An evil face? Verbal evaluative multi-CS conditioning enhances face-evoked mid-latency magnetoencephalographic responses

Markus Junghöfer,¹,² Maimu Alissa Rehbein,¹,² Julius Maitzen,¹ Sebastian Schindler,³,⁴ and Johanna Kissler³,⁴

¹Institute for Biomagnetism and Biosignalanalysis, University Hospital Münster, Münster D-48149, Germany, ²Otto Creutzfeldt Center for Cognitive and Behavioral Neuroscience, University of Münster, Münster D-48151, Germany, ³Department of Psychology, Affective Neuropsychology Unit and ⁴Center of Excellence Cognitive Interaction Technology (CITEC), University of Bielefeld, Bielefeld D-33501, Germany

Correspondence should be addressed to Johanna Kissler, Department of Psychology, Affective Neuropsychology Unit, University of Bielefeld, Bielefeld D-33501, Germany. E-mail: johanna.kissler@uni-bielefeld.de.

Abstract
Humans have a remarkable capacity for rapid affective learning. For instance, using first-order US such as odors or electric shocks, magnetoencephalography (MEG) studies of multi-CS conditioning demonstrate enhanced early (<150 ms) and mid-latency (150–300 ms) visual evoked responses to affectively conditioned faces, together with changes in stimulus evaluation. However, particularly in social contexts, human affective learning is often mediated by language, a class of complex higher-order US. To elucidate mechanisms of this type of learning, we investigate how face processing changes following verbal evaluative multi-CS conditioning. Sixty neutral expression male faces were paired with phrases about aversive crimes (30) or neutral occupations (30). Post conditioning, aversively associated faces evoked stronger magnetic fields in a mid-latency interval between 220 and 320 ms, localized primarily in left visual cortex. Aversively paired faces were also rated as more arousing and more unpleasant, evaluative changes occurring both with and without contingency awareness. However, no early MEG effects were found, implying that verbal evaluative conditioning may require conceptual processing and does not engage rapid, possibly sub-cortical, pathways. Results demonstrate the efficacy of verbal evaluative multi-CS conditioning and indicate both common and distinct neural mechanisms of first- and higher-order multi-CS conditioning, thereby informing theories of associative learning.

Key words: neural plasticity; associative learning; emotion; language; magnetoencephalography; evaluative conditioning

Introduction
Facial expression is a powerful non-verbal communication channel that humans automatically orient to (Brosch et al. 2008) and (Pourtois et al. 2004). EEG experiments reveal preferential visual processing, particularly of negative emotional expressions, at early (<150 ms: Pourtois et al., 2005), mid-latency (150–300 ms: Schupp et al., 2004; Blau et al., 2007), and late processing stages (> 400 ms: Schupp et al., 2004).

Whereas facial expressions convey affective significance via changes in facial feature configuration that have a biological basis and are almost universally understood (Ekman and Friesen, 1971; Jack et al., 2012), faces with neutral expressions can likewise acquire emotional significance via associative
learning. For instance, a recent magnetoencephalography (MEG) study showed that conditioning many different facial identities with unpleasant odors resulted in modulation of early (50–80 ms) and mid-latency (130–190 ms) occipito-temporal and prefrontal cortical responses (Steinberg et al., 2012). Although affective face conditioning had been shown before (e.g. Dimberg, 1987; Morris et al., 1998; Pizzagalli et al., 2003), this study was the first to demonstrate, in the absence of contingency awareness, robust and rapid cortical learning effects for a multitude of neutral-expression CS faces and corresponding US-congruent changes in stimulus evaluation. The paradigm has since been termed “multi-CS conditioning” and its main novel results, namely rapid (starting before 100 ms) conditioning effects in prefrontal and sensory areas for many different CSs, have been replicated with various visual and auditory CS-stimuli (Bröckelmann et al., 2011, 2013; Steinberg et al., 2013; Hintze et al., 2014; Rehbein et al., 2014, 2015; Junghöfer et al., 2015). Multi-CS conditioning effects were also observed with different US, such as odors (Steinberg et al., 2012), electric shock (Rehbein et al., 2014) and simple (Rehbein et al., 2015), or complex auditory emotional stimuli (Bröckelmann et al., 2011).

Faces contain many different social signals (e.g. Mignault and Chaudhuri, 2003) and, as a CS, may predispose learning of particular associations: for instance, Todrank and Chaudhuri, 2003) and, as a CS, may predispose learning of

For evaluating the neural mechanisms of verbal evaluative conditioning, several studies have shown that verbal evaluative conditioning could provide clues as to the pathways involved which are not readily apparent in purely behavioral studies. However, various techniques such as EEG event-related potentials (ERPs), fMRI, and MEG have been used to study verbal evaluative conditioning. ERP studies have shown early (50–100 ms) and later (150–300 ms) effects in the P1/P2 and N1/P3 components, respectively. These effects are often associated with valence and arousal processing of emotional stimuli.

Facial expressions have been found to elicit robust and rapid cortical learning effects (e.g., Bröckelmann et al., 2011). Conditioning faces with two-word language descriptors can even counteract racial prejudice, changing the affective appraisal of outgroup faces (Olson and Fazio, 2006).

However, extant behavioral studies of evaluative conditioning have used relatively small stimulus sets and many repetitions of identical CS–US pairings, whereas multi-CS conditioning with first-order US has been shown to be remarkably efficient, raising questions regarding the possible capacity limits of traditional evaluative conditioning.

Furthermore, investigating the neural mechanisms of verbal evaluative conditioning could provide clues as to the pathways involved which are not readily apparent in purely behavioral studies.

So far, little is known about the cortical dynamics of evaluative conditioning. To our knowledge only two previous studies have explicitly investigated the issue, both using EEG event-related potentials (ERPs): Fritsch and Kuchinke (2013) and Kuchinke et al. (2015) studied evaluative conditioning of pseudoword CS with affective picture US as a model for the acquisition of emotional significance in written language. After multiple associative learning sessions, they found early P1/N1 and late P300 responses to discriminate negatively from neutrally (Fritsch and Kuchinke, 2013) or positively paired stimuli (Kuchinke et al., 2015). Early ERP effects were localized in medial-frontal regions, in line with these regions’ involvement in early valence discrimination (Kuchinke et al., 2015). These results are broadly consistent with the aforementioned MEG data on multi-CS conditioning (e.g. Steinberg et al., 2012, 2013), extending them to abstract pseudoword CS. In fact, because multi-CS conditioning often also results in changes in affective evaluations, multi-CS conditioning can be thought of as a special instance of evaluative conditioning. In sum, extant data indicate that multi-CS conditioning induces rapid plasticity (<100 ms) in perceptual and evaluative brain systems, followed by a second conceptual processing wave at mid-latency stages from around 150 ms, in line with dual-route models of affective processing (LeDoux, 2000; Bullier, 2001).

Thus, whether multi-CS conditioning with language stimuli as higher-order US has a similar capacity as multi-CS conditioning with first-order US and whether verbal evaluative conditioning utilizes the same neural pathways as non-verbal conditioning is hitherto unknown. An answer to these questions would inform models of associative learning.

Evidence suggests that the processing of emotional language, in spite of several similarities with other affective stimuli (for review see Citron, 2012; Kissler, 2013), also differs in important ways: for instance, electrophysiology studies show that processing of verbal emotional stimuli robustly differs from neutral ones only at mid-latency lexical and post-lexical (Kissler and Herbert, 2013; Palazova et al., 2013), but not at early perceptual stages (Trauer et al., 2012, 2015). When they are observed, early perceptual emotion effects in language themselves seem to rely on specific acquisition mechanisms such as conditioning with first-order US (Montoya et al., 1996; Schacht et al. 2012; Fritsch and Kuchinke, 2013; Kuchinke et al., 2015). Hemodynamic studies show that typical fear conditioning with electric shocks results in bilateral and generally large amygdala activation (Olson and Phelps, 2007), whereas affective learning via verbal instructions as in the instructed fear paradigm induces relatively small and left-lateralized amygdala activations (e.g. Phelps et al. 2001). Similarly, classical fear conditioning with masked CS has been shown to induce changes in electrodermal activity, whereas instructed fear learning with masked CS does not (Olson and Phelps, 2004), suggesting a dissociation between classical conditioning and verbal learning pathways.
Against the above background, the current study tests the effectiveness of verbal evaluative conditioning of neutral expression faces in a multi-CS conditioning paradigm. Many different CS faces are used and neuromagnetic correlates of changes in face-evoked responses are investigated to delineate the cortical mechanisms involved. CS faces are paired with different acoustically presented negative or neutral descriptor sentences as US. Thereby, we aim to extend previous findings from multi-CS and evaluative conditioning and address the neural basis of language-mediated social learning, a very common and arguably uniquely human form of learning. We use MEG, which has an excellent temporal resolution and, via source reconstruction, facilitates analysis of regional cortical activation.

Verbal evaluative conditioning of faces is an ecologically valid scenario that is often used in the media, but whose capacity limits and the neural mechanisms involved are hitherto unclear. To elucidate the boundary conditions and mechanisms of this important form of learning, it is key to investigate experimentally whether such learning is successful when many different CS–US associations occur, whether it occurs implicitly, in the absence of learning intention, and whether contingency awareness mediates the effects.

Specifically, the timing of cortical effects can provide important clues regarding the pathways likely involved in verbal evaluative conditioning. Early frontal and sensory effects, occurring around 100 ms, might be indicative of rapid sub-cortical affective learning. Effects occurring in later time-windows would be more in line with recruitment of slower cortical pathways. Therefore, the results of this study can inform models of associative affective learning.

Materials and methods

Participants

Twenty-two women participated in our study. One participant was excluded due to conspicuous results on clinical questionnaires. The remaining 21 participants ($M = 24.70$ years, range 20–50) showed inconspicuous levels of negative affect ($M = 33.43$; s.d. = 6.19) as determined by values on the trait scale of the State–Trait–Anxiety Inventory\(^1\) (Laux et al., 1981) and depression ($M = 16.38$; s.d. = 7.08) as determined by a simplified version of the Beck Depression Inventory (BDI-V; Schmitt and Maes, 2000). All participants had normal hearing, normal or corrected-to-normal vision and gave written informed consent to the protocol approved by the local ethics committee. Since females typically report higher levels of fear of crime than males (e.g. Warr, 2000) and the vast majority of violent crimes are committed by males (e.g. FBI, 2014), only female participants were tested in the study and male faces were used as CS. Towards a potent US and high plausibility for real-life CS–US links (e.g. Todrank et al., 1995), we opted for violent crimes as highly aversive and arousing descriptors of CS faces.

Stimuli

Conditioned stimuli. Sixty frontal view images displaying male Caucasian faces with neutral expression and mouth shut, short hair, no beard and no glasses were used. Faces were taken from the Karolinska Directed Emotional Faces database (Lundqvist et al., 1998), the NimStim set of facial expressions (Tottenham et al., 2009), and the Radboud Faces Database (Langner et al., 2010). Using Adobe\textsuperscript® Photoshop\textsuperscript®, faces were adjusted to a height of 15 cm and a resolution of 72 pixels/inch and were converted to gray scale images.

Unconditioned stimuli. Sixty German sentences describing either the neutral occupation of a person such as ‘This objective biologist investigated a coniferous tree’ (30 US-occu sentences) or an aversive criminal activity such as ‘This cold-blooded prisoner strangled people’ (30 US-crime sentences) were used. The material was taken from a preceding EEG study investigating explicit affective learning (Strehlow and Kissler, 2012). Adjectives, verbs and nouns used in these sentences did not differ between affective conditions regarding word frequency. The material is available from the corresponding author upon request. Auditory US presentation was chosen to facilitate natural, undisturbed and continuous visual perception of the CS faces during parallel CS–US presentation in the learning phase. US sentences were spoken with neutral newsreader-like intonation by one of the female investigators and digitally recorded using a Røde\textsuperscript®-microphone. The duration of the sentences (2.6–3.4 s), the root-mean-square of overall loudness, and the overall ratio of high to low frequencies were measured and analyzed with in-house Matlab\textsuperscript® based software. No differences between US-crime and US-occu sentences regarding these physical properties were found (i.e. no differential CS–US association due to differential physical properties of the auditory US). US loudness was adjusted to 60 dB above the individual’s hearing threshold, which was determined before the MEG recording.

Experimental procedure

The experiment consisted of four subjective tests, administered before and after evaluative multi-CS conditioning, and measurements of neuronal activation conducted before learning, during conditioning, and during extinction (see Figure 1).

Prior to evaluative learning, participants completed subjective ratings of hedonic valence (unpleasant–pleasant) and emotional arousal (calm–arousing) of the US-sentences and the CS-faces, using the 9-point Likert scales of the Self-Assessment–Manikin (SAM, Bradley and Lang, 1994).

The subsequent MEG recording consisted of three phases (see Figure 1B). During the pre-learning and post-learning phases, three blocks of all 60 CS faces were presented in randomized order for 600 ms each and at the center of the screen with 12.6° visual angle. A fixation cross was presented during a randomly jittered, variable, inter-trial interval (ITI, 1100–1700 ms). Participants were instructed to passively view the faces and to keep their eyes focused on the center of the screen.

In the evaluative learning phase, subjects were presented with two differently randomized blocks of sixty audiovisual CS-face/US-sentence pairings, each block containing thirty pairings of faces with aversive crime and thirty with neutral occupation sentences respectively. Neurophysiological responses of the learning phase itself were not investigated here as signal-to-noise would have been insufficient.\(^2\) Within a subject, specific

\(^1\) We would like to note that the State-Trait-Anxiety Inventory is not a pure measure of anxiety, but rather assesses general negative affect, which encompasses both anxiety and depression like symptoms (see, e.g. Nitschke et al., 2001, for more information).

\(^2\) With only two CS-US pairings and US occurrence 1 s after CS onset, just the second presentation of all CS faces during the learning phase discriminated CS-crimp from CS-occu faces, and these have only once been paired before. Thirty trials per experimental condition has been suggested as insufficient to evaluate potential effects of single trial learning.
face-sentence assignments were identical across the two blocks but were randomized across participants so that each subject perceived a different CS-US stimulus set. Faces were presented for a total of 5000 ms. The auditory US started 1000 ms after face-onset, resembling a delay conditioning procedure, and ended 1.4–0.6 s before CS-face offset, depending on US duration. ITIs were identical to the pre-learning phase. Participants were instructed to attend to the audiovisual presentation.

To identify evaluative learning effects of perceived CS valence and arousal, a post-evaluative learning SAM-rating of all CS faces was administered directly after the MEG recording.

Finally, contingency awareness of CS-US pairings was assessed by a surprise-recall task in which participants were presented with each of the 60 faces and had to decide whether it had been associated with a criminal or an occupational context. Participants were then asked to indicate how confident they were about this decision on a visual analogue scale ranging from 1 (not at all confident) to 100 (absolutely confident).

Experimental stimulation was controlled using Presentation® (Neurobehavioral Systems, Albany, CA, USA).

Data recording and analysis

Ratings of US valence and arousal. Paired t-tests were applied to confirm that US-crime sentences were perceived as more aversive and arousing than US-occu sentences.

Ratings of CS valence and arousal. Following the typical analysis approach for evaluative conditioning (e.g., Hermans et al., 2002), ANOVAs including the factors Session (pre-, post-conditioning), CS-type (CS-crime, CS-occu) and Report (CS-crime, CS-occu) were calculated to assess conditioning-induced changes in CS-valence and CS-arousal and test, whether these evaluative changes were driven by the actual pairing (i.e., CS-type), the perceived pairing (i.e., CS-type reported in surprise-recall task), or both.

Awareness of CS-US contingency. The sensitivity index d’ (Green and Swets, 1966) based on the performance in the surprise recall task was tested against the test value 0 by one-sample t-test. A potential response bias (e.g., the tendency to associate a face with a crime rather than occupation context), was assessed by t-test of the log β values (Wickens, 2002).

MEG data. Visual evoked magnetic fields (VEMFs) were recorded using a 275 MEG whole-head sensor system (VSM Medtech Ltd.) with first-order axial gradiometers. The individual head position in the MEG scanner was tracked by markers on the two ear canals and the nasion. The individual head shape and coordinate system was determined by a Polhemus 3Space® Fasttrack. VEMF’s recorded with an A/D rate of 600 Hz were down-sampled offline to 300 Hz and filtered between 0.1 and 48 Hz. Epochs of 800 ms duration (200 ms before to 600 ms after CS onset) were extracted, aligned and baseline-adjusted using a 150-ms pre-stimulus onset interval. Single trials were edited and artifacts were corrected following the method for statistical control of artifacts in high-density EEG/MEG data proposed by Junghöfer et al. (2000). The number of interpolated sensors and rejected trials did not differ between the experimental conditions.

After averaging, radial magnetic field measures were transformed to root-mean-square (RMS) fieldmaps of planar gradiometers (RMS-planar). After transformation, RMS-planar fieldmaps generated post learning were baseline adjusted by the corresponding pre-learning fieldmaps extracted from the first MEG run. Paired t-tests were used to test for differences between these baseline adjusted CS-crime and CS-occu fieldmaps. To correct for multiple comparisons and to consider potential deviations from normal distribution, nonparametric cluster level statistics were applied (Maris and Oostenveld, 2007). T-values of spatially adjacent3 and temporally consecutive4 RMS-planar data, where t-tests exceeded a critical alpha-level of P = .05 (sensor-level criterion), were summed into clusters.

3 To prevent merging of neighboring but not directly adjacent clusters, adjacent sensors were defined using the following procedure: The radius of virtual spheres around all individual sensors was increased until the maximum number of spheres contained five adjacent sensors. Sensors within this final sphere were defined as directly adjacent to the central sensor.

4 Here minimally five consecutive time points representing an interval of 16.7 ms at 300 Hz had to exceed the alpha-level. This setting reduces processing time but still enabled discovery of transient physiological effects. A liberalization of this criterion to three consecutive time points (10 ms) did not change the reported results qualitatively.
Cluster masses were compared against a random permutation cluster-based alpha-level of $P = 0.05$, which was established via Monte Carlo simulations of 1,000 permutations of experimental data sets. Only clusters exceeding an alpha-level of $P = 0.05$ were considered (cluster-level criterion). Based on theoretical expectations, random permutation tests were run separately for early (0–200 ms), mid-latency (200–400 ms) and late (400–600 ms) time intervals. Clusters that survived the non-parametric permutation tests were further evaluated by parametric post-hoc $t$-tests.

The cortical sources of neural activity were estimated using the L2-Minimum-Norm-Estimates method (L2-MNE; Hamalainen and Ilmoniemi, 1994), an inverse modeling technique applied to reconstruct the topography of the primary current underlying the magnetic field distribution. It estimates distributed neural network activity without a priori assumptions regarding the location and/or number of current sources (Hauk, 2004). A spherical shell with evenly distributed dipoles in azimuthal and polar direction at 350 positions was used as source model. A spherical shell is a reasonable approximation of the cortical surface and circumvents the necessity for the regularization of quasi-radial sources in more realistic MEG head modeling (Steinsträter et al, 2010). A source shell radius of 87% of the individually fitted head radius was chosen, roughly

---

Fig. 2. (A) Change in hedonic valence rating across sessions (i.e. pre- and post-multi-CS conditioning) depending on CS-type. Displayed are mean ratings of faces paired with an aversive criminal (red line) or a neutral occupational context (blue line). (B) Change in hedonic valence rating across sessions depending on actual CS-type (i.e. experimental contingency) and reported CS-type (i.e. contingency reported in the surprise recall task). Displayed are mean ratings of faces actually paired with a criminal (red lines) or an occupational context (blue lines) split by whether they were reported as having been paired with a criminal (solid lines) or an occupational context (dashed lines). (C) Change in emotional arousal rating across sessions depending on CS-type (colors as in A). (D) Change in emotional arousal rating across sessions depending on actual CS-type and reported CS-type (colors and lines as in B). For A–D, error bars represent standard errors.

---

Cluster masses were compared against a random permutation cluster-based alpha-level of $P = 0.05$, which was established via Monte Carlo simulations of 1,000 permutations of experimental data sets. Only clusters exceeding an alpha-level of $P = 0.05$ were considered (cluster-level criterion). Based on theoretical expectations, random permutation tests were run separately for early (0–200 ms), mid-latency (200–400 ms) and late (400–600 ms) time intervals. Clusters that survived the non-parametric permutation tests were further evaluated by parametric post-hoc $t$-tests.

The cortical sources of neural activity were estimated using the L2-Minimum-Norm-Estimates method (L2-MNE; Hamalainen and Ilmoniemi, 1994), an inverse modeling technique applied to reconstruct the topography of the primary current underlying the magnetic field distribution. It estimates distributed neural network activity without a priori assumptions regarding the location and/or number of current sources (Hauk, 2004). A spherical shell with evenly distributed dipoles in azimuthal and polar direction at 350 positions was used as source model. A spherical shell is a reasonable approximation of the cortical surface and circumvents the necessity for the regularization of quasi-radial sources in more realistic MEG head modeling (Steinsträter et al, 2010). A source shell radius of 87% of the individually fitted head radius was chosen, roughly
corresponding to gray matter depth. Across all participants and conditions, a Tikhonov regularization parameter k of 0.1 was applied. Topographies of source-direction-independent neural activities—the vector length of the estimated source activities at each position—were calculated for each individual participant, condition and time point. For visualization purposes, L2-MNE results were finally projected onto a model brain.

Results

US rating

Confirming stimulus selection, US-crime sentences were rated as more aversive, t(20) = −15.45, P < .001, and more arousing, t(20) = −9.89, P < .001, than US-occu sentences.

CS rating

Valence. A trend-level main effect of Session, F(1,19) = 4.24, P = .053, a main effect of Report, F(1,19) = 7.52, P = .013 and an interaction of Session × CS-type, F(1,19) = 9.69, P = .006 were found. Effects of Session and Session × CS-type indicated that pleasantness ratings decreased for all faces across sessions, t(19) = −2.06, P = .053, with the decrease in pleasantness being more pronounced for faces paired with a criminal relative to an occupational context, t(19) = −3.11, P = .006 (Figure 2A). The main effect of Report showed that, independently of the actual contingency, participants rated faces as less pleasant when they consciously associated them with a criminal rather than an occupational context (Figure 2B), t(19) = −2.74, P = .013. Importantly, the Session × CS-type interaction was unaffected by perceived contingency, suggesting that differences in CS valence after relative to before conditioning were driven by both emotional learning during multi-CS evaluative conditioning and deliberate categorization.

Arousal. Convergent to valence, participants tended to rate all CSs as more arousing after relative to before conditioning (Figure 2C), but this tendency was somewhat more pronounced for CS-crime than CS-occu [Session × CS-type: F(1,19) = 4.15, P = .056]. Arousal ratings increased, at least in tendency, for “crime faces” only, t(20) = 1.93, P = .067, but not for “occupation faces”, t(20) = 1.33, P = .200.

A main effect of Report, F(1,19) = 7.22, P = .015, and an interaction of Session × Report, F(1,19) = 6.60, P = .019, indicated that throughout the experiment, but more so after relative to before conditioning, t(19) = 2.57, P = .019, all faces consciously associated with a criminal context were perceived as more arousing than faces associated with an occupational context (Figure 2D). Faces reported as having been paired with a criminal context were rated to be in tendency more arousing after conditioning than before conditioning, t(19) = 2.07, P = .053, while arousal ratings did not differ between sessions for faces reported as having been paired with an occupational context, t(19) = 0.70, P = .494. However and importantly, the interaction of Session × CS-type remained unaffected by the reported contingency which again points towards simultaneously contributing, but independent effects of implicit emotional learning and contingency awareness on emotional arousal.

Contingency awareness

Participants were able to report CS-face/US-sentence pairings above chance, as d’-values, M = 0.59, s.d. = 0.47, differed significantly from zero, t(20) = 5.70, P < .001. As such, hit rates, M = 62.14%, s.d. = 17.28%, were on average higher than false alarm rates, M = 42.86%, s.d. = 18.69%. Participants did not show a response bias towards either criminal context or occupational context, because log β was not significantly different from zero, t(20) = −0.77, P = .451. Confidence ratings were higher for hits (i.e. CS-crime correctly categorized as CS-crime) than false alarms (i.e. CS-occu falsely categorized as CS-crime), t(20) = 2.99, P = .007.

MEG sensor space

A widely distributed spatio-temporal cluster covering medial prefrontal, bilateral orbitofrontal, bilateral dorsolateral prefrontal and specifically left hemispheric temporal areas revealed enhanced RMS-planar values of CS-crime compared with CS-occu faces in a time interval between 220 and 320 ms, t(20) = 5.52, P < .00001 (Figure 3 top). While RMS-planar amplitudes evoked by CS-occu faces in this cluster decreased from the pre- to the post evaluative learning phase, t(20) = −3.08, P = .006, corresponding values evoked by CS-crime did not change, t(20) = 1.0, P = .330. A post-hoc separation of the distributed cluster into PFC and left-temporal regions of interest (ROI) revealed qualitatively identical effects for both areas (Figure 3, bottom). For a post hoc laterality test of a seemingly left lateralized temporal ROI, we mirrored the left spatio-temporal cluster to the right hemisphere and calculated a CS (CS-crime, CS-occu) × Hemisphere (left, right) ANOVA. This analysis revealed a trend for a left hemispheric lateralization [CS × Hemisphere: F(1,20) = 3.65, P = .071; bottom right]. No spatiotemporal cluster in the early [0–200 ms; biggest cluster: P = 0.235] or late (400–600 ms; biggest cluster: P = 0.795) time intervals survived the cluster level significance criterion.

MEG source space

Generator estimates revealed a left hemispheric ventro-occipital area with relatively enhanced CS-crime processing, t(20) = 2.31; P = .032 (Figure 4, left). Post hoc t-tests revealed that estimated neural activation generated by CS-occu faces in this cluster decreased from the pre- to post-learning phase, t(20) = −2.12, P = .047, while cluster activations evoked by CS-crime did not change, t(20) = 0.76, P = .456. A post hoc CS × Hemisphere ANOVA—based on left to right mirrored spatio-temporal ventro-occipital cluster—revealed a significant CS × Hemisphere interaction, F(1,20) = 6.80, P = .017, but no main CS effect (P = .12), since CS processing similarly decreased for both affective conditions [CS-crime: t(20) = −2.24, P = .037; CS-occu: t(20) = −2.90, P = .009] in the right, but not in the left hemisphere (Figure 4, right). Additionally, medial prefrontal, right parietal and bilateral auditory sensory cluster showed trends for relatively enhanced CS-crime processing (not shown), indicating activity in a widely distributed network potentially contributing to the convergent, distributed fronto-temporal RMS-planar effects.

Discussion

In order to expand scientific understanding of a common and ecologically valid mechanism of higher-order affective learning, this study investigated short-term verbal evaluative conditioning in a Multi-CS paradigm with neutral-expression faces as CS and negative- or neutral-content sentences as US. During evaluative learning, 60 different faces were paired only twice with
30 negative or 30 neutral person descriptions. Neuromagnetic responses as well as changes in stimulus evaluation and contingency awareness were assessed. In sensor space, a widely distributed cluster covering bilateral prefrontal and left hemispheric occipito-temporal areas revealed enhanced magnetic fields evoked by CS-crime compared with CS-occu faces between 220 and 320 ms after face onset. In source space, these effects were reflected in enhanced CS-crime processing in left visual areas as well as, in tendency, by enhanced medial prefrontal, right parietal, and bilateral auditory sensory activation. No rapid effects, occurring before 100 ms, or late effects after 400 ms, emerged. Behaviorally, crime-paired faces were evaluated more negatively after conditioning than before, and also acquired more arousal than neutrally paired faces. Despite there being a multitude of faces and sentences, participants were able to report contingencies above chance level. Changes in the faces’ affective appraisal were driven both by contingency aware and unaware mechanisms.

The neuromagnetic effects are largely in line with previous reports of mid-latency effects in multi-CS conditioning (Steinberg et al., 2012, 2013), supporting similar mechanisms of learning via first-order and second-order US at this processing stage. In line with previous multi-CS conditioning studies with pre-learning/learning/post-learning design, the Session × CS-type interaction mainly derived from relatively stronger decreases of activation for CS-faces paired with neutral compared with aversive US from the pre- to the post-learning session, suggesting reduced habituation of CS-faces with acquired salience.

Notably however, the absence of early effects suggests potentially important differences to previous multi-CS conditioning studies: whereas conditioning with first-order US such as odors, electric shocks, or aversive noises, modulates face processing already at early perceptual stages (Steinberg et al., 2012, 2013; Rehbein et al., 2015), verbal evaluative conditioning seems to exert its impact from mid-latency conceptual processing stages only. This appears in line with other previous reports of

---

Fig. 3. Top: a widely distributed spatio-temporal sensor cluster covering prefrontal and left occipito-temporal regions in the time interval between 220 and 320 ms post CS-face onset revealed significantly stronger magnetic fields generated by faces previously paired with a criminal context when compared with faces associated with an occupational context. Bottom: cluster masses in the corresponding prefrontal and temporal regions of interest. Error bars depict the standard error.

Fig. 4. Left: statistical analysis of estimated sources in the 220–320 ms time interval showing significant effects in sensor space (see Figure 3) revealed a left occipital cluster with relatively increased neural processing of CS-crime compared with CS-occu faces. Right: a comparison with a right hemispheric mirror region revealed that this difference occurred predominately within the left hemisphere. Error bars depict the standard error.
differences between experience-dependent, observational and verbally mediated learning and in particular a reliance on an extended left-hemisphere network in verbally mediated learning (Phelps et al., 2001; for review Olsson and Phelps, 2007). In fact, the left lateralization of the occipital source space effects in the time interval between 220 and 320 ms diverges from the predominantly right-hemispheric effects found in previous first-order multi-CS face conditioning studies and neurophysiological indices of face processing in general (e.g. Rossion et al., 2003) but converges with the predominant left lateralization of emotion effects in affective word processing (e.g. Kliessl et al., 2007, 2009; Herbert et al., 2009).

Verbally mediated learning, similar to verbally represented affect itself may have less access to rapid affective stimulus processing, putatively mediated via sub-cortical routes (Olsson and Phelps, 2007; Hoffmann et al., 2015) and may require conceptual analysis before emotional responses occur. Visual processing of emotional words themselves differs most robustly from neutral ones at mid-latency processing stages whereas earlier effects are rare and might result from learning experiences akin to conditioning with first-order US (Fritsch and Kuchinke, 2013; Kuchinke et al., 2015). Apparently, in verbal multi-CS evaluative learning, the conceptual significance of the second-order US is transferred to the CS, in turn affecting subsequent conceptual, rather than perceptual CS processing.

The current lack of rapid conditioning effects could be due to specific US properties, such as auditory presentation and US complexity. However, we think this is less likely, because rapid conditioning effects have been observed with simple (Rehbein et al., 2015) as well as complex and temporally extended (6s) auditory US (JADS sounds; Bröckelmann et al., 2011; Junghöfer et al., 2015).

Because evaluative changes were due to both contingency-aware and unaware mechanisms, it is possible that effects of contingency-awareness differentially affect early and mid-latency responses. On the other hand, previous studies reported both early and mid-latency changes in the absence of (Bröckelmann et al., 2011; Steinberg et al., 2012) as well as in the presence of contingency awareness (Fritsch and Kuchinke, 2013; Kuchinke et al., 2015) suggesting that these early effects are not directly correlated with contingency awareness.

Moreover, other data indicate that, at least in the EEG, effects of explicit episodic memory on emotion face recognition (Johansson et al., 2004; Righi et al., 2012) as well as on the recognition of faces varying in explicitly learned associated biographies, appear even later, from around 400 ms after stimulus onset (Abdel Rahman, 2011). Accordingly, in a previous explicit learning and memory study (Strehlow and Kliessler, 2012), using the same material as here, emotion-modulated differences between baseline and test occurred around 400 ms, but not on earlier components. The present contingency awareness test required participants to decide whether a face was paired with a negative or neutral description which participants were able to do above chance and which is indicative of explicit learning contributing to the results. However, the ratings of the incorrectly assigned faces showed the same pattern, albeit less pronounced. Together, this supports the notion that both aware and unaware mechanisms contributed to the present results and to evaluative conditioning in general (Hofmann et al., 2010). Future studies will have to address this issue in greater detail, as verbally mediated learning has been argued to depend on stimulus, and possibly also contingency, awareness (Olsson and Phelps, 2004, 2007). On the other hand, behavioral data have suggested evaluative conditioning to occur even for subliminally presented US words (De Houwer et al., 1994).

To minimize variance and enhance effect sizes only female subjects were committed to this study and only male CS faces were linked to the violent crime US sentences. In spite of possible quantitative differences due to differential US arousal and valence ratings, there are no obvious reasons to expect qualitatively different mechanisms of evaluative learning in males and females. However, sex specific effects of emotional perception and learning are relevant (Cahill, 2006), and their potential effects on evaluative conditioning should be investigated in future research. The current sample size is similar to our previous multi-CS conditioning studies that all had between 20 and 24 participants and reported both rapid and mid-latency effects. Samples of this size are relatively common in the neuroscience literature where many studies contain fewer than 20 participants. On the other hand, we cannot exclude the possibility of false negative effects, due to lack of power or because the sample might in some respects not be representative of the population of young females as a whole. This is particularly true with regard to the null finding of associative learning effects in the early time interval (0–200 ms). Although statistical effects in this time interval were far from significant, weak learning effects could become verifiable within a larger sample.

The question whether real-life plausibility of CS-US matching or some kind of preparedness (Ohman, 2001; Ohman and Mineka, 2003) are relevant for evaluative conditioning is also an interesting though complex issue: of course, the “cold-blooded prisoner” who “strangled people” in our example is in real-life typically male but a female strangler is not impossible and might in fact, because of its rarity, evoke an even stronger, potentially qualitatively different, arousal driven CS-US link. In classic social psychology, evaluative conditioning effects of positive associations have also been shown and, in particular, their potential to counteract stereotypes has been demonstrated (Olsson and Fazio, 2006). Whether the biological mechanisms of positive and negative verbal evaluative learning are identical, with primarily arousal or salience mediating any effects, or differ in important ways, for instance due to differential recruitment of dopamine systems as has been shown for error-based learning (e.g. Holroyd et al., 2004; Yeung and Sanfey, 2004), remains to be determined. Plausibility and preparedness, as well as valence and arousal, are likely to play a role in verbal evaluative learning and the multi-CS paradigm commends itself for the investigation of research questions surrounding this socially important human learning mechanism.

Overall, following evaluative conditioning with verbal higher-order US, the present findings reveal changes of mid-latency face-evoked responses. In cortical source space these effects localized to predominantly left hemispheric visual areas. In this regard, verbal evaluative conditioning is similar to conditioning with first-order USs. However, the absence of rapid perceptual and the left lateralization effects suggests that verbal evaluative conditioning might not modulate face processing via fast sub-cortical and/or rapid magnocellular cortical pathways but only affects later conceptual processing stages. Thereby, these findings help inform models of affective associative learning.

Acknowledgements

We thank Karin Berning, Hildegard Deitermann, Ute Trompeter and Andreas Wollbrink for their help with data acquisition.
Funding
This work was supported by the Deutsche Forschungsgemeinschaft grants SFB-TRR58-C01 to M.J. and DFG KI1283/4-1 to J.K.

Conflict of interest. None declared.

References
Abdel Rahman, R. (2011). Facing good and evil: early brain signatures of affective biographical knowledge in face recognition. Emotion, 11(6), 1397–405.
Adolphs, R. (2002). Neural systems for recognizing emotion. Current Opinion in Neurobiology, 12(2), 169–77.
Balas, R., Swinkel, J. (2012). Evaluative conditioning may occur with and without contingency awareness. Psychological Research, 76(3), 304–10.
Blau, V.C., Maurer, U., Tottenham, N., McCandliss, B.D. (2007). The face-specific N170 component is modulated by emotional facial expression. Behavioral Brain Functions, 3(7), 13.
Bradley, M., Lang, P.J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. Journal of Behavior Therapy and Experimental Psychiatry, 25(4), 49–59.
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. Journal of Neuroscience, 31(21), 7801–10.
Bröckelmann, A.K., Steinberg, C., Dobel, C., et al. (2013). Affect-specific modulation of the N1m to shock-conditioned tones: magnetoencephalographic correlates. European Journal of Neuroscience, 37(2), 303–15.
Brosch, T., Sander, D., Poutou, G., Scherer, K.R. (2008). Beyond fear rapid spatial orienting toward positive emotional stimuli. Psychological Science, 19(4), 362–70.
Bullier, J. (2001). Integrated model of visual processing. Brain Research Reviews, 36, 96–107.
Citron, F.M. (2012). Neural correlates of written emotion word processing: a review of recent electrophysiological and hemodynamic neuroimaging studies. Brain and Language, 122(3), 211–26.
Cahill, L. (2006). Why sex matters for neuroscience. Psychological Bulletin, 132(4), 390–421.
Ciftci, O., Junghöfer, M., Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., Crombez, G. (2010). Evaluative conditioning in humans: a meta-analysis. Psychophysiology, 136(3), 390–421.
Dias, S., Vuilleumier, P., editors. (2004). Keep it simple: a case for using classical minimum norm estimates. Trends in Cognitive Sciences, 4(1), 35–49.
FBI. (2014). Persons arrested, Uniform crime reports, Crime in the United States. https://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u-s/2014/persons-arrested/main.
Fritsch, N., Kuchinke, L. (2013). Acquired affective associations induce emotion effects in word recognition: an ERP study. Brain and Language, 124(1), 75–83.
Haxby, J.V., Hoffman, E.A., Gobbini, M.I. (2000). The distributed human neural system for face perception. Trends in Cognitive Sciences, 4(6), 223–33.
Hauk, O. (2004). Keep it simple: a case for using classical minimum norm estimation in the analysis of EEG and MEG data. Neuroimage, 21(4), 1612–21.
Herbert, C., Ethofer, T., Anders, S., et al. (2009). Amygdala activation during reading of emotional adjectives—an advantage for pleasant content. Social Cognitive and Affective Neuroscience, 4(1), 35–49.
Hintze, P., Junghöfer, M., Bruchmann, M. (2014). Evidence for rapid prefrontal emotional evaluation from visual evoked responses to conditioned gratings. Psychological Bulletin, 99, 125–36.
Holt, C.B., Larsen, J.T., Cohen, J.D. (2004). Context dependence of the event-related brain potential associated with reward and punishment. Psychophysiology, 41(2), 245–53.
Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., Crombez, G. (2010). Evaluative conditioning in humans: a meta-analysis. Psychological Bulletin, 136(3), 390–421.
Hoffmann, M., Mothes-Lasch, M., Milchner, W.H.R., Straube, T. (2015). Brain activation to briefly presented emotional words: effects of stimulus awareness. Human Brain Mapping, 36 (2), 655–66.
Hütter, M., Sweldens, S., Stahl, C., Unkelbach, C., Klauser, K.C. (2012). Dissociating contingency awareness and conditioned attitudes: evidence of contingency-unaware evaluative conditioning. Journal of Experimental Psychology: General, 141(3), 539.
Johansson, M., Mecklinger, A., Treece, A.C. (2004). Recognition memory for emotional and neutral faces: an event-related potential study. Journal of Cognitive Neuroscience, 16(10), 1840–53.
Kissler, J., Herbert, C., Eyrek, P., Junghöfer, M. (2007). ‘Buzzwords’—spontaneous cortical responses to emotional words. Psychological Science, 18(6), 475–80.
Kissler, J., Herbert, C., Winkler, I., Junghöfer, M. (2009). Emotion and attention in visual word processing—An ERP study. Biological Psychology, 80(1), 75–83.
Kissler, J. (2013). Love Letters and Hate Mail. In: Armony, J., Vuilleumier, P., editors. The Cambridge Handbook of Human Affective Neuroscience, Oxford University Press: 304–22.

Conflict of interest. None declared.
Walther, E. (2002). Guilty by mere association: evaluative conditioning and the spreading attitude effect. *Journal of Personality and Social Psychology, 82*(6), 919.

Warr, M. (2000). Fear of crime in the United States: avenues for research and policy. In: Duffee D. (ed.) *Measurement and Analysis of Crime: Criminal Justice 2000*. Washington, DC: US Department of Justice, Office of Justice Programs.

Wickens, T.D. (2002). *Elementary Signal Detection Theory*. New York: Oxford University Press.

Yeung, N., Sanfey, A.G. (2004). Independent coding of reward magnitude and valence in the human brain. *The Journal of Neuroscience, 24*(28), 6258–64.