment 1 was highly similar to that in our initial study (Folyi et al., 2016): We found a general enhancement of the auditory N1 for affectively significant tones relative to neutral ones, and this enhancement did not show any differentiation between positive and negative valence despite the outcome focus manipulation, providing further support to our claim that the general relevance of the valenced tones that governs early auditory attentional processes.

3. Experiment 2

Both in Folyi et al. (2016) and in the present Experiment 1, we presented isolated stimuli with a relatively low presentation rate on the task-relevant channel of the selective listening task. As under these task conditions participants might have switched their voluntary attention between the task-relevant and the task-irrelevant channels, it is possible that (some of the) task-irrelevant valenced tones were in the focus of voluntary attention to some degree. In line with this argumentation, a line of empirical evidence indicates that the fact that a stimulus was irrelevant to the main task does not necessarily mean that it was also initially unattended as "slips" (i.e., covert shifts) of attention could have occurred toward it (Lachter et al., 2004). Accordingly, characteristics of the selective listening task in our former study and in the present Experiment 1 may have allowed for "slips" of attention between the task-relevant and task-irrelevant channel. Therefore, in Experiment 2, we increased the demands for continuous voluntary attentional selection of the task-relevant channel. Although a high rate of stimulus presentation in the selective listening task supports the attentional selection of the task-relevant channel during selective listening (e.g., Woldorff and Hillyard, 1991), this approach also increases the possibility that onset and offset related neural responses elicited by the task-irrelevant and task-relevant stimuli overlap in time. To overcome this possible issue, we presented a continuous white noise mask in the task-relevant channel, and participants' task was to detect infrequent slightly louder target noise bursts in the white noise background that were perceived as slight loudness increments in the ongoing stimulation. Thus, importantly, successful task performance required constant monitoring of the continuous noise delivered to the task-relevant ear thereby preventing "slips" of attention toward the task-irrelevant positive, negative and neutral tones.

In a strict sense of automaticity, attentional enhancement for affective information occurs as an "attentional capture". In this vein, sensory enhancement of valenced stimuli is assumed to operate at least partly independently of voluntary controlled attentional processes, and it is presumably mediated by subcortical structures, involving the amygdala for stimuli with intrinsic emotional meaning, and the basal ganglia, superior colliculus and possibly the amygdala in the case of reward associations, at least in the visual domain (for reviews, see Pourtois et al., 2013; Vuilleumier, 2015; Vuilleumier and Huang, 2009). The assumption that value-based attention can also operate independently of voluntary attention is supported by recent results from the visual modality indicating that reward-associated stimuli can "capture" attention when they are completely task-irrelevant and even when attending these stimuli conflicts with current goals (e.g., Bourgeois et al., 2017; Le Pelley et al., 2015; Wentura et al., 2014). Consequently, if such a "reflexive" preferential enhancement is elicited by the valenced tones, we can expect an early attentional enhancement for these tones even in the task settings of the test phase in Experiment 2, thus when participants' voluntary attention is strictly devoted to a concurrent task. Alternatively, preferential attention to affectively significant stimuli may depend – at least to some degree – on voluntary attentional processes (in the visual domain, see e.g., Eimer and Holmes, 2007; Eimer et al., 2003). If early enhancement of valenced tones depends on voluntary controlled attentional processes, we can expect that the early valence effect will be abolished or substantially reduced by increasing the attentional demands for concurrent task-relevant selection.

3.1. Method

3.1.1. Participants

Twenty-four students from Saarland University (11 females; aged 18–29 years, *Mdn* = 22 years; three left-handers) participated for monetary compensation. All participants were native German 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 (i.e., they 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 "game", see Design, materials, and procedure; after the first 1.5 h, participants received further €4.50 for each additional half-hour). Since Experiment 1 and 2 were planned in parallel as two follow-ups to Folyi et al. (2016), power planning was the same as for Experiment 1. Admittedly, the effect in Experiment 1 was smaller than the one in Folyi et al. (2016). We will return to this issue in the section of Across-experiments analyses.

3.1.2. Design, materials, and procedure

Design, materials and procedure were essentially the same as in Experiment 1, with the following exceptions.

In the valence induction phase, we only changed a detail. We did not further employ a no-go tone (i.e., tone with the highest tone-frequency) in order to support the acquisition of tone-to-valence associations (i.e., three tone-frequencies can be mapped easily into the more salient representations of "low", "middle", and "high" tones). To still ensure that participants encode tone-frequency, six additional tone-discrimination trials (each of the three tones once in standard and once in target form in each block) replaced the visual detection trials (see Materials and procedure of Experiment 1). Tone-discrimination trials were identical to the experimental trials with the following exceptions: After 400 ms following the onset of the tones, the black fixation cross turned red and remained on the screen during the sound presentation (i.e., 1000 ms long). Participants' task was to choose whether the presented tone was a "danger", "chance", or "neutral" tone by clicking to the corresponding term on the screen. Error feedback was presented visually if the participant responded incorrectly or did not respond within 1500 ms. Failure was associated with a penalty of 10 points.

The test phase (Fig. 1B) comprised ten blocks of the selective listening task with an additional block of the valence induction task interspersed after every two blocks of selective listening. Now, there was no outcome focus manipulation. On the task-relevant channel, a continuous white noise was presented with an intensity level of approximately 37 dB SL. Additionally, 15 target white noise bursts were presented on the task-relevant channel in each selective listening block in random temporal positions with a minimum ISI of 500 ms. Target noise bursts were presented with a duration of 200 ms and intensity level of approximately of 39 dB SL[8] and they were perceived as abrupt loudness increments of the continuous white noise. Participants' task was to detect the target noise bursts as quickly and accurately as possible by pressing the space bar. Thereby, participants had to monitor the task-relevant channel continuously to detect the infrequent target noise bursts that were perceived as slight increments in the loudness of the continuous noise. Visual feedback was given about the mean hit rate

[8] Target intensity was set according to three pilot sessions that aimed to set the loudness of the target noise bursts slightly above the threshold of detectability.

after each block.

Finally, as Folyi et al. (2016) we conducted a manipulation check phase, in which we tested the effectiveness of the valence induction with an auditory Affective Simon Task (AST; Folyi and Wentura, 2017; see Houwer et al., 2001, for the original procedure with visual stimuli). This procedure and the results are presented in Appendix C.

3.1.3. EEG recording and analysis

EEG recording and data preparation was the same as for our former study (Folyi et al., 2016) and epochs were rejected according to the criteria described above (however, note, that in the current study only target and no standard noise bursts were presented on the task-relevant channel). On average ERPs were based on 248 ($SD = 40$), 247 ($SD = 37$), and 249 ($SD = 39$) trials per participant in the positive, neutral, and negative valence conditions, respectively. N1-amplitudes were again measured on the frontocentral ROI (see e.g., Näätänen and Picton, 1987, and above) as mean voltage in a 20-ms window centered at the latency of the group-average peak (144 ms, experimental conditions collapsed).

3.2. Results

Results for the behavioral performance in the valence induction and selective listening tasks can be found in Appendix B.

Prototypical P1-N1-P2 waveform was clearly observable in the group average ERPs (see Fig. 3, and for the mean amplitudes for the components of interests, see Table 2). A 3 × 3 MANOVA for repeated measures with valence (positive, neutral, negative) as within-participants factor and counterbalancing group as between-participants factor of the N1-amplitudes did not yield significant valence main effect, $F(2,20) = 1.38$, $p = .275$, $\eta_p^2 = .121$. The a priori contrast of valenced (positive and negative) vs. neutral condition showed a tendency of a difference, $F(1,21) = 2.23$, $p = .075$ (one-tailed; see Footnote 6), $\eta_p^2 = .096$; with more negative N1-amplitudes in the valenced conditions compared with the neutral condition ($M = -4.33\,\mu V$, $SD = 1.68\,\mu V$ in the valenced condition; $M = -3.98\,\mu V$, $SD = 2.00\,\mu V$ in the neutral condition). The a priori contrast of the two valenced conditions (positive vs. negative) was not significant, $F < 1$, *n.s.*

Table 2
ERP results of Experiment 2. N1-amplitudes, and P1-N1 peak-to-peak amplitudes (in μV) in each valence condition; *SD* in parentheses.

Valence condition	N1	P1-N1
Positive	−4.30 (1.87)	−5.72 (2.14)
Neutral	−3.98 (2.00)	−5.34 (1.69)
Negative	−4.36 (1.75)	−5.88 (1.96)

Since the main analysis yielded an ambiguous result – the valenced vs. neutral contrast for the N1 was non-significant (with $p = .075$) although numerically the pattern fits perfectly to Folyi et al. (2016) and the present Experiment 1, for an additional exploratory analysis, we applied peak-to-peak amplitude measurement for the P1 and N1 peaks (i.e., using the adjacent peak as a reference, e.g., Handy, 2005) to better quantify the apparent relative difference between the P1 and N1 as indicated by the grand average waveform (Fig. 3). The auditory P1 and N1 are typically elicited together, thereby often termed as P1-N1 complex (Burkard et al., 2007); and there is evidence indicating that

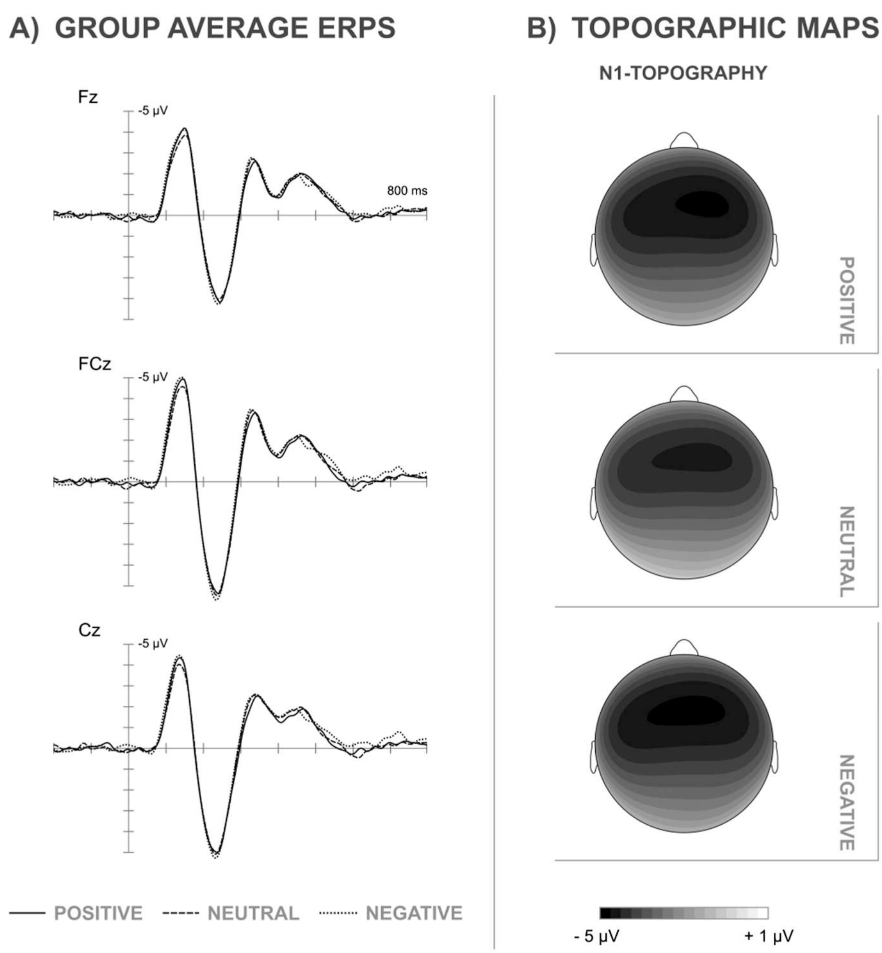

Fig. 3. ERP-results of Experiment 2. (A) Group average ERP waveforms to the positive, neutral, and negative tones on the representative Fz, FCz and Cz electrode sites. The physical onset of the tones is at the crossing of the axes (0 ms). Negative polarity is plotted upwards. (B) Group average topography maps in the N1 time window (134–154 ms) in positive, neutral and negative conditions.

auditory attention can enhance both components at least in the case of highly focused attention (e.g., Woldorff and Hillyard, 1991). Peak-to-peak amplitudes were derived for each participant as the voltage difference between the P1 and N1 peaks (see Table 2; P1 and N1 peaks were identified in the individual ERPs as the strongest positive/negative local peaks within a 40/60 ms long time window centered around the latency of the group-average P1/N1 peaks; 62 ms and 144 ms, respectively). A 3 × 3 MANOVA for repeated measures with valence (positive, neutral, negative) as within-participants factor and counterbalancing group as between-participants factor of the P1-N1 peak-to-peak amplitudes showed a significant a priori contrast of valenced (positive and negative) vs. neutral condition (i.e., representing the general relevance bias hypothesis of affective attention), $F(1,21) = 4.43$ $p = .048$, $\eta_p^2 = .174$, indicating enhanced P1-N1 amplitudes for valenced compared with neutral tones ($M = -5.80\,\mu V$, $SD = 1.86\,\mu V$ in the valenced condition; and $M = -5.34\,\mu V$, $SD = 1.69\,\mu V$ in the neutral condition). The contrast of the two valenced conditions (positive vs. negative; i.e., representing the negativity bias hypothesis of affective attention) was not significant, $F < 1$, *n.s.* ($F(2,20) = 2.74$, $p = .089$, $\eta_p^2 = .215$ for the overall main effect).

3.3. Discussion

Experiment 2 targeted the question whether sensory encoding of valenced tones can be enhanced independently of voluntary attention (i.e., similarly to an independent "emotional" attention, for a review, see e.g., Pourtois et al., 2013). To do so, we increased the demands for voluntary attentional selection of the task-relevant channel to reduce "slips" of attention to the task-irrelevant positive, negative, and neutral tones. To reach this aim, we presented a continuous white noise to the task-relevant ear. In consequence, successful detection of the embedded infrequent target noise bursts required constant monitoring of the continuous white noise. At the same time, positive, negative, and neutral tones were presented to the other, task-irrelevant ear, ensuring that they are not in the focus of voluntary attention in this task design. Of main interest in the present study, under these task conditions, we found only a non-significant ($p = .075$) evidence in our standard analysis on the N1-amplitude for a differential attentional enhancement of valenced compared with neutral tones. However, there is an additional indicator that points toward an early valence-related attentional enhancement: The P1-N1 complex together was enhanced for valenced compared with neutral tones, suggesting a moderate early attentional effect. Although the N1 is a reliable indicator of early attentional enhancement in the auditory ERP (e.g., Herrmann and Knight, 2001), there is evidence that enhancement of auditory P1 can index very early attentional processes (e.g., Fritz et al., 2007; Woldorff and Hillyard, 1991). Thus, P1 and N1 components can be enhanced together in the case of an attentional effect with an early temporal locus. However, although this analysis tests our general hypotheses (i.e., general relevance bias and negativity bias hypothesis of early auditory attention), the measure of the P1-N1 peak-to-peak amplitude was chosen based on an apparent difference of the grand average waveforms. Hence, these results should be treated with some caution and they need further replication (see e.g., Luck and Gaspelin, 2017).

The behavioral results of Experiment 2 (see Appendix B) also support our claim that in this task design participants held their voluntary attention more constantly on the task-relevant channel. In line with our effort to increase task demands, we indeed observed substantially lower accuracy on the selective listening task of Experiment 2 compared with the experiments in which we did not present a task-relevant continuous noise (93.2% in our initial study and 96.1% in Experiment 1[9] while 78.4% in Experiment 2; $F(2,69) = 32.31$ $p < .001$, $\eta_p^2 = .484$, for the comparison of the three experiments; while follow-up comparisons, with Bonferroni-adjusted alpha = .017, revealed significant difference in mean accuracy between the present Experiment 1 and 2, $t(46) = 6.56$, $p < .001$, $d = 1.89$; and our initial study and the present Experiment 2, $t(46) = 5.23$, $p < .001$, $d = 1.51$; while the comparison of our initial study and the present Experiment 1 fell short of significance, $t(46) = -2.45$, $p = .018$, $d = 0.71$). Moreover, despite the perceptually more challenging task, mean RT was considerably shorter in Experiment 2 (590 ms our initial study and 502 ms in Experiment 1, while 367 ms in Experiment 2; $F(2,69) = 114.47$, $p < .001$, $\eta_p^2 = .768$, for the comparison of three experiments; while the post-hoc comparisons, with Bonferroni-adjusted alpha = .017, revealed significant difference in mean RT between the present Experiment 1 and 2, $t(46) = 13.55$, $p < .001$, $d = 3.91$; and our initial study and the present Experiment 2, $t(46) = 13.96$, $p < .001$, $d = 4.03$; the comparison of our initial study and the present Experiment 1 was also significant, $t(46) = 5.00$, $p < .001$, $d = 1.44$), with markedly lower standard deviation (74 ms in our initial study and 44 ms in Experiment 1, versus 24 ms in Experiment 2). This pattern is consistent with the interpretation that while in Experiment 2 participants focused their attention constantly to the continuous stimulation delivered to the task-relevant ear, in our initial study, however, participants' attentional focus might have been on the task-irrelevant channel on some of the trials, and prolonged RTs reflect the time cost of shifting attention back to the task-relevant channel when a target was presented. The intermediate values for the present Experiment 1 might reflect that participants attended the task-relevant channel also more continuously in order to achieve better performance in line with their higher task motivation (we will further discuss this point below).

As we found an admittedly weak valence effect in our standard analysis of Experiment 2, the question arises whether the current pattern of results supports an interpretation that in our initial study voluntary attentional processes likely contributed to the N1-enhancement for task-irrelevant valenced tones, and only a weaker involuntary enhancement can emerge when "slips" of voluntary attention are prevented. In other words, in a situation when "slips" of attention to the task-irrelevant channel are possible, there might be an additive effect of voluntary and involuntary attention on early auditory processing of valenced tones (for a review on additive effects of voluntary and more automatic "emotional" attention, see Pourtois et al., 2013). When "slips" are prevented, only involuntary attentional processes could contribute to the results.

Moreover, the finding of a numerically weaker valence effect on the N1-amplitudes relative to our initial study holds true also for Experiment 1. One possible explanation for this outcome is that increasing the complexity of the design by introducing outcome focus manipulation reduced test power for detecting valence differences. A further possibility, as already mentioned above, is that participants might have voluntarily attended the task-relevant channel more constantly in Experiment 1 relative to our initial experiment in order to achieve a high performance and thereby ensure better monetary outcome. As a consequence, similarly to Experiment 2, we might have decreased a possible contribution of voluntary attention to the early enhancement of valence tones in Experiment 1.

To target these issues, we conducted a combined analysis across our previous study (Folyi et al., 2016) and the present Experiment 1 and 2. Besides clarifying the question of potential differences between experiments (i.e. whether potentially different attentional processes could contribute to the sensory enhancement of the valenced tones), a combined analysis allows a better estimate of the average valence effect and better statistical base for rejecting the hypotheses of N1-differences between positive and negative valence.

[9] However, note that in Experiment 1 participants' motivation for successful performance on the selective listening task was increased by additional performance-dependent monetary rewards and losses.

4. Across-experiments analyses

We conducted a 3 × 3 MANOVA for repeated measures with valence (positive, neutral, negative) as within-participants factor and experiments (our initial study, Folyi et al., 2016, and Experiment 1 and 2 of the current study) as a between-participants factor of the N1-amplitudes (each participant took part only in one experiment). Beside a significant experiment main effect, $F(2,69) = 5.69$, $p = .005$, $\eta_p^2 = .142$ (lowest N1-amplitudes in our initial experiment), the analysis yielded a significant valence main effect, $F(2,68) = 5.77$, $p = .005$, $\eta_p^2 = .145$. The a priori planned contrast of valenced (positive and negative) vs. neutral condition was significant, $F(1,69) = 11.70$, $p = .001$, $\eta_p^2 = .145$, indicating enhanced N1-amplitudes in the valenced conditions compared with the neutral condition across experiments. The Bayes factor in favor of our directed hypothesis is $BF_{10} = 53.21$. This can be considered "a very strong evidence" according to Jeffreys (1961, p. 432; see also Wagenmakers et al., 2011). The a priori contrast of the two valenced conditions (positive vs. negative) did not show any differences, $F(1,69) = 0.10$, $p = .750$, $\eta_p^2 = .001$. The Bayes factor in favor of the null hypothesis is $BF_{01} = 7.34$. This can be considered "substantial evidence" for the H_0. N1-differences based on our a priori planned comparisons for all three experiments are depicted in Fig. 4.

Importantly, there was no indication for an interaction between valence and experiment, $F(4,138) = 0.26$, $p = .904$, $\eta_p^2 = .007$, indicating a homogenous pattern of valence results across the three experiments. Accordingly, the a priori planned contrast of valenced (positive and negative) vs. neutral condition was not moderated by experiment, $F(2,69) = 0.33$, $p = .723$, $\eta_p^2 = .009$. The Bayes factor in favor of the null hypothesis is $BF_{01} = 6.62$. This can be considered "substantial evidence" for H_0. Similarly, the contrast of the two valenced conditions was also not moderated by experiment, $F(2,69) = 0.23$, $p = .799$, $\eta_p^2 = .006$. The Bayes factor in favor of the null hypothesis is $BF_{01} = 7.14$. Again, this can be considered "substantial evidence" for H_0.

Additionally, we calculated weighted averages of the effect sizes from our three experiments (by using Exploratory Software for Confidence Intervals, ESCI; Cumming, 2011) for the effects of most interest: For the valenced (positive and negative)-minus-neutral difference it resulted an overall effect size of $d_z = 0.41$ (95% CI: 0.17 to 0.65); while the overall effect size associated with the positive-minus-negative difference was $d_z = 0.06$ (95% CI: -0.17 to 0.29).

5. General discussion

Taken together, the results of Experiment 1 and 2 are in accordance with our previous study indicating early, involuntary attentional enhancement for valenced tones in general without differentiation between positive and negative valence. Although the attentional effect for valence tones on the N1 amplitude appeared admittedly weaker in the single Experiment 1 and 2 compared with our initial study, a combined analysis of these experiments showed a clear valence effect across the three experiments. Taken together, the across-experiments analysis showed a clear-cut pattern: (1) An early attentional effect for valenced compared with neutral tones as reflected in enhanced N1 amplitudes for these tones; (2) no indication of a difference between positive and negative valences. Importantly, this pattern of findings was consistent across the three experiments as there was no indication for an interaction between the factor of valence and experiment. As these single experiments employed different task-relevant attentional and motivational demands concurrent with the task-irrelevant positive, negative and neutral tones, the absence of moderation by experiment factor (i.e., a relative "immunity" to concurrent task-demands) gives support to the interpretation that early attentional enhancement for valenced tones can occur independently of voluntary attention. Thus, although we cannot rule out that voluntary attention took effect in our initial study, based on the present results, it is not needed to assume a necessary contribution of voluntary attentional processes to an early facilitation of valenced tones.

Thus, the pattern of results concerning affective attentional biases at the level of sound encoding appeared unequivocal: In three experiments, we found only indications for valenced versus neutral differential effect at the early stage of sound encoding, but we found no indication for preferential attention to a specific valence category over another (e.g., a negativity bias). Consequently, our results suggest that it is the general relevance of the valenced tones that governs early attentional processes rather than the priority of negative valence specifically. While the early attentional effect on the N1-amplitude was comparable for positive and negative tones, behavioral measures indicated differentiation between valence categories (see Appendix B). Behavioral measures of the valence induction phase in both Experiment 1 and 2 showed a differentiation between positive and negative valence as reflected in more false alarms for positive compared with negative tones, indicating that participants strategically differentiated the valence conditions in order to maximize their monetary outcome. Besides, participants could identify the valence-related meaning of the tones with high precision on the tone-discrimination trials of Experiment 2, indicating successful contingency learning. Additionally, we administered a manipulation check phase in Experiment 2. By using an auditory AST for accessing implicit evaluations, we found evidence that simple tones had indeed acquired positive and negative affective valence during the valence induction (see Appendix C).

Additionally, the across-experiments analysis on the N1 amplitude revealed a general between-experiments difference: The N1-response had generally higher amplitude in the present two experiments compared with our former study (Folyi et al., 2016). Nonetheless, importantly, as there was no moderation of the valence effect by experiment, the differential enhancement for valenced relative to neutral tones was independent of this general enhancement. A tentative

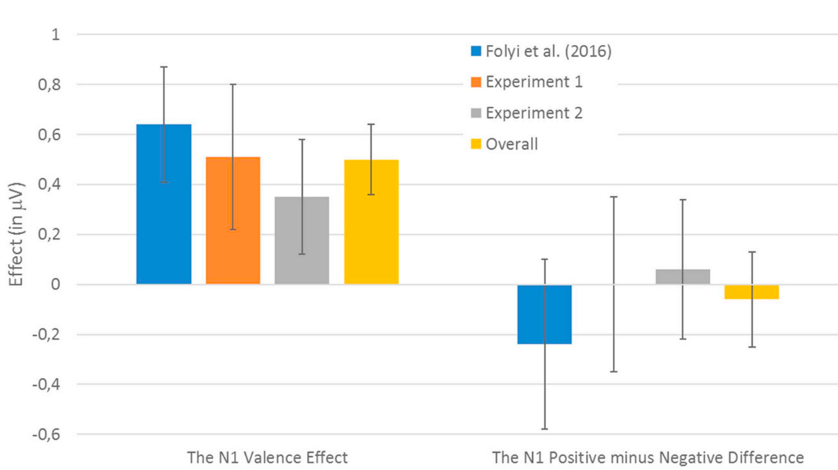

Fig. 4. For all studies (Folyi et al., 2016 and the present Experiments 1 and 2), difference on N1-amplitudes (in µV) representing the two specific hypotheses: the neutral-minus-valenced (averaged across positive and negative conditions) difference representing the general relevance bias hypothesis; and the positive-minus-negative difference representing the negativity bias hypothesis. Error bars depict one standard error above/below the mean.

explanation is that an increase of nonspecific arousal or alertness–due to generally increased motivational state by performance-related rewards and losses (Experiment 1) and by increased task difficulty (Experiment 2)–resulted in generally enhanced auditory responses in the present study (see e.g., Näätänen and Picton, 1987). Furthermore, recent findings suggest that auditory selective attention can enhance feature selectivity independently of non-specific gain modulations through rapid attention-driven short-term plasticity (e.g., Ahveninen et al., 2011). In the present paradigm, selective tuning to the low-level features of reward- and loss-associated tones might have taken place independently of a nonspecific sensory gain modulation.

To conclude, the present study provided evidence on a pooled sample of 72 participants that valenced tones can receive attentional enhancement at a perceptual stage of sound encoding. These results suggest that this early attentional enhancement for valenced auditory stimuli can emerge involuntarily. Hence, our results give support to the notion that prioritized processing of auditory sensory input is not a fixed function of their physical salience or a priori meaning (e.g., "evolutionary preparedness" for example for threat-related vocalizations or associations with such a "prepared" emotional stimuli), but already brief previous experiences with–-admittedly mild–losses and gains are sufficient to induce changes in the perceptual encoding of tones. While explicit and implicit behavioral measures differentiated between positive and negative valence, our ERP-results suggest that the general relevance of the motivationally significant tones, irrespective of their specific valence category, governs attentional processes at the early level of sound encoding. The exact underlying mechanisms of this rapid "relevance signal" cannot be specified in the present paradigm; however, in the visual domain, it was demonstrated that reward-history can induce plastic changes in spatial priority maps (Chelazzi et al., 2014). As short-term plasticity in the auditory cortices after shock conditioning and associations with natural emotional sounds was suggested in the auditory domain (Bröckelmann et al., 2013, 2011), tentatively, an analogous short-term plasticity in auditory frequency maps could be a possible mechanism to enable selective tuning already to the low-level features of reward- and loss-associated sounds, independently of general attentional gain modulations (e.g., Ahveninen et al., 2011).

Appendix A

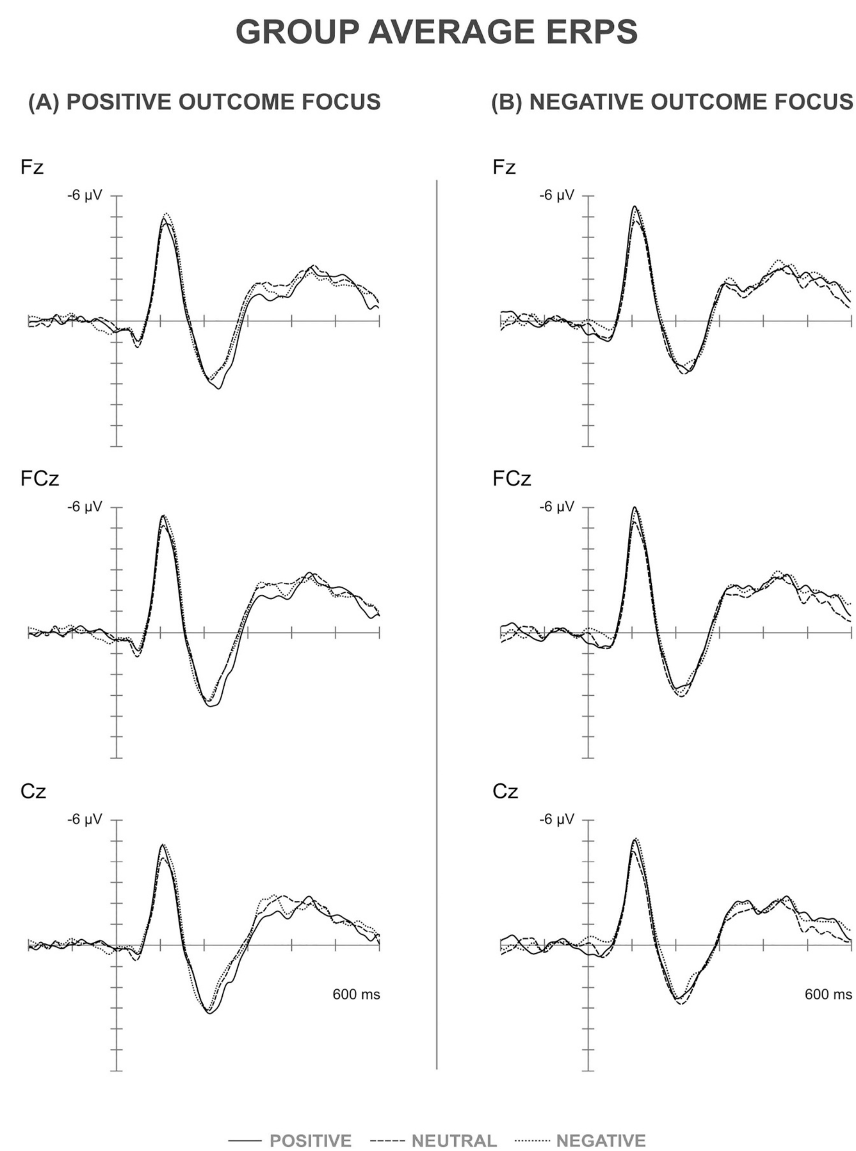

Fig. A1. ERP-results of Experiment 1. Group average ERP waveforms to the positive, neutral, and negative tones in positive (A) and negative (B) outcome focus conditions on the representative Fz, FCz and Cz electrode sites. The physical onset of the tones is at the crossing of the axes (0 ms). Negative polarity is plotted upwards.

Appendix B. Results of behavioral performance in the valence induction and selective listening tasks

In this appendix, we report results in the valence induction and selective listening tasks of Experiment 1 and 2. Although these results are not focal to our hypotheses, they are reported for the sake of completeness and to show that participants complied with instructions, and also (as expected) applied differential strategies in different valence and outcome focus conditions in order to maximize their monetary outcome.

Experiment 1

Behavioral performance was adequate in the valence induction and selective listening tasks. The behavioral results of the valence induction task are presented in Table A1. RT analyses were restricted to target trials with correct responses (6.9% and 4.6% of all target trials were excluded because of incorrect responses in the valence induction and selective listening task, respectively). In the valence induction task, RTs were calculated from the beginning of the loudness increase in the ongoing tone. RTs below 100 ms were excluded as an a priori criterion (8.9% and 0% of all trials in the valence induction and selective listening tasks, respectively).

Table A1
Behavioral results of the valence induction phase in Experiment 1: Mean reaction times (RTs), accuracy rates (ACCs), false alarm rates (FARs), and hit rates (HRs) in each valence condition; *SD* in parentheses.

Valence	RT	ACC	HR	FAR
Positive	333 (50)	88.5 (12.9)	93.7 (7.8)	16.7 (25.1)
Neutral	335 (46)	91.3 (9.8)	92.9 (8.4)	10.3 (19.8)
Negative	333 (41)	92.0 (6.9)	92.7 (7.6)	8.6 (10.4)

We conducted a 3 × 3 MANOVA for repeated measures with valence as within-participants factor and counterbalancing group as between-participants factor on the behavioral measures of the valence induction task. RTs did not show valence differences, $F < 1$, *n.s.* The average accuracy (ACC) across participants was adequately high ($M = 90.6\%$, $SD = 9.0\%$). ACC did not show a significant valence main effect, $F(2,20) = 2.30$, $p = .126$, $\eta_p^2 = .187$. However, ACC was significantly lower for positive compared with negative tones, $F(1,21) = 4.63$, $p = .044$, $\eta_p^2 = .181$, while valenced vs. neutral conditions did not differ significantly, $F < 1$, *n.s.* Similarly to the behavioral results of our previous study, the relatively low mean ACC in the positive condition emerged also due to high false alarms (see Table A1), as false alarms had no negative consequences in this condition. Overall, false alarms did not show a significant valence main effect, $F(2,20) = 2.30$, $p = .126$, $\eta_p^2 = .187$; however, as expected, participants made significantly more false alarms in the positive compared with negative condition, $F(1,21) = 4.77$, $p = .040$, $\eta_p^2 = .185$, while valenced versus neutral conditions did not differ significantly, $F < 1.24$, *n.s.* Analogue analysis on the hit rates (HR) did not show any differences; all $Fs < 1.11$, *n.s.* Altogether, the behavioral results reflect that participants were engaged in the valence induction task and that they made use of strategic differentiations between valence conditions.

Behavioral results of the selective listening task are presented in Table A2. RTs did not differ between positive and negative outcome focus conditions, $t(23) = -1.16$, $p = .257$, $d_z = -0.24$. The average ACC across participants was adequately high ($M = 96.1\%$, $SD = 3.1\%$). ACC and HR did not differ between positive and negative outcome focus conditions, $t(23) = -1.03$, $p = .313$, $d_z = -0.21$, for ACCs; and $t(23) = -0.57$, $p = .572$, $d_z = -0.12$ for HRs. However, participants made more false alarms on positive outcome blocks compared to negative outcome blocks, $t(23) = 2.54$, $p = .018$, $d_z = 0.52$; thus, suggesting that participants applied different behavioral strategies in the different outcome focus conditions.

Table A2
Behavioral results of the selective listening task of Experiment 1: Mean reaction times (RTs), accuracy rates (ACCs), false alarm rates (FARs), and hit rates (HRs) in positive and negative out-come focus conditions; *SD* in parentheses.

Outcome focus	RT	ACC	HR	FAR
Positive	500 (41)	95.7 (3.9)	95.0 (6.8)	3.5 (1.5)
Negative	504 (46)	96.5 (3.4)	95.8 (6.7)	2.8 (1.0)

Experiment 2

Behavioral performance was adequate in the valence induction and selective listening tasks. The behavioral results of the valence induction task are presented in Table A3. RT analyses were restricted to target trials with correct responses (2.8% and 21.8% of all target trials were excluded because of incorrect responses in the valence induction and selective listening task, respectively). In the valence induction task, RTs were calculated from the beginning of the loudness increase in the ongoing tone. RTs below 100 ms were excluded as an a priori criterion (6.8% and 0% of all trials in the valence induction and selective listening tasks, respectively).

We conducted a 3 × 3 MANOVA for repeated measures with valence as within-participants factor and counterbalancing group as between-participants factor (see Results section of Experiment 1) on the behavioral measures of the valence induction task. RTs showed a non-significant valence main effect, $F(2,20) = 2.97$, $p = .074$, $\eta_p^2 = .229$; $F(1,21) = 2.44$, $p = .128$, $\eta_p^2 = .104$, and, $F(1,21) = 1.27$, $p = .274$, $\eta_p^2 = .057$, for valenced vs. neutral and positive vs. negative, respectively. The average accuracy (ACC) across participants was adequately high ($M = 92.6\%$, $SD = 7.3\%$). ACCs showed a non-significant valence main effect, $F(2,20) = 3.02$, $p = .071$, $\eta_p^2 = .232$ as in Experiment 1, ACC was significantly lower for positive compared with negative tones, $F(1,21) = 6.12$, $p = .022$, $\eta_p^2 = .226$. The valenced vs. neutral comparison failed the criterion of significance, $F(1,21) = 4.02$, $p = .058$, $\eta_p^2 = .161$. The relatively low mean ACC in the positive condition emerged again due to high false alarms (see Table A3). False alarms showed a marginally significant valence main effect, $F(2,20) = 3.13$, $p = .066$, $\eta_p^2 = .238$; and, as expected, participants made significantly more false alarms in the positive compared with negative condition, $F(1,21) = 5.67$, $p = .028$, $\eta_p^2 = .213$. The valenced vs. neutral comparison was also significant, $F(1,21) = 5.73$, $p = .026$, $\eta_p^2 = .214$. Analogue analysis on the hit rates (HR) did not show any differences;

all *Fs* < 1.05, *n.s.* Additionally, participants achieved adequately high accuracy on the tone-discrimination trials of the valence induction task ($M = 93.8\%$, $SD = 5.4\%$).

Table A3
Behavioral results of the valence induction phase in Experiment 2: Mean reaction times (RTs), accuracy rates (ACCs), false alarm rates (FARs), and hit rates (HRs) in each valence condition; *SD* in parentheses.

Valence	RT	ACC	HR	FAR
Positive	343 (48)	89.3 (12.9)	97.6 (6.6)	19.0 (26.3)
Neutral	348 (33)	94.0 (6.5)	96.5 (6.0)	8.6 (10.3)
Negative	337 (34)	94.5 (5.4)	97.4 (6.0)	8.5 (8.2)

On the selective listening task, the average RT across participants was 367 ms ($SD = 24$ ms) with an average ACC of 78.4% ($SD = 12.9\%$), suggesting that the continuous selective listening task was–in line with our effort to increase task-demands–indeed more difficult compared to the selective listening task without continuous noise mask (in our initial study, Folyi et al., 2016, with comparable selective listening task without noise mask the average RT was 590 ms [$SD = 74$ ms], $t(27.76) = 13.96$, $p < .001$, $d = 4.03$, for the comparison of our initial study and the present Experiment 2 on RTs; while the average ACC was 93.2% [$SD = 5.1\%$] in our former study, $t(29.92) = 5.23$, $p < .001$, $d = 1.51$ for the between-experiments comparison on ACC).

Appendix C. Manipulation check in Experiment 2

In Experiment 2, we repeated the manipulation check of Folyi et al. (2016). Because of the long duration of Experiment 1 (i.e., the duration of the test phase was increased after introducing the outcome focus manipulation), we did not administer a manipulation check in Experiment 1.

Method

On the experimental trials of the AST, the positive, negative, and neutral tones were presented in two versions: with an illusory movement from a central position toward the right (“moving to the right” tones) and left (“moving to the left” tones) side of the perceiver, respectively. Participants' task was to categorize the direction of this 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. An experimental trial of the AST started with a 1000 ms long presentation of a fixation cross without auditory stimuli and it remained on the screen until the end of the trial. After 1000 ms, a positive, negative or neutral tone was played. From 500 ms after stimulus onset, the 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 creating a movement illusion toward the right or the left side of the perceiver. Participants had to categorize the direction of this illusory movement by uttering “good” (right direction) or “bad” (left direction). While a voice key apparatus recorded the onset of the vocal response (RT), the experimenter registered the response category online. Tones were terminated as soon as a response was detected by the voice key (maximal tone duration was 6000 ms). Finally, error feedback was presented visually in the case of an incorrect response. The AST comprised 12 practice and 30 experimental trials. Positive, negative, and neutral tones were presented ten 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 were identical to the experimental trials except that natural sounds with positive, negative, and neutral meaning (e.g., laughing, attack, office noise) were presented instead of the experimental tones. 18 natural sounds (six of each valence category) were selected from the IADS battery (Bradley and Lang, 2007). Averaged normative valence ratings on a 9-point scale ranging from most unpleasant (1) to most pleasant (9) were: $M = 7.16$ ($SD = 0.37$) for positive, $M = 2.19$ ($SD = 0.54$) for negative, and $M = 4.75$ ($SD = 0.37$) for neutral sounds, respectively. From this pool of filler sounds, one positive, one negative, and one neutral sound were selected for each participant in a way that each sound is presented four times across the sample of twenty-four participants. We created “moving to the right” and “moving to the left” versions of the natural sounds in the same way as for the experimental tones. Filler sounds were also presented ten times each in random order (five times in “moving to the left” and five times in “moving to the right” version).

At the end of the experiment, participants were asked to rate the valence of the auditory stimuli presented in the AST task on a 9-point scale ranging from most unpleasant (1) to most pleasant (9). During the valence rating, auditory stimuli were presented with constant intensity.

Results

For the AST, RTs were calculated from the beginning of the illusory movement in the ongoing tone. RT analyses were restricted to correct trials (5.6% of the trials was excluded be-cause 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 (4.3% of the trials), and RT outliers were excluded (1.1% of the trials; for each participant an upper and a lower outlier criterion were calculated based on the individual distribution of RTs Tukey, 1977). One participant was excluded during data analysis because of an insufficient number of trials remaining after this procedure (73% of the trials were excluded due to generally high RTs, individual mean RT was 2836 ms).[10] A further participant had incomplete data in the affective Simon task (AST) due to technical failure of the voice key apparatus. Note that datasets from these two participants were included in the analysis of the main phases of the experiment (i.e., valence induction and test phase) as we had no reason to assume that the data emerging from the main phases of the experiment would be invalid for these participants (i.e., low task performance of the participant was specific to the AST; and the voice key device was used only during the AST). Consequently, we did not want to reduce test power for detecting valence differences on the N1 (i.e., our central research question) due to exclusion of valid datasets from the main analysis.

Mean RTs and ERs for congruent, incongruent, and neutral trials, as well as the mean AS effect are presented in Table B1 for the experimental and

[10] Excluding this dataset from the analysis of the valence induction and test phases did not change the pattern of results reported above.

filler trials, respectively. On the experimental trials, the incongruent-minus-congruent RT difference showed the expected AS effect, $t(21) = 1.77$, $p = .046$ (one-tailed), $d_z = 0.38$. (A cross-experiments analysis with the corresponding data of our previous study, Folyi et al., 2016, showed a significant AS effect, $F(1,42) = 7.92$, $p = .007$, $\eta_p^2 = .159$; which was not moderated by Experiment, $F(1,42) = 0.16$, $p = .696$, $\eta_p^2 = .004$.) Similar analysis on the filler trials (natural sounds) did not yield significant results, $|t| < 1$, *n.s.* However, this corresponds exactly to the findings by Folyi et al. (2016). Analogue analyses on the ERs did not show significant AS effects for the experimental trials, $t(21) = 1.16$, $p = .257$, $d_z = 0.25$, or filler trials, $t(21) = 1.30$, $p = .208$, $d_z = 0.28$.

Table B1
Mean RTs (in ms) and mean ERs (%) as a function of stimulus and response valence congruency for experimental and filler trials in the AST of Experiment 2; *SD* in parentheses.

	Experimental tones		Fillers	
	RT	ER	RT	ER
Neutral	917 (316)	6.4 (10.5)	814 (233)	6.4 (11.8)
Congruent	880 (273)	3.6 (6.6)	802 (240)	4.6 (11.4)
Incongruent	912 (281)	5.5 (8.6)	799 (251)	7.3 (10.3)
AS effect[a]	32 [18]	1.8 [1.6]	−3 [19]	2.7 [2.1]

[a] Incongruent-minus-congruent difference; standard errors in brackets.

The means of the explicit valence ratings reflected the learned valences ($M_{pos} = 4.10$, $M_{neut} = 4.08$, $M_{neg} = 4.05$). However, we have to admit that differences are smaller compared to the previous study. A 3 × 3 MANOVA for repeated measures with valence (positive, neutral, negative) as within-participants factor and counterbalancing group as between-participants factor did not show a valence main effect, $F < 1$. As in Folyi et al. (2016), it was observable that explicit ratings were dominated by tone frequency.

Discussion

For the AST, we were able to replicate the result by Folyi et al. (2016): if affective valence of the tones were congruent to the affective valence of the response (i.e., responding by uttering "good"/"bad" to a positive/negative tone), RTs were faster than in the case of incongruence (i.e., responding by uttering "bad"/"good" to a positive/negative tone). This finding can be considered as evidence that positive and negative tones indeed diverge with regard to valence, but show equivalent N1-enhancement relative to neutral tones at the early level of sound encoding.

As in our previous study, natural emotional sounds of the filler trials did not reveal an AS effect. This might appear surprising, given that we demonstrated the expected AS-effect using natural emotional sounds in an independent study (Folyi and Wentura, 2017). However, the AST of the present study was not designed to replicate the results of Folyi and Wentura (2017), that is reflected in our power planning and in important differences between the two tasks. First, our power planning focused on the central questions of the present study (i.e., valence effect on the N1), while to replicate an AS-effect of $d_z = 0.33$ found by Folyi and Wentura (2017), with a probability of $1 - \beta = .80$ and an α-value of .05 (one-tailed), a sample size of $N = 59$ would have been needed (calculated with the aid of G*Power 3 software; Faul et al., 2007). Second, there were important differences between the ASTs in these two studies: In the study Folyi and Wentura (2017), a variety of 60 natural sounds was presented only once per participant. In the present version, we randomly selected one positive, one negative, and one neutral natural sound per participant, which were repeated several times (in correspondence to the fact that there were only three tones in the experimental condition). Thus, it might be that repetition leads to habituation for natural sounds, but not (to the same extent) for the tones. Furthermore, we have to keep in mind that the manipulation check phase was administered at the end of a long and demanding experimental session, while in our independent study (Folyi and Wentura, 2017), participants had to work through only the AST. This fact can have important consequences: Certainly, at the end of a long experimental session participants' motivation and ability to focus on the task at hand might have been more limited. Furthermore, in the current study, participants were presented with simple sinusoid tones during the main phases of the experiment and on half of the AS-trials, thus the reach, natural emotional sounds might have been more surprising, that could lead to different strategies on these AS-trials (e.g., more effort to suppress the sound meaning). Because of the long experimental session, we presented relatively few AST trials (10 trials in each valence condition in the experimental and filler trials, respectively; thus, the AS-effect for natural emotional sounds was based only on 10 incongruent and 10 congruent trials per participant, while on 20 incongruent and 20 congruent trials per participant in the study of Folyi and Wentura, 2017) that could again contribute to the lack of finding significant AS-effect for the natural emotional sounds in the present Experiment 2. In the present study, we did not find support for a valence differentiation in explicit self-reports. However, as in Folyi et al. (2016) the explicit ratings were oriented on the objective tone frequencies.

References

Ahveninen, J., Hämäläinen, M., Jääskeläinen, I.P., Ahlfors, S.P., Huang, S., Lin, F.-H., ... Belliveau, J.W., 2011. Attention-driven auditory cortex short-term plasticity helps segregate relevant sounds from noise. Proc. Natl. Acad. Sci. 108 (10), 4182–4187. https://doi.org/10.1073/pnas.1016134108.

Anderson, B.A., 2016a. The attention habit: how reward learning shapes attentional selection. Ann. N. Y. Acad. Sci. 1369 (1), 24–39. https://doi.org/10.1111/nyas.12957.

Anderson, B.A., 2016b. Value-driven attentional capture in the auditory domain. Atten. Percept. Psychophys. 78 (1), 242–250. https://doi.org/10.3758/s13414-015-1001-7.

Anderson, B.A., Laurent, P.A., Yantis, S., 2011. Value-driven attentional capture. Proc. Natl. Acad. Sci. 108 (25), 10367–10371. https://doi.org/10.1073/pnas.1104047108.

Asutay, E., Västfjäll, D., 2016. Auditory attentional selection is biased by reward cues. Sci. Rep. 6, 36989. https://doi.org/10.1038/srep36989.

Bendixen, A., SanMiguel, I., Schröger, E., 2012. Early electrophysiological indicators for predictive processing in audition: a review. Int. J. Psychophysiol. 83 (2), 120–131. https://doi.org/10.1016/j.ijpsycho.2011.08.003.

Bourgeois, A., Neveu, R., Bayle, D.J., Vuilleumier, P., 2017. How does reward compete with goal-directed and stimulus-driven shifts of attention? Cognit. Emot. 31 (1), 109–118. https://doi.org/10.1080/02699931.2015.1085366.

Bradley, M.M., Lang, P.J., 2007. The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective ratings of sounds and instruction manual. University of Florida, Gainesville, FL (Tech. Rep. B-3).

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. J. Neurosci. 31 (21), 7801–7810. https://doi.org/10.1523/JNEUROSCI.6236-10.2011.

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. Eur. J. Neurosci. 37 (2), 303–315. https://doi.org/10.1111/ejn.12043.
Brosch, T., Sander, D., Pourtois, G., Scherer, K.R., 2008. Beyond fear rapid spatial orienting toward positive emotional stimuli. Psychol. Sci. 19 (4), 362–370. https://doi.org/10.1111/j.1467-9280.2008.02094.x.
Bucker, B., Theeuwes, J., 2016. Appetitive and aversive outcome associations modulate exogenous cueing. Atten. Percept. Psychophys. 78 (7), 2253–2265. https://doi.org/10.3758/s13414-016-1107-6.
Burkard, R.F., Eggermont, J.J., Don, M., 2007. Auditory Evoked Potentials: Basic Principles and Clinical Application. Lippincott Williams & Wilkins.
Chelazzi, L., Eštočinová, J., Calletti, R., Gerfo, E.L., Sani, I., Libera, C.D., Santandrea, E., 2014. Altering spatial priority maps via reward-based learning. J. Neurosci. 34 (25), 8594–8604. https://doi.org/10.1523/JNEUROSCI.0277-14.2014.
Cumming, G., 2011. Exploratory Software for Confidence Intervals (Version Meta-Analysis). Retrieved from. www.thenewstatistics.com.
Eimer, M., Holmes, A., 2007. Event-related brain potential correlates of emotional face processing. Neuropsychologia 45 (1), 15–31. https://doi.org/10.1016/j.neuropsychologia.2006.04.022.
Eimer, M., Holmes, A., McGlone, F.P., 2003. The role of spatial attention in the processing of facial expression: an ERP study of rapid brain responses to six basic emotions. Cogn. Affect. Behav. Neurosci. 3 (2), 97–110. https://doi.org/10.3758/CABN.3.2.97.
Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A., 2007. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods 39 (2) Retrieved from. http://link.springer.com/article/10.3758/BF03193146.
Folyi, T., Wentura, D., 2017. Fast and unintentional evaluation of emotional sounds: evidence from brief segment ratings and the affective Simon task. Cognit. Emot. 31 (2), 312–324. https://doi.org/10.1080/02699931.2015.1110514.
Folyi, T., Liesefeld, H.R., Wentura, D., 2016. Attentional enhancement for positive and negative tones at an early stage of auditory processing. Biol. Psychol. 114, 23–32. https://doi.org/10.1016/j.biopsycho.2015.12.001.
Fritz, J.B., Elhilali, M., David, S.V., Shamma, S.A., 2007. Auditory attention — focusing the searchlight on sound. Curr. Opin. Neurobiol. 17 (4), 437–455. https://doi.org/10.1016/j.conb.2007.07.011.
Gilmore, C.S., Clementz, B.A., Berg, P., 2009. Hemispheric differences in auditory oddball responses during monaural versus binaural stimulation. Int. J. Psychophysiol. 73 (3), 326–333. https://doi.org/10.1016/j.ijpsycho.2009.05.005.
Handy, T.C., 2005. Event-related Potentials: A Methods Handbook. MIT Press, Cambridge, MA.
Herrmann, C.S., Knight, R.T., 2001. Mechanisms of human attention: event-related potentials and oscillations. Neurosci. Biobehav. Rev. 25 (6), 465–476. https://doi.org/10.1016/S0149-7634(01)00027-6.
Horváth, J., Czigler, I., Sussman, E., Winkler, I., 2001. Simultaneously active pre-attentive representations of local and global rules for sound sequences in the human brain. Cogn. Brain Res. 12 (1), 131–144. https://doi.org/10.1016/S0926-6410(01)00038-6.
Horváth, J., Maess, B., Baess, P., Tóth, A., 2012. Action–sound coincidences suppress evoked responses of the human auditory cortex in EEG and MEG. J. Cogn. Neurosci. 24 (9), 1919–1931. https://doi.org/10.1162/jocn_a_00215.
Houwer, J.D., Crombez, G., Baeyens, F., Hermans, D., 2001. On the generality of the affective Simon effect. Cognit. Emot. 15 (2), 189–206. https://doi.org/10.1080/02699930125883.
Jacobsen, T., Schröger, E., Horenkamp, T., Winkler, I., 2003. Mismatch negativity to pitch change: varied stimulus proportions in controlling effects of neural refractoriness on human auditory event-related brain potentials. Neurosci. Lett. 344 (2), 79–82. https://doi.org/10.1016/S0304-3940(03)00408-7.
JASP Team, 2018. JASP (Version 0.9.0.1) [Computer software].
Jeffreys, H., 1961. Theory of Probability. Oxford University Press, Oxford, England.
King, A.J., Nelken, I., 2009. Unraveling the principles of auditory cortical processing: can we learn from the visual system? Nat. Neurosci. 12, 698–701. https://doi.org/10.1038/nn.2308.
Lachter, J., Forster, K.I., Ruthruff, E., 2004. Forty-five years after Broadbent (1958): still no identification without attention. Psychol. Rev. 111 (4), 880–913. https://doi.org/10.1037/0033-295X.111.4.880.
Le Pelley, M.E., Pearson, D., Griffiths, O., Beesley, T., 2015. When goals conflict with values: counterproductive attentional and oculomotor capture by reward-related stimuli. J. Exp. Psychol. Gen. 144 (1), 158–171. https://doi.org/10.1037/xge0000037.
Le Pelley, M.E., Mitchell, C.J., Beesley, T., George, D.N., Wills, A.J., 2016. Attention and associative learning in humans: an integrative review. Psychol. Bull. 142 (10), 1111–1140. https://doi.org/10.1037/bul0000064.
Luck, S.J., Gaspelin, N., 2017. How to get statistically significant effects in any ERP experiment (and why you shouldn't). Psychophysiology 54 (1), 146–157. https://doi.org/10.1111/psyp.12639.
Maxwell, S.E., Delaney, H.D., 2004. Designing Experiments and Analyzing Data: A Model Comparison Perspective, 2nd ed. Psychology Press, New York.
Méndez-Bértolo, C., Moratti, S., Toledano, R., Lopez-Sosa, F., Martínez-Alvarez, R., Mah, Y.H., ... Strange, B.A., 2016. A fast pathway for fear in human amygdala. Nat. Neurosci. 19 (8), 1041–1049. https://doi.org/10.1038/nn.4324.
Müller, S., Rothermund, K., Wentura, D., 2016. Relevance drives attention: attentional bias for gain- and loss-related stimuli is driven by delayed disengagement. Q. J. Exp. Psychol. 69 (4), 752–763. https://doi.org/10.1080/17470218.2015.1049624.
Murphy, S., Fraenkel, N., Dalton, P., 2013. Perceptual load does not modulate auditory distractor processing. Cognition 129 (2), 345–355. https://doi.org/10.1016/j.cognition.2013.07.014.
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 (4), 375–425. https://doi.org/10.1111/j.1469-8986.1987.tb00311.x.
Näätänen, R., Astikainen, P., Ruusuvirta, T., Huotilainen, M., 2010. Automatic auditory intelligence: an expression of the sensory–cognitive core of cognitive processes. Brain Res. Rev. 64 (1), 123–136. https://doi.org/10.1016/j.brainresrev.2010.03.001.
O'Brien, R.G., Kaiser, M.K., 1985. MANOVA method for analyzing repeated measures designs: an extensive primer. Psychol. Bull. 97 (2), 316–333. https://doi.org/10.1037/0033-2909.97.2.316.
Öhman, A., 2005. The role of the amygdala in human fear: automatic detection of threat. Psychoneuroendocrinology 30 (10), 953–958. https://doi.org/10.1016/j.psyneuen.2005.03.019.
Öhman, A., Soares, S.C., Juth, P., Lindström, B., Esteves, F., 2012. Evolutionary derived modulations of attention to two common fear stimuli: serpents and hostile humans. J. Cogn. Psychol. 24 (1), 17–32. https://doi.org/10.1080/20445911.2011.629603.
Pinheiro, A.P., Barros, C., Dias, M., Kotz, S.A., 2017. Laughter catches attention. Biol. Psychol. 130, 11–21. https://doi.org/10.1016/j.biopsycho.2017.09.012.
Pollatsek, A., Well, A.D., 1995. On the use of counterbalanced designs in cognitive research: a suggestion for a better and more powerful analysis. J. Exp. Psychol. Learn. Mem. Cogn. 21 (3), 785–794. https://doi.org/10.1037/0278-7393.21.3.785.
Pourtois, G., Schettino, A., Vuilleumier, P., 2013. Brain mechanisms for emotional influences on perception and attention: what is magic and what is not. Biol. Psychol. 92 (3), 492–512. https://doi.org/10.1016/j.biopsycho.2012.02.007.
Rohr, M., Degner, J., Wentura, D., 2012. Masked emotional priming beyond global valence activations. Cognit. Emot. 26 (2), 224–244. https://doi.org/10.1080/02699931.2011.576852.
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 (1), 34–46. https://doi.org/10.1037/1528-3542.8.1.34.
Rothermund, K., Gast, A., Wentura, D., 2011. Incongruency effects in affective processing: automatic motivational counter-regulation or mismatch-induced salience? Cognit. Emot. 25 (3), 413–425. https://doi.org/10.1080/02699931.2010.537075.
Rouder, J.N., Speckman, P.L., Sun, D., Morey, R.D., Iverson, G., 2009. Bayesian t tests for accepting and rejecting the null hypothesis. Psychon. Bull. Rev. 16 (2), 225–237. https://doi.org/10.3758/PBR.16.2.225.
Rouder, J.N., Morey, R.D., Verhagen, J., Swagman, A.R., Wagenmakers, E.-J., 2017. Bayesian analysis of factorial designs. Psychol. Methods 22 (2), 304–321. https://doi.org/10.1037/met0000057.
Schwager, S., Rothermund, K., 2013. Counter-regulation triggered by emotions: positive/negative affective states elicit opposite valence biases in affective processing. Cognit. Emot. 27 (5), 839–855. https://doi.org/10.1080/02699931.2012.750599.
Schwager, S., Rothermund, K., 2014. On the dynamics of implicit emotion regulation: counter-regulation after remembering events of high but not of low emotional intensity. Cognit. Emot. 28 (6), 971–992. https://doi.org/10.1080/02699931.2013.866074.
Spence, C.J., Driver, J., 1994. Covert spatial orienting in audition: exogenous and endogenous mechanisms. J. Exp. Psychol. Hum. Percept. Perform. 20 (3), 555–574. https://doi.org/10.1037/0096-1523.20.3.555.
Theeuwes, J., 1992. Perceptual selectivity for color and form. Percept. Psychophys. 51 (6), 599–606. https://doi.org/10.3758/BF03211656.
Tukey, J.W., 1977. Exploratory Data Analysis. Addison-Wesley, Reading, MA.
Vuilleumier, P., 2015. Affective and motivational control of vision. Curr. Opin. Neurol. 28 (1), 29–35. https://doi.org/10.1097/WCO.0000000000000159.
Vuilleumier, P., Huang, Y.-M., 2009. Emotional attention uncovering the mechanisms of affective biases in perception. Curr. Dir. Psychol. Sci. 18 (3), 148–152. https://doi.org/10.1111/j.1467-8721.2009.01626.x.
Wagenmakers, E., Wetzels, R., Borsboom, D., van der Maas, H.L.J., 2011. Why psychologists must change the way they analyze their data: the case of psi: comment on Bem (2011). J. Pers. Soc. Psychol. 100 (3), 426–432. https://doi.org/10.1037/a0022790.
Wentura, D., Voss, A., Rothermund, K., 2009. Playing TETRIS for science counter-regulatory affective processing in a motivationally "hot" context. Acta Psychol. 131 (3), 171–177. https://doi.org/10.1016/j.actpsy.2009.05.008.
Wentura, D., Müller, P., Rothermund, K., 2014. Attentional capture by evaluative stimuli: gain- and loss-connoting colors boost the additional-singleton effect. Psychon. Bull. Rev. 21 (3), 701–707. https://doi.org/10.3758/s13423-013-0531-z.
Wentura, D., Müller, P., Rothermund, K., Voss, A., 2018. Counter-regulation in affective attentional biases: evidence in the additional singleton paradigm. Q. J. Exp. Psychol. 71 (ja), 1209–1218. https://doi.org/10.1080/17470218.2017.1315147.
Winkler, I., Teder-Sälejärvi, W.A., Horváth, J., Näätänen, R., Sussman, E., 2003. Human auditory cortex tracks task-irrelevant sound sources. Neuroreport 14 (16), 2053.
Woldorff, M.G., Hillyard, S.A., 1991. Modulation of early auditory processing during selective listening to rapidly presented tones. Electroencephalogr. Clin. Neurophysiol. 79 (3), 170–191. https://doi.org/10.1016/0013-4694(91)90136-R.