A rift between implicit and explicit conditioned valence in human pain relief learning

Marta Andreatta[1,*], **Andreas Mühlberger**[1], **Ayse Yarali**[2],
Bertram Gerber[2] **and Paul Pauli**[1,*]

[1]*Department of Psychology, and* [2]*Department of Neurobiology and Genetics, University of Würzburg,
Würzburg, Germany*

Pain is aversive, but does the cessation of pain ('relief') have a reward-like effect? Indeed, fruitflies avoid an odour previously presented before a painful event, but approach an odour previously presented after a painful event. Thus, event-timing may turn punishment to reward. However, is event-timing also crucial in humans who can have explicit cognitions about associations? Here, we show that stimuli associated with pain-relief acquire positive implicit valence but are explicitly rated as aversive. Specifically, the startle response, an evolutionarily conserved defence reflex, is attenuated by stimuli that had previously followed a painful event, indicating implicit positive valence of the conditioned stimulus; nevertheless, participants explicitly evaluate these stimuli as 'emotionally negative'. These results demonstrate a rift between the implicit and explicit conditioned valence induced by pain relief. They might explain why humans in some cases are attracted by conditioned stimuli despite explicitly judging them as negative.

Keywords: relief; pain; startle reflex

1. INTRODUCTION

Primary reinforcers are either positive (e.g. food) or negative (e.g. noxious events), and stimuli associated with primary reinforcers will generally acquire similar qualities.

However, event-timing is a crucial determinant for this process. For example, Dickinson & Dearing (1979) pointed out that a stimulus predicting the absence of a threat can have similar behavioural effects as a stimulus predicting an appetitive event. Corroborating this assumption, fruitflies were found to avoid an odour previously associated with the onset of an electric shock, as it predicts danger, but approach an odour previously associated with the shock's offset, as it predicts safety (Tanimoto *et al.* 2004; Yarali *et al.* 2008). Using such a backward conditioning paradigm, similar results have been found in mice and rats as well (Cunningham *et al.* 2002; Salvy *et al.* 2004). In humans, fear studies corroborate that a stimulus preceding a shock becomes a predictor for danger (Lipp *et al.* 1994; Grillon 2002). However, it remains unknown whether in humans a stimulus following an aversive event later on predicts safety.

Notably, dual-process theories (Strack & Deutsch 2004; Bechara 2005) propose that human behaviour is determined by the output of two systems, an impulsive, implicit system working on associative principles, and a system operating on the basis of reflective, explicit knowledge about facts and values. Importantly, these systems can operate in a synergistic or antagonistic fashion and it remains to be clarified how event-timing affects these systems and/or their interaction. Understanding event-timing in humans and its potentially dissociated effects

on the impulsive and reflective system may be relevant for psychopathologies like drug addiction (Koob & LeMoal 2001; Weiss 2005), or anxiety disorders (Bouton *et al.* 2001; Mineka & Oehlberg 2008): both disorders are characterized by emotionally intensely negative episodes such as drug withdrawal and anxiety attacks, respectively, and stimuli associated with the offset of these events may become appetitive.

To examine the effects of event-timing in humans, we compared three groups of participants who had undergone stimulus-discrimination learning reinforced by an aversive unconditioned event, i.e. a moderately painful electric shock; what differed between the three experimental groups is the relative timing of stimulus and this shock. On the one hand, the valence of the stimulus associated with this electric shock is assessed on the basis of explicit, subjective reports. On the other hand, the implicit valence of the stimulus is assessed by probing for its capacity to modulate the startle response (Lang *et al.* 1998); this is an especially suitable method because it allows us to assess both positive and negative valence effects within one setting, and because it allows translational studies between humans and rodents. That is, the startle response is an evolutionarily conserved defence reflex, the neural mechanisms of which have been studied in detail (Koch 1999; Davis 2006). Significantly, if animals or humans are trained to associate a stimulus (e.g. a light) with a painful electric shock, the startle response amplitude is potentiated in the presence of that shock-predicting stimulus (Lipp *et al.* 1994; Koch 1999; Grillon 2002; Davis 2006). Such an associative increase of startle amplitude indicates fear and is mediated by connections from the amygdala impinging upon the startle-reflex circuitry (Davis 2006). However, when the stimulus is trained to predict a rewarding sucrose solution, the startle

*Authors for correspondence (marta.andreatta@mail.uni-wuerzburg.de; pauli@mail.uni-wuerzburg.de).

amplitude is attenuated in its presence (Koch *et al.* 1996; Schneider & Spanagel 2008). This associative attenuation may involve dopaminergic projections of the nucleus accumbens (NAcc) onto the startle circuitry (Schultz 2006). Also in humans, attenuation of the startle response has been found in the presence of pleasant stimuli (Lang *et al.* 1998) or stimuli signalling monetary gain (Skolnick & Davidson 2002). Therefore, we reasoned that a safety-predicting stimulus may attenuate the startle response in humans. Thus, regarding danger- or safety-predicting stimuli we complement explicit valence judgments with startle modulation as an implicit measure of valence.

2. MATERIAL AND METHODS

(a) *Participants*

A total of 101 healthy volunteers (68 females: mean age = 23.2 years, s.d. = 4.6, range = 18–43 years) were divided into three groups, which only differed in the relative timing of the conditioned stimulus (CS+) and the unconditioned stimulus (US) during *acquisition* (training phase): one group of 34 participants underwent forward delay conditioning (FORWARD: US onset 8 s after CS+ onset), one group of 34 participants a forward trace conditioning (CONTROL: US onset 14 s after CS+ onset), and the third group of 33 participants received backward conditioning (BACKWARD: US onset 6 s before CS+ onset). Participants were free of neurological, psychiatric or chronic pain diseases.

(b) *Stimulus material and apparatus*

The US was a single unipolar electric shock of 200 ms duration generated by a battery-driven constant-current stimulator (maximum of 140 V and of 10 mA) delivered via a surface bar electrode (which consisted of two durable gold-plated stainless steel disc electrodes with 9 mm diameter and 30 mm spacing) attached to the left forearm (Neumann & Waters 2006). US intensity was assessed before the experimental session. Each participant received two series of electrical stimuli with ascending and two with descending intensity in steps of 0.5 mA (Reiff *et al.* 1999). Participants evaluated the intensity of each electrical stimulus on a rating scale ranging from 0 (no pain at all) until 10 (unbearable pain). The mean value of the intensities rated as 'just noticeable pain' (i.e. 4) was defined as pain threshold and increased by 1 mA.

CSs were three simple geometrical shapes (a square, a circle and a equilateral triangle), all solid yellow, 12 cm in width and 12 cm in height (see Lipp *et al.* 1994). These stimuli were always presented on a 19″ black computer screen for 8 s. The screen was located 140 cm from the participants at eye level. Shapes could function either as a reinforced CS (CS+), which was always associated with the US (in either a FORWARD, CONTROL or BACKWARD way), as a non-reinforced CS (CS−), which was never associated with the US, or a new stimulus (NEW), which was presented only during the test phase (i.e. extinction). The shapes' functions were counterbalanced among participants.

The *startle stimulus* was a burst of white noise (50 ms, 105 dB) delivered binaurally with headphones. The eyeblink component of the startle response was measured through electromyography (EMG) of the left *orbicularis oculi* muscle with two 5 mm Ag/AgCl electrodes; one placed under the pupil of the left eye and the other

approximately 1 cm lateral. Both the ground and the reference electrodes were placed on the forehead. Before attaching the electrodes, the skin was cleaned with alcohol and slightly abraded to keep all electrode impedances below 5 kΩ (measured with Vision Recorder V-Amp Edition Software). The raw signal was sampled at 400 Hz. Startle responses were registered continuously with a V-Amp 16 using Vision Recorder V-Amp Edition Software (v. 1.03.0004). EMG activity was filtered online with a 50 Hz notch filter to eliminate 50 Hz interference.

(c) *Procedure and experimental design*

After having signed an informed consent form, participants were seated in a comfortable chair in a sound-attenuated room next to the experimenter room. After electrode attachment, the pain threshold was assessed. Participants were then informed that a series of geometrical shapes would be presented and that they should keep these pictures in their visual focus. Participants were also told that electrical stimuli would be delivered occasionally.

The experiment consisted of two phases, the training phase (i.e. acquisition) and the test phase (i.e. extinction). Additionally, before and after the training phase, the participants were asked to rate the three visual stimuli (CS+, CS− and NEW) in terms of their valence and arousal value. This *rating procedure* was realized as follows: first, the visual stimulus was presented for 1.2 s. Then, the valence and the arousal of the stimulus were assessed in succession. In both cases, a scale ranging from 1 to 9 appeared on the screen in front of the participants. For valence ratings, 1 was labelled 'very unpleasant' and 9 'very pleasant'; for arousal ratings, 1 was labelled 'calm' and 9 'exciting'. Participants had to respond by using a numeric keyboard. For analyses, valence and arousal data were transformed by subtracting 5 (as a consequence, negative values represent negative valence or low arousal, respectively, whereas positive values represent positive valence or high arousal; a value of zero represents 'neutral' ratings).

The *training phase* (acquisition) consisted of 32 trials: 16 presentations of CS+, always associated with the US and 16 presentations of CS−, never associated with US. The trial length varied between 28 and 44 s (mean of 36 s), the intertrial interval (ITI) varied between 20 and 30 s. The three experimental groups only differed in the time between CS+ onset and US onset (i.e. the interstimulus interval, ISI); for the FORWARD conditioning group, US onset coincided with CS+ offset, such that the US was delivered 8 s after CS+ onset (ISI = 8 s); for the BACKWARD conditioning group, US onset preceded CS+ onset by 6 s (ISI = −6 s); for the CONTROL conditioning group, US onset followed CS+ onset by 14 s (ISI = 14 s). No startle stimuli were delivered during the training phase.

The *test phase* (extinction) started with nine startle stimuli delivered every 7–15 s to decrease initial startle reactivity. Then, 48 extinction trials were run in a way identical to all participants; during the test phase, no US was delivered. Trial's length varied between 28 and 38 s (mean length 33 s), the ITI varied between 20 and 30 s. The three visual stimuli (CS+, CS− and NEW) were presented 16 times for 8 s each. In summary, 32 startle stimuli were delivered, eight in the presence of CS+, eight in the presence of CS− and eight in the presence of the NEW stimulus; thus, the respective visual stimuli were presented with or without the startle stimulus in half of the cases. These startle stimuli

occurred 3–7 s after visual stimulus onset. In order to enhance the unpredictability of the startle stimuli, eight additional startle stimuli were delivered during the ITI (not analysed).

(d) Data reduction and statistical analysis

Startle responses EMG data were analysed offline with the Brain Vision Analyser Software (v. 1.05, BrainProducts Inc.). Data were first filtered (low cut-off filter 28 Hz, high cut-off 500 Hz, moving average of 50 ms) and rectified. Then, startle response amplitude was determined for each trial as the peak startle response (the maximum in the 20–120 ms time window following the startle stimulus) relative to baseline defined as mean EMG activity over 50 ms preceding stimulus onset (see Grillon *et al.* 2006). Trials were excluded if the baseline EMG was not stable, or if the onset of the startle response was not within 20–60 ms after the startle probe onset. Startle response amplitude of each participant were standardized as a z-score ($z = (x - \mu)/\sigma$), where x is a raw score, μ is the mean which is zero and σ is the standard deviation which is 1) in order to normalize data and to reduce the influence of between-subjects variability unrelated to psychological processes (see Blumenthal *et al.* 2005). Finally, mean startle response amplitude for each participant and each CS type (CS+, CS− and NEW) were calculated on the basis of this z-score.

Startle data were analysed with an analysis of variance (ANOVA) including the between-subjects factor group (FORWARD, BACKWARD, CONTROL) and the within-subjects factor stimulus (CS+, CS− and NEW); the ANOVA for the valence and arousal ratings had the additional within-factor time (BEFORE, AFTER acquisition). The α level was set at 0.05 for all statistical tests. Greenhouse–Geisser corrections (GG-ε) were used for main effects and interactions involving factors with more than two levels.

All data were analysed using SPSS for Windows (Release 17.0).

3. RESULTS

Implicit valence ratings assayed by the modulation of startle revealed a significant interaction of stimulus × group ($F_{4,196} = 3.49$, $p = 0.009$; figure 1), underscoring the crucial role of the temporal sequence of events experienced during training; no other ANOVA effect reached significance. Follow-up tests (comparison of the group's mean with the overall mean of 0, i.e. the mean of the z-normalized distribution) indicate that a visual stimulus which during training had been presented briefly before (group FORWARD) an aversive event, later on induces a potentiation of the startle responses ($t_{33} = 2.91$, $p = 0.006$) indicating negative implicit valence. In the FORWARD group, startle potentiation by CS+ and CS− did not differ ($t_{33} = 0.72$, $p = 0.479$). Importantly, however, if a visual stimulus had been following an aversive event during training (group BACKWARD), the startle response is attenuated in its presence ($t_{32} = -2.1$, $p = 0.044$), indicating implicit positive valence (i.e. safety). Supportively, the attenuation of startle by CS+ was stronger than by the CS− ($t_{32} = -2.75$, $p = 0.010$). Training with a long ISI (group CONTROL) did not subsequently affect the startle responses ($t_{33} = -0.13$, $p = 0.894$). Consistently, in the
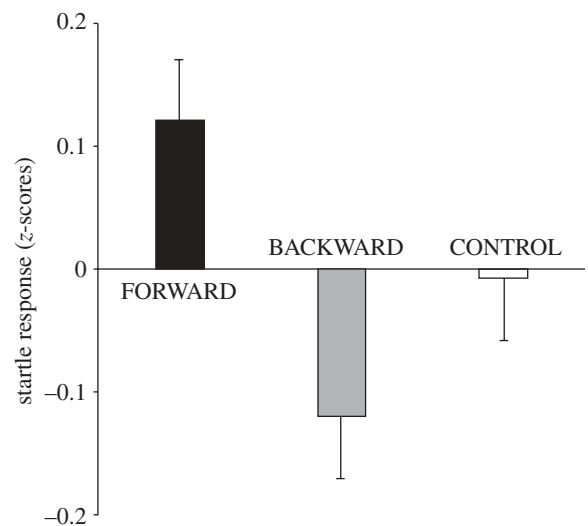

Figure 1. Bars represent mean (with standard errors) startle amplitudes in z-scores assessed in the test phase (i.e. extinction) in the presence of visual CS+; the black bar represents the FORWARD group, the grey bar the BACKWARD group, and the white bar the CONTROL group. Positive values indicate startle response potentiation; negative values startle attenuation relative to the mean. CS+ has been reinforced during the preceding training phase (i.e. the acquisition) with a different timing relative to the US (i.e. a mild electric shock). In the FORWARD and the BACKWARD group, the CS+ briefly preceded or followed the US, respectively, and consequently the CS+ had opposite effects on startle response modulation during test. In the CONTROL group, the CS+ preceded the US with a long delay and this did not subsequently affect the startle response.

CONTROL group, startle responses in the presence of CS+ did not differ from the one in the presence of CS− ($t_{33} = -0.71$, $p = 0.481$).

Exploiting the capacity of humans for explicit reports, we asked our participants to rate the valence and the arousal of the visual stimuli before and after the training phase. Importantly, both stimulus valence and arousal were rated as neutral before the training phase and the ratings at this time did not differ significantly between CS+, CS− and NEW. However, as indicated by significant interactions of stimulus × time ($F_{2,196} = 22.43$, $p = 0$, GG-ε = 0.91) and of stimulus × time × group ($F_{4,196} = 4.12$, $p = 0.004$, GG-ε = 0.91), CSs changed their valence through training (figure 2a). Comparisons within groups revealed that the CS+ was rated as 'emotionally' more negative after both forward *and* backward training compared with its initial pre-training rating (FORWARD: $t_{33} = 5.55$, $p < 0.001$; BACKWARD: $t_{32} = 2.3$, $p = 0.028$), but not after the control training (CONTROL: $t_{33} = 1.23$, $p = 0.226$). Comparisons between groups indicate that the FORWARD CS+ after training was rated as more negative when compared with the CONTROL CS+ ($t_{66} = 2.34$, $p = 0.023$), while this difference was not found before training ($t_{66} = -0.11$, $p = 0.910$). Although the valence of the BACKWARD CS+ after training valence did not differ significantly from the CONTROL CS+ valence after training ($t_{65} = 0.54$, $p = 0.590$), we note that both the FORWARD CS+ apparently acquired negative valence (FORWARD: $t_{33} = -4.65$, $p = 0$) and that the BACKWARD CS+ also acquired negative valence,
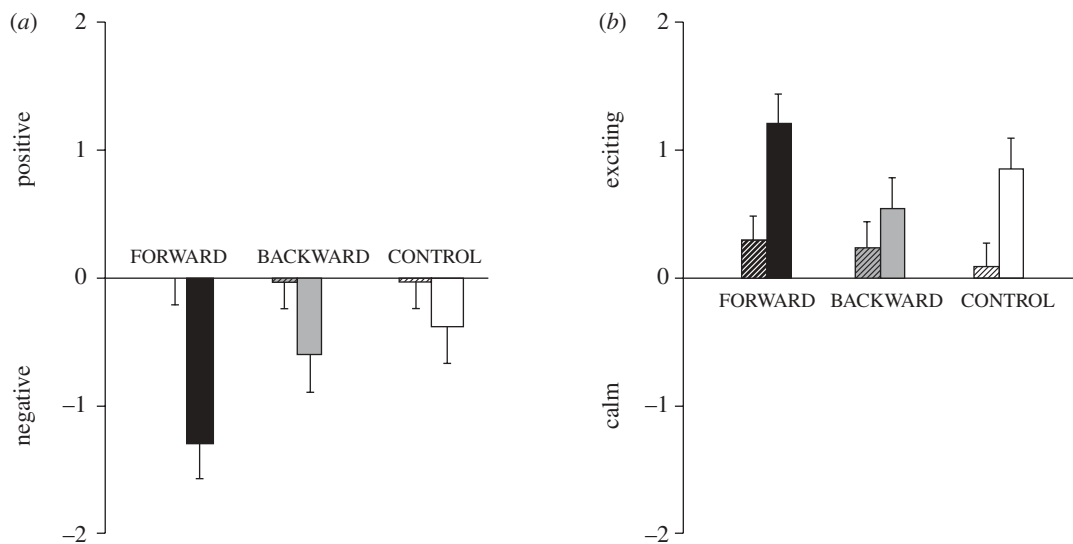
Figure 2. Bars represent mean scores (with standard errors) of (a) valence and (b) arousal ratings of the CS+ with the neutral value (i.e. 5) subtracted from the ratings. Hatched bars represent the ratings before the acquisition phase (i.e. conditioning) and filled bars after the acquisition phase. Black fill represents the FORWARD group, grey fill the BACKWARD group and white fill the CONTROL group. Negative values represent negative valence and low arousal, whereas positive values represent positive valence and high arousal. (a) Ratings of valence before training were neutral and after training were consistently negative indicating that the CS+ acquired negative explicit valence, independent of event-timing. (b) Ratings of arousal were in all three groups higher after compared with that before training, indicating that the CS+ became arousing independent of event-timing.

which just failed to reach significance (BACKWARD: $t_{32} = -1.95$, $p = 0.060$); no such trend was seen for the CONTROL CS+ ($t_{33} = -1.4$, $p = 0.172$). This is confirmed by analyses of the differences between pre- and post-training scores on the basis of differences between the valence of the forward CS+ or the backward CS+ and the control CS+ revealed for both the FORWARD (post-training (forward − control) − pre-training (forward − control): $t_{33} = 4.71$, $p < 0.001$) and the BACKWARD group (post-training (backward − control) − pre-training (backward − control), $t_{32} = 2.42$, $p = 0.021$), arguing that the CS+ acquired negative explicit valence in both groups.

For arousal ratings, the ANOVA also indicated a significant stimulus × time interaction ($F_{2,196} = 16.69$, $p = 0$, GG-ε = 0.91), but the interaction stimulus × time × group was not significant ($F_{4,196} = 0.56$, $p = 0.673$, GG-ε = 0.91; figure 2b). Follow-up tests indicate that both the FORWARD CS+ and the CONTROL CS+ were rated as more arousing after the training phase compared with their initial pre-training ratings (FORWARD: $t_{33} = -3.14$, $p = 0.004$; CONTROL: $t_{33} = -3.06$, $p = 0.004$); this was not the case for the BACKWARD CS+ ($t_{32} = -1.26$, $p = 0.216$). There were no significant differences comparing FORWARD or BACKWARD CS+ arousal with CONTROL CS+ arousal (FORWARD: $t_{66} = -1.13$, $p = 0.263$; BACKWARD: $t_{65} = 0.93$, $p = 0.355$). However, in absolute terms, the FORWARD, the BACKWARD and the CONTROL CS+ acquired higher arousal since the arousal ratings after training did significantly differ from 0 (FORWARD: $t_{33} = 4.91$, $p < 0.001$; BACKWARD: $t_{32} = 2.03$, $p = 0.051$; CONTROL: $t_{33} = 4.4$, $p < 0.001$).

4. DISCUSSION
Our findings reveal that event-timing determines the implicit valence of a CS in an opponent manner. On the one hand, a stimulus *signalling* an aversive event later potentiates startle responses, indicating that this stimulus acquired negative implicit valence. This result replicates animal as well as human fear conditioning studies (Lipp *et al.* 1994; Grillon 2002). On the other hand, a stimulus *following* an aversive event later attenuates startle responses, indicating that this stimulus acquired positive implicit valence. This new finding suggests that a stimulus associated with pain relief may activate reward circuits (Seymour *et al.* 2005; Leknes & Tracey 2008; Brischoux *et al.* 2009). Importantly, this observed opponency does not seem to be related to the omission of an expected aversive event (Dickinson & Dearing 1979), but rather to its termination (Solomon 1980; Seymour *et al.* 2005), since participants could correctly verbalize the association between the CS+ and the aversive US after both kinds of training. Finally, in humans this timing-dependent opponency appears restricted to implicit processes assessed on the basis of startle modulation. Explicit ratings of valence did not show timing-dependent opponency as both stimuli were rated as 'emotionally negative'.

The observed timing-dependent bidirectional modulation of human startle behaviour conforms to the opponent-process theory of acquired motivation (Solomon 1980), which suggests that an aversive stimulus such as a painful electric shock generates two opponent processes: an initial negative affect upon onset and an after-process entailing the opposite state, i.e. positive affect. A stimulus presented after the aversive event thus may become associated with this latter process and therefore acquire positive valence. Please note however that the present experiment cannot directly verify these assumed opponent processes.

Notably, on the neural level our results concerning startle modulation could be explained on the basis of spike timing-dependent plasticity (Drew & Abbott 2006). That is, the temporal sequence of two inputs

determines whether synapses are potentiated or depressed. This process can conceivably act at behaviourally relevant time scales (Abbott & Nelson 2000; Drew & Abbott 2006), in particular, if it was at operation in the amygdala and/or the dopamine neurons of NAcc (Brischoux *et al.* 2009), both structures relevant for associative startle response modulation (Davis 2006; Schneider & Spanagel 2008). Additionally, converging evidence indicates that potentiation or depression of synaptic firing in both amygdala (Delgado *et al.* 2006; Kim & Jung 2006) and NAcc (Wise 2004) are underlying mechanisms of associative memory. However, while amygdala responses have been found to habituate rapidly because of repeated stimulation (Quirk *et al.* 1997; LaBar *et al.* 1998; Büchel & Dolan 2000), no such results are reported for the NAcc. Given that we averaged across repeated extinction trials, this difference might explain why we observed no difference in startle amplitude in the presence of CS+ versus CS− after forward conditioning, but did find such a difference after backward conditioning. We further note that, in flies, effects of backward conditioning are apparently more stable over extended retention periods than the effects of forward conditioning (Yarali *et al.* 2008).

Importantly, our results suggest an event-timing-specific dissociation of implicit from explicit valence. Based on dual-process theories (e.g. Strack & Deutsch 2004), it might be speculated that after the offset of an aversive event, the impulsive, implicit system processes the experienced relief while the reflective, explicit system processes the overall aversiveness of the event. After all, the electric shock was painful and therefore aversive. This may be why a previously neutral stimulus presented briefly after an aversive event acquires positive implicit valence but nevertheless is explicitly evaluated as negative.

Further analyses of the implicit reward-like after-effects of aversive events in humans seem desirable in particular in the context of psychopathologies, e.g. anxiety disorders or drug addiction, including its modulation by genotype (see Yarali *et al.* 2009). For example, stimuli associated with the offset of a panic attack (e.g. the clinic or a physician) or the offset of withdrawal symptoms (e.g. the drug intake environment) may contribute to the maintenance of the disorder because they become appetitive and will be approached. Finally, although speculative, reward-like after-effects of aversive events may contribute to our understanding of 'paradoxical' human behaviours, like approaching stimuli that are explicitly evaluated as negative or dangerous (e.g. rollercoaster ride or bungee jumping).

In summary, the present study demonstrates that event-timing in humans crucially determines the implicit valence of a conditioned stimulus associated with an aversive event. Notably, stimuli associated with the offset of the aversive event acquire positive implicit valence. Moreover, the behaviour mediated by the impulsive, implicit system can dissociate from the expression of the reflective, explicit system. These findings should prompt studies to clarify the psychological and neuronal mechanisms behind the processes involved in event-timing and their implicit–explicit dissociation.

REFERENCES

Abbott, L. F. & Nelson, S. B. 2000 Synaptic plasticity: taming the beast. *Nat. Neurosci.* **3**, 1178–1183. (doi:10.1038/81453)

Bechara, A. 2005 Decision making, impulse control and loss of willpower to resist drugs: a neurocognitive perspective. *Nat. Neurosci.* **8**, 1458–1463. (doi:10.1038/nn1584)

Blumenthal, T. D., Cuthbert, B. N., Filion, D. L., Hackley, S., Lipp, O. V. & Van Boxtel, A. 2005 Committee report: guidelines for human startle eyeblink electromyographic studies. *Psychophysiology* **42**, 1–15. (doi:10.1111/j.1469-8986.2005.00271.x)

Bouton, M. E., Barlow, D. H. & Mineka, S. 2001 A modern learning theory perspective on the etiology of panic disorder. *Psychol. Rev.* **108**, 4–32. (doi:10.1037/0033-295X.108.1.4)

Brischoux, F., Chakraborty, S., Brierley, D. I. & Ungless, M. A. 2009 Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proc. Natl Acad. Sci.* **106**, 4894–4899. (doi:10.1073/pnas.0811507106)

Büchel, C. & Dolan, R. J. 2000 Classical fear conditioning in functional neuroimaging. *Curr. Opin. Neurobiol.* **10**, 219–223. (doi:10.1016/S0959-4388(00)00078-7)

Cunningham, C. L., Clemans, J. M. & Fidler, T. L. 2002 Injection timing determines whether intragastric ethanol produces conditioned place preference or aversion in mice. *Pharm. Biochem. Behav.* **72**, 659–668. (doi:10.1016/S0091-3057(02)00734-7)

Davis, M. 2006 Neural systems involved in fear and anxiety measured with fear-potentiated startle. *Am. Psychol.* **61**, 741–756. (doi:10.1037/0003-066X.61.8.741)

Delgado, M. R., Olsson, A. & Phelps, E. A. 2006 Extending animal models of fear conditioning to humans. *Biol. Psychol.* **73**, 39–48. (doi:10.1016/j.biopsycho.2006.01.006)

Dickinson, A. & Dearing, M. F. 1979 Appetitive-aversive interactions and inhibitory processes. In *Mechanisms of learning and motivation: a memorial volume to Jerzy Konorsky* (eds A. Dickinson & R. A. Boakes), pp. 203–232. Mahwah, NJ: Lawrence Erlbaum Associates.

Drew, P. J. & Abbott, L. F. 2006 Extending the effects of spike-timing-dependent plasticity to behavioral time-scales. *Proc. Natl Acad. Sci.* **103**, 8876–8881. (doi:10.1073/pnas.0600676103)

Grillon, C. 2002 Startle reactivity, and anxiety disorders: aversive conditioning, context and neurobiology. *Biol. Psychiatry* **56**, 958–975. (doi:10.1016/S0006-3223(02)01665-7)

Grillon, C., Baas, J. M., Cornwell, B. & Johnson, L. 2006 Context conditioning and behavioral avoidance in a virtual reality environment: effect of predictability. *Biol. Psychiatry* **60**, 752–759. (doi:10.1016/j.biopsych.2006.03.072)

Kim, J. J. & Jung, M. W. 2006 Neural circuits and mechanisms involved in Pavlovian fear conditioning: a critical review. *Neurosci. Biobehav. Rev.* **30**, 188–202. (doi:10.1016/j.neubiorev.2005.06.005)

Koch, M. 1999 The neurobiology of startle. *Prog. Neurobiol.* **59**, 107–128. (doi:10.1016/S0301-0082(98)00098-7)

Koch, M., Schmid, A. & Schnitzler, H. 1996 Pleasure-attenuation of startle is disrupted by lesions of the nucleus accumbens. *NeuroReport* **7**, 1442–1446. (doi:10.1097/00001756-199605310-00024)

Koob, G. F. & LeMoal, M. 2001 Drug addiction, dysregulation of reward, and allostasis. *Neuropsychopharmacology* **24**, 97–129. (doi:10.1016/S0893-133X(00)00195-0)

LaBar, K. S., Gatenby, J. C., Gore, J. C., LeDoux, J. E. & Phelps, E. A. 1998 Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fRMI study. *Neuron* **20**, 937–945. (doi:10.1016/S0896-6273(00)80475-4)

Lang, P. J., Bradley, M. M. & Cuthbert, B. N. 1998 Emotion motivation and anxiety: brain mechanisms and psychophysiology. *Biol. Psychiatry* **44**, 1248–1263. (doi:10.1016/S0006-3223(98)00275-3)

Leknes, S. & Tracey, I. 2008 A common neurobiology of pain and pleasure. *Nat. Rev. Neurosci.* **9**, 314–320. (doi:10.1038/nrn2333)

Lipp, O. V., Sheridan, J. & Siddle, D. A. T. 1994 Human blink startle during aversive and nonaversive Pavlovian conditioning. *J. Exp. Psychol. Anim. Behav. Proc.* **20**, 380–389. (doi:10.1037/0097-7403.20.4.380)

Mineka, S. & Oehlberg, K. 2008 The relevance of recent developments in classical conditioning to understanding the etiology and maintenance of anxiety disorders. *Acta Psychol.* **127**, 567–580. (doi:10.1016/j.actpsy.2007.11.007)

Neumann, D. L. & Waters, A. M. 2006 The use of the unpleasant sound as an unconditional stimulus in a human aversive Pavlovian conditioning procedure. *Biol. Psychol.* **73**, 175–185. (doi:10.1016/j.biopsycho.2006.03.004)

Quirk, G. J., Armony, J. L. & LeDoux, J. E. 1997 Fear conditioning enhances different temporal components of tone-evoked spike trains in auditory cortex and lateral amygdala. *Neuron* **19**, 613–624. (doi:10.1016/S0896-6273(00)80375-X)

Reiff, S., Katkin, E. S. & Friedman, R. 1999 Classical conditioning of the human blood pressure response. *Int. J. Psychophys.* **34**, 135–145. (doi:10.1016/S0167-8760(99)00071-9)

Salvy, S. J., Pierce, W. D., Heth, D. C. & Russell, J. C. 2004 Taste avoidance induced by wheel running: effects of backward pairings and robustness of conditioned taste aversion. *Physiol. Behav.* **82**, 303–308. (doi:10.1016/j.physbeh.2004.03.017)

Schneider, M. & Spanagel, R. 2008 Appetitive odor-cue conditioning attenuates the acoustic startle response in rats. *Behav. Brain Res.* **189**, 226–230. (doi:10.1016/j.bbr.2007.12.017)

Schultz, W. 2006 Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.* **57**, 87–115. (doi:10.1146/annurev.psych.56.091103.070229)

Seymour, B., O'Doherty, J. P., Koltzenburg, M., Wiech, K., Frackowiak, R., Friston, K. & Dolan, R. 2005 Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nature Neurosci.* **8**, 1234–1240. (doi:10.1038/nn1527)

Skolnick, A. J. & Davidson, R. J. 2002 Affective modulation of eyeblink startle with reward and threat. *Psychophysiology* **39**, 835–850.

Solomon, R. L. 1980 The opponent-process theory of acquired motivation. The cost of pleasure and the benefit of pain. *Am. Psychol.* **35**, 691–712. (doi:10.1037/0003-066X.35.8.691)

Strack, F. & Deutsch, R. 2004 Reflective and impulsive determinants of social behavior. *Person. Soc. Psychol. Rev.* **8**, 220–247. (doi:10.1207/s15327957pspr0803_1)

Tanimoto, H., Heisenberg, M. & Gerber, B. 2004 Event timing turns punishment to reward. *Nature* **430**, 983. (doi:10.1038/430983a)

Weiss, F. 2005 Neurobiology of craving, conditioned reward and relapse. *Curr. Opin. Pharm.* **5**, 9–19. (doi:10.1016/j.coph.2004.11.001)

Wise, R. A. 2004 Dopamine, learning and motivation. *Nat. Rev. Neurosci.* **5**, 483–494. (doi:10.1038/nrn1406)

Yarali, A., Niewalda, T., Chen, Y., Tanimoto, H., Duerrnagel, S. & Gerger, B. 2008 'Pain relief' learning in fruit flies. *Anim. Behav.* **76**, 1173–1185. (doi:10.1016/j.anbehav.2008.05.025)

Yarali, A., Krischke, M., Michels, B., Saumweber, T., Mueller, M. J. & Gerber, B. 2009 Genetic distortion of the balance between punishment and relief learning in *Drosophila*. *J. Neurogenet.* **23**, 235–247. (doi:10.1080/01677060802441372)