

HUMAN EVALUATIVE FLAVOR-FLAVOR CONDITIONING IS NOT SENSITIVE TO POST-ACQUISITION US-INFLATION

Frank BAEYENS, Wouter VANHOUCHE,
Geert CROMBEZ, and Paul EELEN
University of Leuven

Post-conditioning US-inflation was used to determine the nature of the US representation underlying human evaluative flavor-flavor conditioning. In Experiment 1, exposure to inflated USs did not result in the acquired CS+/CS-discrimination becoming more pronounced than in two control groups that received exposure to non-inflated USs or no further USs, suggesting that only the emotional-affective but not the sensory-descriptive US characteristics are encoded. In Experiment 2, two alternative accounts for this lack of a US-inflation effect were investigated, namely that the asymptote of conditioning was reached before US-inflation, or that the inflated US was not perceived as related to the acquisition US. Because neither of these possibilities was supported, it is concluded that this conditioning preparation evidences learning mediated by the referential system, whereby only the emotional-affective US characteristics are encoded.

Evaluative conditioning refers to the observation that the mere spatio-temporal co-occurrence of a neutral stimulus (conditioned stimulus, CS) with a positively or negatively valenced stimulus (unconditioned stimulus, US) may result in the originally neutral CS itself acquiring a valence congruent with the affective value of the US (Levey & Martin, 1987). For example, after a neutral picture of a human face has repeatedly been paired with a liked (or disliked) picture of another face, research participants typically demonstrate an acquired liking (or disliking) for the originally neutral face (e.g., Baeyens, Eelen, Crombez, & Van den Bergh, 1992). The evaluative flavor-flavor learning paradigm represents another conditioning preparation that has proven to be a very efficient tool for the study of human evaluative learning. In this differential conditioning paradigm, which was first developed by Zellner, Rozin, Aron, and Kulish (1983) and extended at our laboratory, research participants are exposed

Frank Baeyens, Wouter Vanhouche, Geert Crombez, and Paul Eelen, Department of Psychology, University of Leuven, Leuven, Belgium. Frank Baeyens and Geert Crombez are postdoctoral researchers for the Fund of Scientific Research (FWO).

We wish to express our feelings of gratitude towards Hilde Hendrickx, Jan De Houwer, Dirk Hermans, Deb Vansteenwegen, and Anneloes Vandenbroucke for their valuable contributions to this research and/or for their critical comments on an earlier version of this paper. Special thanks towards Stijn Collaerts for his assistance in the data collection of Experiment 2, and towards an anonymous reviewer for his critical comments on a previous version of this manuscript.

Correspondence concerning this article should be addressed to Frank Baeyens, Department of Psychology, University of Leuven, Tiensestraat 102, B-3000 Leuven. Electronic mail may be sent to Frank.Baeyens@psy.kuleuven.ac.be.

during acquisition to a series of flavored and colored water solutions containing the appropriate CS+/US and CS- contingencies. The results of several experiments showed that after an originally neutral fruit-flavor A (CS+) has been presented several times in simultaneous compound with the very bad-tasting substance Tween20 (Polysorbate 20; US), research participants develop a dislike for flavor A relative to a similar fruit-flavor B (CS-) which has been presented equally often in plain water (e.g., Baeyens, Eelen, Van den Bergh, & Crombez, 1990b). In the original paradigm, Zellner et al. (1983) obtained evidence for the acquisition of flavor likes due to contingent flavor-sugar pairings, even though the effect seems more fragile than the negative Tween-effect (e.g., Baeyens et al., 1990b).

We are beginning to obtain a quite detailed understanding of the conditions and the characteristics of this type of evaluative flavor learning. Available evidence first of all indicates that this effect is quite *robust*, in that (a) it involves rapid learning; (b) it survives the mere passage of time; (c) it is not under spontaneous contextual control of the color of the acquisition solutions; (d) it can be obtained with sequential rather than simultaneous presentations of CS and US flavors, albeit in a weaker form; and (e) it also allows for an observational rather than for a direct mode of presenting the crucial CS-US information (Baeyens et al., 1990b; Baeyens, Crombez, Hendrickx, & Eelen, 1995a; Baeyens, Vansteenwegen, De Houwer, & Crombez, 1996b; Baeyens, Hendrickx, Crombez, & Hermans, in press). Next, there is evidence within this paradigm that *qualitatively different classes of CSs* may be more or less, or even not at all appropriate to enter into association with the flavor US. Indeed, whereas reliable evaluative conditioning effects are obtained when the flavors of the liquids act as the critical CS+/CS- dimension and the colors function as an irrelevant distracter variable, no evaluative learning effect is obtained when the colors of the drinks are designated to be the CS+/CS- and the flavors represent the irrelevant distracter variable (Baeyens et al., 1990b; Baeyens et al., 1996b). Another characteristic of evaluative flavor-flavor learning is that it appears to be impossible to bring about *modulation* (or "occasion-setting") of the evaluative CS-US association. Thus, whereas research participants' evaluative behavior very easily comes under control of a simple CS-US association, we were not successful in establishing conditional evaluative discriminations, in which the validity of the CS-US association is dependent on the presence of another cue X (Baeyens, Crombez, De Houwer, & Eelen, 1996a; Baeyens et al., in press).

Two additional properties of evaluative flavor-flavor learning bring it in line with the functional characteristics of other evaluative conditioning preparations, and at the same justify to distinguish evaluative learning from Pavlovian expectancy-learning (Baeyens, Eelen, & Crombez, 1995b). The first issue bears on the role of explicit knowledge about the CS-US relation, c.q.

contingency-awareness. Expectancy learning, involving the acquisition of preparatory (defensive, appetitive, or orienting) conditioned responses, has repeatedly been shown to require awareness of the CS-US contingency (e.g., Hamm & Vaitl, 1996; for review see Dawson & Shell, 1987). Quite the opposite, evidence is accumulating from several paradigms that contingency awareness is *not* a necessary condition for evaluative conditioning (e.g., Baeyens, Eelen, & Van den Bergh, 1990a; De Houwer, Baeyens, & Eelen, 1994; De Houwer, Hendrickx, & Baeyens, in press; Hammerl & Grabitz, 1996; but see Baeyens & De Houwer, 1995; Baeyens, De Houwer, & Eelen, 1994; Davey, 1994; Shanks & Dickinson, 1990; Shanks & St. John, 1994 for critical comments and further discussion). In line with this, evaluative flavor-flavor learning has been demonstrated to proceed orthogonal to the conscious identification of the CS-US flavor contingencies. For example, in two studies (Baeyens et al., 1990b; Baeyens et al., 1996b) a double dissociation between evaluative learning and the acquisition of conscious knowledge was observed, in that with flavors functioning as CS+/CS-, none of the participants demonstrated awareness of the stimulus relations, but they evidenced clear conditioning; on the other hand, when colors acted as CS+/CS- a substantial number of participants was contingency-aware, but no evaluative conditioning was observed (see also Baeyens et al., 1996a; Baeyens et al., in press). The second important functional property of evaluative flavor learning pertains to the interrelated phenomena of *extinction* and *contingency-sensitivity*. In case of expectancy learning, unreinforced CS presentations after the acquisition phase (i.e., an extinction procedure) result in a decrease or complete disappearance of the conditioned preparatory responses. Similarly, the presentation of unreinforced CSs during the learning phase and hence lowering the CS-US contingency clearly attenuates the acquisition of conditioned responding. Descriptively, both phenomena may be thought of as the result of expectancy-disconfirmation (the expected US does not occur). At a process level, they result from an inhibitory component or inhibitory strength being conditioned to the CS (see Baeyens et al., 1995b, for a more detailed discussion). The main reason to conceive of evaluative learning as qualitatively different from expectancy-learning, is that unreinforced CS presentations, either during or after acquisition, do *not* seem to have an effect on this type of conditioning: evaluative conditioning is resistant to extinction (Baeyens, Crombez, Van den Bergh, & Eelen, 1988; Baeyens et al., 1995b; Baeyens, Eelen, Van den Bergh, & Crombez, 1989; Ledoux, Romanski, & Xagoraris, 1989; but again see Davey, 1994; Baeyens & De Houwer, 1995 for a critical discussion) and to CS-US contingency manipulations (Baeyens, Hermans, & Eelen, 1993). In line with this, in the evaluative flavor-flavor paradigm it was observed that postacquisition unreinforced CS presentations (extinction procedure) do not attenuate the acquired flavor dislike (Baeyens et al., 1995a; Baeyens et al.

1996a), and that an acquisition schedule containing an equal number of Tween20-reinforced and non-reinforced flavor trials results in an equally strong acquired flavor dislike as an acquisition schedule containing the same number of reinforced but no unreinforced learning trials (Baeyens et al., 1996a). Finally, there is some evidence that temporal contiguity may even overrule statistical contingency (Baeyens et al., 1996a; Baeyens et al., in press).

These findings justify to consider evaluative flavor-flavor learning as a prototypical example of a Pavlovian learning process mediated by what we named the Referential System (Baeyens et al., 1995b). According to this hypothesis, during acquisition an association is build up (consciously or unconsciously) between CS and US representations following a rudimentary Hebbian learning algorithm, the parameters of which only include temporal contiguity and stimulus salience: The synchronous activation of CS and US representations results in an increase of associative strength, whereas the activation of the representation of the CS alone (or of the US alone) is causally ineffective. Due to this associative link, activation of the representation of the CS flavor activates (consciously or unconsciously) a representation of the US, c.q. the bad-tasting substance Tween20, which is behaviorally expressed in a dislike for the CS flavor. Thus, it is hypothesized that this whole process can (but should not necessarily) take place in the absence of generating the expectation that the US is really going-to-occur-here-and-now, and without triggering defensive preparatory responses.

What this account leaves unanswered, however, and what the studies reported here tried to answer, is the *content of what is learned* (Rescorla, 1988). In the flavor-flavor paradigm, this comes down to the question how the Tween-US is encoded, or what is the content of the representation of the Tween-US which is activated by the flavor CS. Namely, several authors have argued that a US should be conceived of as a multi-property event involving several levels of representation, each of which can entertain separate associative links with a CS, and each of which can have different behavioral consequences. Wagner and Brandon (1989) proposed a distinction between *emotional-affective* and *sensory-descriptive* properties and corresponding levels of a US representation (see also Konorski, 1967; Mackintosh, 1983, for a related point of view). The emotional representation is argued to be more general than the sensory representation of a US, in that it may be shared by several, perceptually different discrete USs. For example, quinine and salt solutions have clearly distinctive perceptual characteristics and as such involve different sensory-descriptive representations, but probably share the same emotional-affective representation, in that both are bad-tasting, disliked stimuli. A privileged strategy to assess whether a CS is only associated with the emotional-affective representation or also with the sensory-descriptive representation of a US, involves post-acquisition modification of the value of the US ("US-revaluation",

or "US-inflation"; see Mackintosh, 1983; Rescorla, 1988; White & Davey, 1989). In this procedure, a CS-US relation is first presented. Next, the US is revaluated in the absence of the CS. For example, an originally positive US is made neutral or negative (revaluation), or an originally weakly negative US is presented at a high intensity, making it strongly negative (inflation). Finally, the CS is tested alone. The crucial question then becomes whether or not the US-revaluation has had any influence on the acquired valence of the CS. To the extent the CS *tracks* the changed valence of the US, it has to be concluded that the CS is associated with a sufficiently detailed representation of the specific sensory-descriptive characteristics of the US. Namely, the CS has to activate a representation of "this particular (dis)liked US with these particular sensory characteristics", not only the representation of "something (dis)liked", in order for the newly acquired knowledge - that "the valence of this particular US with these particular sensory characteristics has changed" - to be able to affect the acquired valence of the CS. If, on the other hand, the CS *does not track* the changed value of the US, this suggests that the CS is only associated with the emotional-affective level of the US-representation. Namely, if subjects associate a CS only with the general affective-properties of the US ("something (dis)liked"), and not with any sensory attributes of this event, and if they later learn that "the evaluative value of this particular US has changed", there is indeed no way to combine these two different pieces of knowledge, so that the acquired valence of the CS cannot be affected (Mackintosh, 1983). We already followed this rationale to determine the content of learning in the evaluative picture-picture paradigm (Baeyens et al., 1992). We first exposed participants to pairings of pictures of neutral faces (CS) and either strongly liked or disliked faces (US). On a within-subject base, a US-revaluation technique was next applied to half of the USs (e.g., an originally liked US was made neutral), whereas the valence of the other USs was left unchanged. We observed that the post-conditioning revaluation of the USs *did* affect the acquired valence of the CSs previously paired with them, even in subjects who were unaware of the original CS-US contingency. Thus, we concluded that the conditioned valence shifts of the CS pictures were based on an association with the representation of the US including not only the emotional-affective, but also the specific sensory-descriptive characteristics of the (dis)liked US face. In the first experiment reported here, a similar approach was followed to determine the content of learning in the evaluative flavor-flavor learning paradigm.

Experiment 1

Experiment 1 contained three conditions: the crucial US-Inflation group, the *no_US-Control* group, and the *US-Control* group (see Table 1, top). All groups were treated identically throughout all phases of the experiment, except for the stimuli that were presented during the US-inflation phase. Participants were first exposed to a standard differential CS+/CS- flavor conditioning schedule using fruit-flavors as CSs and a moderate concentration of Tween20 as the disliked US. After a first assessment of acquired differential evaluative responding (Post-Acquisition test), participants in the *US-Inflation* group were exposed to a series of "inflated" US trials, consisting of a very strong concentration of Tween20 presented without the CS flavors. To the extent that evaluative flavor-flavor learning involves an encoding of the sensory-descriptive characteristics of the Tween US, this US-inflation procedure was expected to result in the acquired dislike for the CS+ flavor becoming stronger and, consequently, in the evaluative CS+/CS- discrimination becoming more pronounced at the second, Post-Inflation assessment phase. If on the other hand only the emotional-affective characteristics of the Tween US are encoded, the acquired dislike for the CS+ flavor (and the CS+/CS- discrimination) was expected not to change: Learning that "this particular flavor/taste stimulus (c.q. Tween20) has become very dislikeable" cannot affect the knowledge that "the CS+ flavor is associated with something disliked".

The two control groups (see Table 1, top) were included to allow us to infer that the eventual post-inflation enhancement of discriminative responding towards CS+ and CS- flavors in the *US-Inflation* group was actually due to the inflation of the US and not to any other aspect of the procedure. The *no_US-Control* group received a series of plain water trials (instead of inflated US trials) after the acquisition phase. This group was included to assess the influence of repeated testing on the acquired CS+/CS- differentiation, c.q. to obtain a baseline of the course of discriminative evaluative responding from the Post-Acquisition to the Post-Inflation test phases. The *US-Control* group received an equal number of post-acquisition US-only trials as the *US-Inflation* group, but Tween20 was presented at the same moderate concentration as used during acquisition rather than at the very strong concentration used in the *US-Inflation* group. This control group was added to ensure that an eventually observed US-inflation effect in the *US-Inflation* group was indeed due to the *inflated* character of the post-acquisition US presentations and not to the mere fact that participants were exposed to additional post-acquisition USs. To the extent, however, that no differences would show up between the *US-Control* and the *no_US-Control*, it was planned to merge the data from both conditions into one single Control for the *US-Inflation* manipulation.

The experiment contained one further within-subject manipulation (see

again Table 1, top). After the Post-Inflation test, participants in all three groups were exposed to an extinction schedule, consisting of a series of unreinforced CS+ and CS- flavor presentations, which was followed by a final Post-Extinction evaluation of the flavors. The most important aim of the extinction manipulation was to replicate the previously reported finding that post-acquisition non-reinforced CS presentations do not affect acquired flavor dislikes. Hence, we expected that the Post-Extinction ratings of both Control groups would demonstrate that evaluative flavor conditioning is resistant to extinction. The Post-Extinction ratings of the US-Inflation group would additionally inform us whether the eventual US-inflation induced enhancement of discriminative evaluative responding would survive extinction treatment, or whether discriminative responding would rather return to the pre-US-inflation level.

Method

General Lay-Out of Experiment 1

In the US-inflation condition as well as in the US-Control and the no_US-Control conditions, the experiment consisted of six consecutive phases (see Table 1, top): Acquisition, Post-Acquisition test, US-Inflation, Post-Inflation test, Extinction, and finally the Post-Extinction test. During Acquisition, participants were exposed to nine Tween-reinforced CS+ and to nine unreinforced CS- trials (both CS+ and CS- flavors each presented three times in each of three different colors). Within each condition, for half of the participants flavor A functioned as CS+ and flavor B as CS-, whereas the reverse was true for the other half. The Post-Acquisition test consisted of an evaluation of four CS+ and of four CS- liquids (both CS+ and CS- flavors each presented once in each of the three different acquisition colors and once colorless). During the US-inflation phase, the US-Inflation group consumed three inflated US-only liquids (high dose of Tween20), the US-Control condition three normal US-only liquids (moderate acquisition dose of Tween20), whereas the no_US-Control was offered three plain water trials. The Post-Inflation test was identical to the Post-Acquisition test. During extinction, research participants were exposed to four unreinforced CS+ and CS- trials (both CS+ and CS- flavors each presented once in each of the three different colors and once colorless). The final Post-Extinction test was identical to both previous test phases.

Table 1. *Design of Experiments 1 and 2*

Group	Acquisition	Post-Acq. Test	Inflation	Post-Infl. Test	Extinction	Post-Ext. Test
EXPERIMENT 1						
US-Inflation						
US-Control	9 CS+ _{Tween_2ml/L} 9 CS-	4 CS+ 4 CS-	3 Tween_20ml/L 3 Tween_2ml/L	4 CS+ 4 CS-	4 CS+ 4 CS-	4 CS+ 4 CS-
No_US Control			3 Water			
EXPERIMENT 2						
US-Inflation						
Ass.-US-Inflation	9 CS+ _{Tween_0.45ml/L} 9 CS-	4 CS+ 4 CS	4 Tween_10ml/L 4 Tween_0.45ml/L → Tween_10ml/L	4 CS+ 4 CS-		
US-Control			4 Tween_0.45ml/L			

Research Participants

A total of 48 undergraduate psychology students (14 males, 34 females) participated as a partial fulfilment of course requirements. None of them had any previous experience with a flavor-conditioning study. Twenty-four students were randomly selected to participate in the US-Inflation condition, whereas 12 randomly selected participants were run in the US-Control and the no-US-control conditions. Participants were tested in groups of six, and test duration was about 45 minutes. No information was provided with respect to the true purpose of the experiment.

Stimuli

Two colorless artificial concentrate flavors, Lemon and Apricot, were used as CS+ or CS-. In an independent rating group ($N = 14$) which was run before this study started, the average evaluative rating for Lemon was +6.43 on a +100/-100 scale, whereas Apricot was rated +5.71. The flavors were provided by Universal Flavors (Belgium). Three commercial flavorless food colorings (Vahiné), Yellow (E 102), Red (E 122), and Green (E 102 and E 131) were added to the liquids as a distracting color context. All stimuli were presented in 5 ml tap water at room temperature. Lemon flavor was presented at a concentration of 0.3 ml/L water, Apricot flavor at a concentration of 1 ml/L. A concentration of 1 ml/L water was used for the Yellow and the Green food colorants, and 1.5 ml/L for Red. A concentration of 2 ml/L of Tween20 (*Polysorbate 20*) functioned as the disliked US used for acquisition and for post-acquisition US-only presentations in the US-Control group, whereas a 10 times stronger concentration of 20 ml/L of Tween20 was used for the inflation of the US in the US-Inflation group. A group of 10 judges (staff members) independently evaluated the 20 ml/L dose of Tween20 as much more aversive than the 2 ml/L dose.

Apparatus

The experiment took place in a lab containing seven identical separated test booths arranged in a semi-circle, such that visual contact between participants was precluded, but each of them could clearly see and hear the experimenter. Each test booth contained a table and a chair. The experimenter was seated behind a table at the centre of the semi-circle. The 18 transparent plastic 30-ml cups containing the acquisition solutions were presented in a wooden cup-holder, and were arranged in the appropriate order of presentation for each participant. A second wooden cup-holder contained first the eight cups for the Post-Acquisition test, next the three cups for the US-inflation phase, followed

by eight cups for the Post-Inflation test, eight cups for the extinction phase, and finally eight cups for the Post-Extinction test. All cups on the second tray were numbered 1 to 35, the numbers being written in black on top of each cup. Within easy reach of the participants, the tables contained dishes filled with small pieces of white bread, to be eaten between each conditioning or test trial.

Procedure

Acquisition. Upon entering the laboratory, participants were told that we were interested in "some variables that may play a role in humans' perception of flavors". The experimenter said that they should first attentively look at and consume a series of drinks, without having to give any explicit responses. After this, they would be presented with a second series of drinks, about which some questions should be answered. Participants were also informed that some of the drinks might have a bad taste, but that these drinks were completely harmless. Finally, it was stressed that they should follow the instructions as precisely as possible, such that each of them would be presented in the same way and for the same amount of time with the experimental stimuli. Next participants consumed the nine reinforced CS+ and the nine unreinforced CS- cups on the tray. The cups were placed in a random sequence, with the restriction that no more than two consecutive trials were allowed to be a CS+ (or CS-). A different random sequence was used for each participant. Each consumption trial lasted for 40 s and consisted of the following five elements. During the 5 first s of each trial, participants had to take a cup and look at it; during the next 5 s they moved around the cup under their nose and smelled; during the next 5 s, they took the liquid into their mouth and swirled it around in their mouth until the experimenter signalled them to swallow; the next 5 s period was a waiting period, after which Ss took a piece of bread and ate it; the remainder of the final 20 s of each trial consisted of a waiting phase. The eating of the piece of bread was introduced to avoid after-tastes and the mixing up of flavors of consecutive trials as much as possible. Participants were guided through each acquisition trial by means of the experimenter's verbal instructions: "next cup and looking", "smelling", "in mouth and tasting", "swallow", and finally "piece of bread".

Post-acquisition test. The participants received a response booklet containing a rating scale for each of the test stimuli, and were presented with the second tray containing the liquids for this and all following phases. They were instructed that they should consume each of the drinks in the same way as the first series of drinks, the only difference being that they should now simply indicate how much they liked or disliked each drink. The rating scale was a line 222 mm long, labelled *very disliked/very bad* at the left, *neutral* in the middle, and *very liked/very good* at the right. The line was subdivided into 21 categories,

labelled by numbers above the scale (-100/-90/-80/.../0/.../+80/+90/+100). Participants were asked to mark the category indicating their global, spontaneous evaluation of the drinks. As during the conditioning phase, they rated a solution every 40 s on the experimenter's signal. Hence during the first 5 s of each trial, participants took a cup and looked at it; during the next 5 s they smelled; during the next 5 s they took the solution into their mouths and tasted it; after 5 s of waiting, they rated the solution and next ate a piece of bread; the remainder of the last 20 s was waiting time. The order of presentation of the eight test stimuli of the Post-Acquisition test (four CS+, four CS- flavors) was randomized, with the restriction that no more than two consecutive trials were allowed to be a CS+ (or CS-). A different random sequence was used for each participant.

US-inflation. Research participants were by no means informed about the crucial US-inflation manipulation nor about the fact that the trial following the last Post-Acquisition test trial represented a transition from testing to US-inflation. Depending on the specific condition, participants consumed and evaluated the three liquids containing the inflated USs (US-Inflation), the standard USs (US-Control), or plain water (no_US-Control), and this in exactly the same way as they consumed and evaluated the Post-Acquisition test liquids. The reason why participants were required to rate the three inflation stimuli was mainly to create smooth transitions from testing to the inflation manipulation to subsequent testing, making the crucial experimental manipulation less transparent and less sensitive to experimental demand. Even though it might be tempting to believe that these ratings can also be used to evaluate whether the inflated Tween concentration indeed was perceived as more aversive than the acquisition dose, we expected that these data would probably *not* reflect this difference, due to a high likelihood of a differential anchoring of evaluative, stimulus, and response continua in the three groups. Namely, using the standard -100/+100 rating scale the extremes of which are anchored to the terms "very disliked/very bad" and "very liked/very good", the 2ml/L acquisition dose normally already obtains extreme (near-minimal) ratings, and the manipulation of the Tween concentration was between-group, such that participants in the US-control group could not compare the negativity of the 2 ml/L dose with the 20 ml/L dose and adjust (heighten) their ratings accordingly.

Post-inflation test. Again, no mention was made about the transition from US-inflation to testing. The eight-trial Post-Inflation test was identical to the Post-Acquisition test, except for the fact that a different, individually determined randomized order of stimulus presentation was used.

Extinction. It was said to the participants that, just like for the stimuli of the first tray (acquisition tray), no evaluations were required for the subsequent

eight trials. Hence, four CS+ and four CS- liquids, all unreinforced, were consumed (but not rated) in an individually determined random sequence.

Post-extinction test. The final rating of the four CS+ and four CS- ratings proceeded in the same way as the two previous test phases.

Results

US-Inflation Phase Ratings

The average ratings of the inflated US (20 ml/L), the acquisition US (2 ml/L), and of the plain water stimuli were -90.42, -94.44, and 6.67, resp. A between-subject ANOVA on these data revealed a significant effect of the Group variable (US-Inflation/US-Control/no_US-Control), $F(2, 45) = 192.76$; $p < .0001$. Tukey a posteriori pairwise comparisons demonstrated a reliable ($p < .001$) difference between the plain water ratings and the ratings of both the inflated and the acquisition USs, which did not differ from each other, however. Hence, as expected, participants in the US-Inflation group as well as in the US-Control group gave the inflated or the non-inflated Tween stimuli extreme, near-minimal ratings. Therefore, with respect to the successfulness of the inflation manipulation, we had to rely on the pre-experimentally obtained opinion of the 10 staff members, who were exposed to both concentrations and independently but unanimously evaluated the 20 ml/L dose of Tween20 as much more aversive than the 2 ml/L dose (see also results of Experiment 2 for additional empirical arguments justifying this decision).

CS Ratings in US-Control Versus in No_US-Control Groups

A first analysis of the CS ratings concentrated on the question whether there were any reliable differences between the two different control groups. To that purpose, the Post-Acquisition, Post-Inflation, and Post-Extinction CS+ and CS- evaluations of the US-Control and of the no_US-Control groups were analyzed by means of a $2 \times 2 \times 2 \times 2 \times 4 \times 3$ [CS+_Flavor (Apricot=CS+/Lemon=CS+) $\times$ Group (US-Control/no_US-Control) $\times$ Conditioning (CS+/CS-) $\times$ Color (Red/Yellow/Green/Colorless) $\times$ Phase (Post-Acquisition/Post-Inflation/Post-Extinction)] ANOVA with repeated measures on the last three variables. An alpha level of .05 was used for all statistical tests throughout this and the subsequent experiment, and Greenhouse-Geiser correction of degrees of freedom was applied whenever appropriate. Neither the main effect of Group, nor the more crucial Group $\times$ Conditioning, nor the Group $\times$ Conditioning $\times$ Phase interactions were reliable (all F 's < 1). Furthermore, none of the other

interactions involving the Group variable reached significance. Therefore, it was decided to merge the data from the US-Control and the no_US-Control into one single Control ($N = 24$) for the US-Inflation group ($N = 24$).

CS Ratings in Control versus Inflation Groups

The mean Post-Acquisition, Post-Inflation, and Post-Extinction CS+ and CS- flavor ratings of US-Inflation and Control groups are presented in Figure 1, averaged over the different levels of the test Color variable. As predicted, in the Control group the CS+ ratings were more negative than the CS- ratings. This effect appeared to remain quite stable throughout the three different test moments, suggesting no effect of the postconditioning US-only or water presentations, nor an effect of the extinction manipulation. Contrary to what would be expected if evaluative flavor learning would be sensitive to US-inflation, the same appeared to be true in the US-Inflation condition. The reliability of these and of some other observations was confirmed by the results of a $2 \times 2 \times 2 \times 2 \times 4 \times 3$ [CS+_Flavor (Apricot=CS+/Lemon=CS-) $\times$ Inflation (Control/US-Inflation) $\times$ Conditioning (CS+/CS-) $\times$ Color (Red/Yellow/Green/Colorless) $\times$ Phase (Post-Acquisition/Post-Inflation/Post-Extinction)] ANOVA with repeated measures on the last three variables. The main effects of Phase, $F(2, 88) = 10.32$; $p < .0001$; $Ndf = (1.86, 81.66)$, $p < .0002$; $MSe = 1548.23$, of Color, $F(3, 132) = 3.76$; $p < .02$; $Ndf = (2.53, 111.36)$, $p < .02$; $MSe = 794.61$, and, theoretically more important, of Conditioning, $F(1, 44) = 7.79$; $p < .008$; $MSe = 4258.41$, were significant. However, there was not the slightest evidence for a Conditioning $\times$ Phase interaction, nor for the crucial interaction between Conditioning $\times$ Phase $\times$ Inflation (both F 's < 1). None of the other main effects or interactions reached significance.

The main effect of Phase was due to the fact that the overall rating of the flavors became gradually more negative as testing proceeded (means at Post-Acquisition, Post-Inflation, and Post-Extinction: -2.66, -8.23, and -15.52, resp.; linear trend, $F(1, 44) = 16.51$; $p < .0002$; $MSe = 1925.05$), a phenomenon which most probably reflects simple satiation and fatigue. The Color main effect involved different ratings for red (mean = -5.66), colorless (mean = -7.12), yellow (mean = -9.38), and green drinks (mean = -13.06); however, only the difference between red and green drinks was reliable (*Tukey* a posteriori pairwise comparisons). This result most probably reflects simple baseline differences. Finally, the Conditioning main effect implied overall more negative ratings of the CS+ flavor (mean = -14.17) than of the CS- flavor (mean = -3.44), a fact which guarantees that the standard evaluative flavor-flavor learning effect was replicated in this study. Even though neither the Conditioning $\times$ Phase, nor the Conditioning $\times$ Phase $\times$ Inflation interaction reached significance, based on our *a priori* hypotheses we next calculated

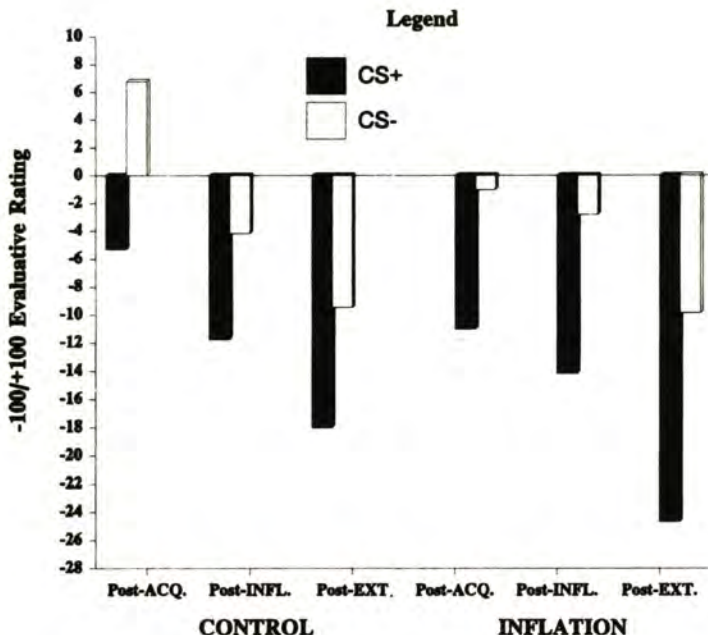


Figure 1. Mean Post-Acquisition, Post-Inflation, and Post-Extinction CS+ and CS- flavor ratings of the US-Inflation and Control groups (data of US-Control and no_US-Control are merged), averaged over the different levels of the test Color variable.

several contrasts that provide a more detailed picture of the results obtained at the different phases of the experiment. The results of these additional analyses are reported in the next sections.

Basic evaluative flavor-flavor conditioning effect. At the first (Post-Acquisition) test a reliable Conditioning effect was observed, $F(1, 44) = 5.71$; $p < .03$; $MSe = 2048.49$ (mean CS+ = -8.18, mean CS- = 2.86) which was not modified by an Inflation X Conditioning interaction, $F < 1$ (Control mean CS+ = -5.31, CS- = 6.77; Inflation mean CS+ = -11.04, mean CS- = -1.04).

Post-acquisition US-inflation sensitivity. In the Control condition, the extent of the evaluative CS+/CS- differentiation did not change from Post-Acquisition to Post-Inflation; the same was true in the US-Inflation group, and both groups did not differ in this respect (all three interactions F 's < 1). In line with this, the Conditioning effect remained reliable at the Post-Inflation test, $F(1, 44) = 4.67$; $p < .04$; $MSe = 1845.88$ (mean CS+ = -12.97, mean CS- = -3.49),

but again was not modified by an Inflation X Conditioning interaction, $F < 1$ (Control mean CS+ = -11.77, CS- = -4.17; Inflation mean CS+ = -14.17, mean CS- = -2.81). In sum, there was not the slightest evidence for the acquired evaluative flavor dislike to be sensitive to US-inflation.

Extinction sensitivity. In the same vein, the subsequent extinction trials did not have any demonstrable effect on the acquired CS+/CS- discrimination. Over Control and US-Inflation groups, and in both conditions separately, the CS+/CS- discrimination did not reliably change from Post-Inflation to Post-Extinction (all F 's < 1). The same was true when we looked for a shift from the first Post-Acquisition to the final Post-Extinction test. Thus, after eight unreinforced CS+ and CS- test presentations (four times during Post-Acquisition test and four times during Post-Inflation test), and despite an additional four unreinforced exposures during the formal extinction phase, a highly reliable Conditioning effect was still present at the final Post-Extinction test, $F(1, 44) = 8.18$; $p < .007$; $MSe = 1597.16$ (mean CS+ = -21.35, mean CS- = -9.69). As in the two previous tests, this effect was unmodified by an Inflation X Conditioning interaction, $F < 1$ (Control mean CS+ = -18.02, CS- = -9.48; Inflation mean CS+ = -24.69, mean CS- = -9.90).

Discussion

In Experiment 1, during a first phase participants were exposed to a differential evaluative flavor-flavor conditioning schedule, which resulted in reliable evaluative flavor learning. The subsequent exposure to a series of inflated USs (US-Inflation group) did not result in the evaluative CS+/CS- discrimination becoming more pronounced at the Post-Inflation test. The same unmodified but still reliable discriminative responding was obtained in two control groups that received exposure to non-inflated USs (US-Control) or to plain water trials (no_US-Control) at the moment of the US-inflation manipulation. A subsequent extinction manipulation did not affect the acquired evaluative discrimination in either of the groups. This aspect of the results replicates what we argued to be the hallmark of a learning process mediated by the Merely Referential system (see above): resistance to extinction, or more generally, ineffectivity of unreinforced CS presentations. The fact that no evidence was obtained for a US-inflation effect deserves some further comments, however.

As argued in the introduction, a straightforward conclusion would be that only the emotional-affective and not the specific sensory-descriptive characteristics of the Tween US are encoded in evaluative flavor learning: the CS+ flavor is associated with and associatively activates the representation of "something disliked", not of "this particular disliked taste/flavor with these

particular sensory characteristics". Before accepting this conclusion, several alternative accounts of the failure to observe a US-inflation effect have to be considered, however.

First, the absence of US-inflation is concluded from a nonsignificant effect for the critical Conditioning X Phase X Inflation interaction, and statistically represents an acceptance of the null hypothesis. Consequently, a critical issue is the sensitivity or *statistical power* of the critical test. Following a corrected procedure of Cohen's (1988) treatment of power calculations in ANOVA (see Hendrickx, De Houwer, Baeyens, Eelen, & Van Avermaet, 1997 (*Appendix*)), we therefore calculated the power ($1-\beta$) for the critical *F*-test, and arrived at values of .85, 1, and 1 for a small ($f = .10$), a medium ($f = .25$), or a large ($f = .40$) effect size, resp. Hence, even if the effect size of an eventual US-inflation effect would be small, power exceeds Cohen's recommended criterion of .80. Moreover, we estimated the effect size of the only available demonstration of US-revaluation in evaluative learning, c.q. the revaluation effect obtained in the previously discussed (Baeyens et al., 1992) picture-picture paradigm, and arrived at a value falling within the range of Cohen's operational definition of a *large* effect, namely, $f = .41$. Taken together, this power analysis suggests a more than satisfying ($1-\beta$) for the critical test in Experiment 1, such that the acceptance of the null hypothesis and the conclusion that no US-inflation was obtained seems justified, at least on statistical grounds.

Next, the question might be raised whether, as a matter of fact, there exists any evidence within this paradigm for a logical precondition of US-inflation: the sensitivity of evaluative flavor-flavor conditioning to the strength of the US. Namely, if post-conditioning inflation of the US is to have any effect on differential responding to the CSs, it should at least be possible to demonstrate that groups initially conditioned with a weak or with an inflated US reliably differ in the extent of differential (CS+/CS-) responding at the end of the acquisition phase. Even though the experiment reported here does not provide a direct comparison of two such groups, in the Baeyens et al. (1995a) flavor-flavor parameter study, post-acquisition differential responding to the test flavors (CS+/CS-) was indeed shown to be significantly affected by the strength of the evaluative differentiation of the acquisition compounds (CS+_US/CS-_noUS). This clearly suggests that, at least in principle, the prerequisite of evaluative flavor learning showing sensitivity to the strength of the US, is fulfilled in the paradigm used in Experiment 1.

Two other explanations for the absence of a US-inflation effect cannot so easily be dismissed based on the available evidence, however. The first, related to but distinguishable from the point discussed above, bears on the possibility of a ceiling-effect: maybe, the asymptote of conditioned differential evaluative responding was already reached after the acquisition phase of Experiment 1, leaving no room for the subsequent US-inflation to further enhance

discriminative responding. This critique may be supported by the results of the same parameter study as mentioned above, in which augmenting the number of acquisition trials from 6 to 12 likewise did not increase the evaluative flavor conditioning effect (Baeyens et al., 1995a). The second possibility is that participants somehow failed to "link" the inflated US to the acquisition US, i.e. to perceive the inflated US as a stronger and more aversive version of the acquisition US. Namely, if the 20 ml/L dose of Tween20 would be perceived as a qualitatively different stimulus bearing no relation whatsoever to the 2 ml/L dose, no US-inflation effect could reasonably be expected even in case the CS had been associated with a representation of the US comprising its specific sensory characteristics. The purpose of Experiment 2 was to assess the validity of these last two alternative interpretations of the absence of a US-inflation effect in Experiment 1.

Experiment 2

Theoretically, there are two options to deal with the problem of a potential ceiling-effect hindering the induction of a US-inflation effect. A first is to reduce CS-US associative strength by decreasing the number of acquisition trials, a second to decrease US intensity during acquisition. Rather than trying to create a CS entertaining a weak association with an already strong US, entailing the risk that no CS-US association at all develops, we tried to establish a CS having a strong association with a weak US. We expected that this conditioning manipulation would result in the Post-Acquisition ratings showing only a very weak or even no differential conditioning effect at all. However, to the extent that flavor-flavor learning would entail an encoding of the sensory characteristics of the US, the presence of a strong associative link between CS and US representations would allow the subsequent US-inflation manipulation to show up in reliable Post-Inflation discriminative evaluative responding towards the CS flavors (for a successful demonstration of this strategy in human expectancy learning, see White & Davey, 1989). Therefore, during the Acquisition phase of Experiment 2 we retained the number of learning trials used in Experiment 1, but lowered the acquisition dose of Tween20 to a very weak concentration.

After acquisition and a Post-Acquisition test, participants in the *US-Inflation* and in the *US-Control* group were treated as in Experiment 1, and thus received a series of US-only presentations at either the inflated or the low acquisition concentration (see Table 1, bottom). A third condition, the *Associative-US-Inflation* group was added in an attempt to deal with the problem of ensuring a "linking" of the weak acquisition US and the inflated US. At the Inflation phase, these participants were exposed to sequential pairings of the weak

acquisition US immediately followed by the inflated US, a strategy making it probably more likely that the two stimuli are perceived as qualitatively related stimuli presented at a low versus at a high concentration. Moreover, given the fact that evaluative flavor-flavor conditioning has also been demonstrated using sequential presentations of CS and US (see introduction), this Associative-US-Inflation technique might also be conceptualized as an associative route to modify the valence of the weak US into negative direction (i.e., to reevaluate it). Whether by means of genuine US-inflation or by means of US-revaluation, this condition was expected to optimize chances to obtain post-acquisition modification of the value of the US.

A final departure from Experiment 1 was that this study contained no Extinction phase, and consequently no third, Post-Extinction measurement phase.

Method

Research Participants

A total of 38 undergraduate psychology students (11 males, 27 females) participated as a partial fulfilment of course requirements. None of them had any previous experience with a flavor-conditioning study. Twelve students were randomly selected to participate in the US-Control condition, whereas 13 randomly selected participants were run in the Associative-US-Inflation and the US-Inflation conditions. Participants were tested in groups of six or seven, and test duration was about 35 minutes. No information was provided with respect to the true purpose of the experiment.

Stimuli

All CS flavor and color stimuli were identical to those used in Experiment 1. The concentrations of Tween20 used in this study were arrived at after substantial piloting. A concentration of 0.45 ml/L of Tween20 was selected as the Acquisition US, and as US at the Inflation phase in the US-Control and Associative-US-inflation groups (first member of pair). A panel of independent judges (10 staff members) evaluated drinks containing a 0.45 ml/L of Tween20 as still clearly containing the taste characteristics of Tween20, without however evoking a strong reaction of dislike. A concentration of 10 ml/L Tween20 was used as the inflated US in US-inflation and Associative-US-Inflation groups (second cup). The inflated US concentration was decreased from 20 ml/L (Experiment 1) to 10 ml/L, because the panel of judges unanimously rated the 10 ml/L concentration of Tween20 as equally strong and bad as the 20 ml/L concentration.

Procedure

Acquisition and Post-Acquisition testing proceeded in exactly the same way as in Experiment 1. At the US-Inflation phase, participants in the US-Control and the US-Inflation groups consumed four liquids (rather than three in Experiment 1) containing 0.45 ml/L or 10 ml/L of Tween20. In the Associative-US-Inflation condition, each of the four inflation trials lasted for 60 s and proceeded as follows. Immediately after looking at, smelling, tasting, and swallowing the liquid in the first cup (15 s), participants took the second cup of a pair (5 s), visually inspected it, smelled the liquid, tasted and swallowed it (15 s). After another 5 s of waiting, they evaluated both stimuli of a pair, and took a piece of bread, and waited for the remainder of the final 20 s of a trial. The final Post-Inflation test proceeded in exactly the same way as in Experiment 1.

Results

US-Inflation Phase Ratings

Like in Experiment 1, it was problematic to rely on a between-group comparison of the US ratings in the US-Control and US-Inflation groups in order to assess the successfulness of the manipulation of the Inflation independent variable, because again the US-Control group was never exposed to the much stronger inflated US-concentration. Consequently, the same problem of a differential anchoring of the subjective evaluation to the rating scale showed up here. The mean evaluation of the 0.45 ml/L concentration (US-Control) was nearly as negative (mean = -79.50) as the evaluation of the 10 ml/L concentration (US-Inflation group) (mean = -82.68) (difference: $F < 1$). However, because participants in the Associative-US-inflation group were exposed to *both* concentrations, the interpretation of their ratings of the weak and the strong US was less problematic. In this group, the 0.45 ml/L concentration of Tween20 was indeed rated only weakly negative (mean = -36.16) and definitely much less negative than the strong 10 ml/L concentration (mean = -83.10) (difference: $F(1, 11) = 116.99$; $p < 0.000001$; $MSe = 121.67$). This both confirms the successfulness of the manipulation of the Inflation independent variable, and testimonies to the validity of our interpretation of the absence of a between-group difference between US-Control and US-Inflation conditions in Experiment 2 (as well as in Experiment 1).

CS Ratings in US-Control, US-Inflation, and Associative-US-Inflation Groups

The mean Post-Acquisition and Post-Inflation CS+ and CS- flavor ratings of the US-Control, US-Inflation, and Associative-US-Inflation groups are presented in Figure 2, averaged over the different levels of the test Color variable. In general, and with the possible exception of the Post-Acquisition ratings of the US-Control group, little or no evaluative differentiation was observed between CS+ and CS- flavors, and this in all groups and both at Post-Acquisition and at Post-Inflation. Thus, in line with what was expected when a very weak US is used during acquisition, there appeared to be no clear differential conditioning after the Acquisition phase. Neither in the US-Inflation nor in the Associative-US-Inflation conditions was there any apparent tendency for an evaluative CS+/CS- discrimination to show up at the Post-Inflation test.

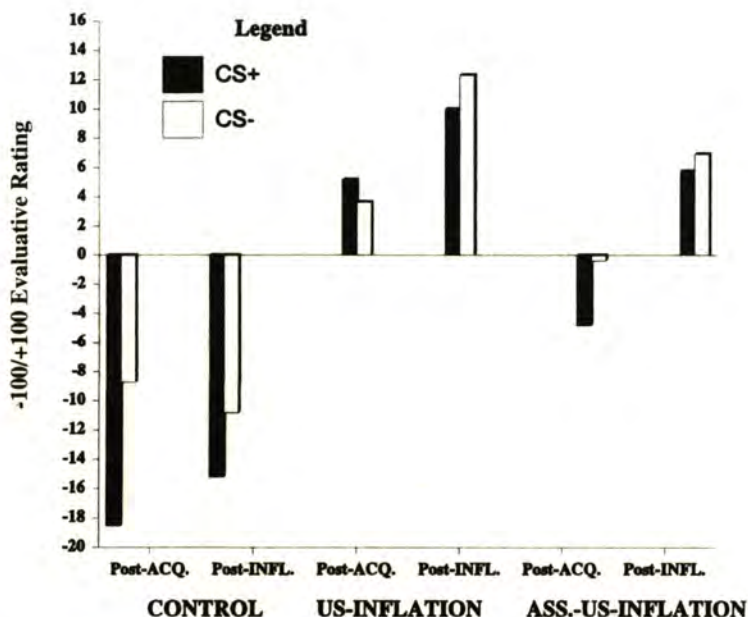


Figure 2. Mean Post-Acquisition and Post-Inflation CS+ and CS- flavor ratings of the US-Control, US-Inflation, and Associative-US-Inflation groups, averaged over the different levels of the test Color variable.

In line with this, a 2 X 3 X 2 X 4 X 2 [CS+ _Flavor (Apricot=CS+/Lemon=CS+) X Inflation (US-control/US-Inflation/Associative-US-Inflation) X Conditioning (CS+/CS-) X Color (Red/Yellow/Green/Colorless) X Phase (Post-Acquisition/Post-Inflation)] ANOVA with repeated measures on the last three variables failed to provide any evidence for a main effect of Conditioning, $F(1, 32) = 1.66$; $p > .20$; $MSe = 1635.12$, nor for a Conditioning X Phase, nor for the most crucial Conditioning X Phase X Inflation interactions, both F 's < 1 . None of the other main effects or interactions were significant. A set of contrasts testing for a Conditioning effect at Post-Acquisition or at Post-Inflation, or for a Phase X Conditioning interaction, revealed no single significant effect, nor in the US-Control, nor in the US-Inflation, nor in the Associative-US-Inflation group.

Finally, power analysis again confirmed that the absence of a significant Conditioning X Phase X Inflation interaction was most likely not a result of insufficient power of the test, in that, except for a small effect size, $1-\beta$ again proved to be satisfying ($f = .55, .99$, and 1 for a small, a medium, or a large effect, resp.).

Discussion

In Experiment 2, two alternative accounts for the lack of a US-inflation effect in Experiment 1 were dealt with. The first possibility was that the asymptote of conditioned differential evaluative responding was already reached before the US-inflation manipulation, the second that the inflated US was not perceived as being related to the acquisition US. In order to avoid a ceiling-effect, a relatively high number of acquisition trials with a very weak US was used. As expected, the Post-Acquisition ratings showed that this weak US was not strong enough to affect the valence of the flavor CSs. However, given the fact that a number of acquisition trials was used that standardly results in a sufficiently strong CS-US associative link, and on condition that flavor-flavor learning would entail an encoding of the sensory characteristics of the US, the subsequent inflation of the US should have caused the hitherto behaviorally silent CS-US association to be expressed in reliable discriminative evaluative responding towards the CS flavors. This was not observed, neither by using the standard US-inflation procedure, nor by using the Associative-US-Inflation technique, which was designed to maximize chances that the strong and the weak US would be perceived as related stimuli. Like in Experiment 1, power analysis confirmed that the absence of a reliable effect should not be attributed to insufficient sensitivity of the statistical test. Therefore, available evidence makes it plausible that this evaluative flavor-flavor conditioning preparation is genuinely not sensitive to post-acquisition US-inflation. This

makes it probable that only the emotional-affective and not the specific sensory-descriptive characteristics of the Tween US are encoded in this type of evaluative flavor learning.

General Discussion

In previous papers (e.g., Baeyens et al., 1995b), we noticed that the distinction between an Expectancy System and a Referential System does not presuppose any assumption concerning the nature of the representation of the US. That is, from an *a priori* point of view, it seems plausible that a CS may entertain an expectancy or a referential relation with only the affective, or also with the sensory-descriptive aspects of a US representation. Given the findings reported in this paper, we now have empirical evidence that it is in line with the supposed orthogonality of the Expectancy/Referential-system and the Affective/Sensory-levels-of-US-representation dichotomies. Namely, whereas the animal and human Pavlovian conditioning literature provides plenty of examples of (extinction-sensitive) expectancy learning preparations which are either or not sensitive to US-revaluation/inflation, up to now no conditioning preparation had been identified which showed both resistance-to-extinction (and therefore should be categorized as being based on referential learning), and no sensitivity to US-revaluation/inflation. The results of the present studies suggest that the flavor-Tween paradigm may fill this empty cell.

Some might question whether the very concept of "a merely referential relation with the emotional-affective US characteristics" should not better be replaced by the idea of the CS being changed *intrinsically*. This would imply that in the flavor-Tween paradigm, the acquired valence of the CS is not based on or mediated by an association between CS and US representations, but rather that, due to the contingent presentation of a neutral and a valued event, the representation of the flavor CS itself is — in one way or another — endowed with affective value (see for example Garcia, Rusinak, & Brett, 1977; Grings, 1983; Martin & Levey, 1987; Rozin & Zellner, 1985). We prefer our own view for two main reasons. The first has to do with parsimony: if evaluative flavor-flavor learning is conceived of as resulting in an intrinsic change process, a different concept is required to conceptualize those instances of evaluative conditioning that do show sensitivity to US-inflation (e.g., picture-picture paradigm). The second reason is the lack of explanatory power of the intrinsic change concept: claiming that a CS is intrinsically changed is one thing, but explaining or understanding such an intrinsic change process is quite another. As far as we are aware of, up to now no single satisfactory theoretical account of an "intrinsic change process" has been proposed. Rather than replacing the notion of "a merely referential relation with the emotional-affective US

characteristics" by the intrinsic change theory, we would argue that the former associative structure provides a plausible way to conceptualize a so-called intrinsic change in CS valence.

Independently of this conceptual issue, a question that immediately imposes itself but at the same time is hard to answer given the evidence available at present, is *why* this type of evaluative flavor-flavor learning apparently involves only an encoding of the affective US characteristics, whereas evaluative learning in the face-face paradigm (Baeyens et al., 1992) was shown to be based on an associative link with a representation of the US specifying its specific sensory-descriptive characteristics. We hypothesize that this difference may be due to the different nature of the US in both paradigms. At a phenomenological level, a colored picture of a human face has much more distinct and detailed sensory-descriptive characteristics than the flavor of Tween20. Additionally, intuitively it makes sense that the affective-evaluative characteristics of the stimulus dominate the perception of Tween20, or at least more so than in case of a picture of a human face: Tween20 essentially is "a disgusting substance", whereas a liked picture of a human face may be perceived more as "that lovely girl with long blond hairs, head slightly aside, with the sad look in her pale grey eyes, her sensual pink-red mouth, upper white teeth visible, an almost transparent skin, and a glimpse of a blue cotton dress appearing". It may be the case that this phenomenological impression is paralleled at a representational level by a picture of a human face US having a much more elaborate and richer sensory-descriptive representation than a Tween-US, which may be represented as a negative, disliked entity without much further sensory-descriptive specification. From this admittedly speculative account follows that we hypothesize that US-inflation insensitivity is not necessarily a characteristic of flavor-flavor learning in general (no more than that US inflation sensitivity would be a characteristic of the picture-picture paradigm in general), but largely depends on the specific nature of the flavor US. Hence, when a flavor CS would be paired with a flavor US that has more pronounced and complex sensory-descriptive characteristics (for example, by adding citric acid and quinine to Tween20, making the US a sour, bitter, disgusting stimulus), we would expect the acquired dislike for the CS to be sensitive to US-inflation.

Finally, we would like to add that apart from its theoretical relevance, the fact that there is now some evidence for evaluative learning which does not seem to involve a representation of the sensory-descriptive characteristics of the US, may have an important consequence for evaluative-conditioning inspired suggestions to the clinical practice. It has been argued that, to the extent that disturbed affective functioning as seen in phobic, preference, or compulsive disorders, is mainly or partially based on an evaluative conditioning process, it may be predicted that extinction-based exposure therapies will be of little help to modify the basic acquired valence of the CSs involved (Baeyens

et al., 1988, 1989; Merckelbach, de Jong, Arntz, & Schouten, 1993). Therefore, aside from counter-conditioning procedures (e.g., Eifert, Craill, Carey, & O'Connor, 1988), US-revaluation based intervention techniques have been proposed as a potentially fruitful way to deal with evaluative-conditioning based disturbances in stimulus (dis)liking (see Merckelbach, de Jong, & Van den Hout, 1994, for a critical discussion). Valid as this suggestion may be, the data presented here immediately make clear the potential limitations of this approach: a US-revaluation based intervention should only be expected to be effective *insofar as* the CS is associated with a sufficiently detailed representation of the sensory-descriptive characteristics of the US. In this respect, it becomes critically important for future research to focus at unravelling the factors determining whether evaluative conditioning results in only the emotional-affective US characteristics being encoded, or also the sensory-descriptive properties of the US.

References

- Baeyens, F., Crombez, G., De Houwer, J., & Eelen, P. (1996a). No evidence for modulation of evaluative flavor-flavor associations in humans. *Learning and Motivation*, 27, 200-241.
- Baeyens, F., Crombez, G., Hendrickx, H., & Eelen, P. (1995a). Parameters of human evaluative flavor-flavor conditioning. *Learning and Motivation*, 26, 141-160.
- Baeyens, F., Crombez, G., Van den Bergh, O., & Eelen, P. (1988). Once in contact always in contact: Evaluative conditioning is resistant to extinction. *Advances in Behaviour Research and Therapy*, 10, 179-199.
- Baeyens, F., & De Houwer, J. (1995). Evaluative conditioning is a qualitatively distinct form of classical conditioning: A reply to Davey (1994). *Behaviour Research and Therapy*, 33, 825-831.
- Baeyens, F., De Houwer, J., & Eelen, P. (1994). Awareness inflated, evaluative conditioning underestimated. *Behavioral and Brain Sciences*, 17, 396-397.
- Baeyens, F., Eelen, P., & Crombez, G. (1995b). Pavlovian associations are forever: On classical conditioning and extinction. *Journal of Psychophysiology*, 9, 127-141.
- Baeyens, F., Eelen, P., Crombez, G., & Van den Bergh, O. (1992). Human evaluative conditioning: Acquisition trials, presentation schedule, evaluative style, and contingency awareness. *Behaviour Research and Therapy*, 30, 133-142.
- Baeyens, F., Eelen, P., & Van den Bergh, O. (1990a). Contingency awareness in evaluative conditioning: A case for unaware affective-evaluative learning. *Cognition and Emotion*, 4, 3-18.
- Baeyens, F., Eelen, P., Van den Bergh, O., & Crombez, G. (1989). Acquired affective-evaluative value: Conservative but not unchangeable. *Behaviour Research and Therapy*, 27, 279-287.
- Baeyens, F., Eelen, P., Van den Bergh, O., & Crombez, G. (1990b). Flavor-flavor and color-flavor conditioning in humans. *Learning and Motivation*, 21, 434-455.

- Baeyens, F., Hendrickx, H., Crombez, G., & Hermans, D. (in press). Neither extended sequential nor simultaneous feature positive training result in modulation of evaluative flavor conditioning in humans. *Appetite*.
- Baeyens, F., Hermans, D., & Eelen, P. (1993). The role of CS-US contingency in human evaluative conditioning. *Behaviour Research and Therapy*, 31, 731-737.
- Baeyens, F., Vansteenwegen, D., De Houwer, J., & Crombez, G. (1996b). Observational conditioning of food valence in humans. *Appetite*, 27, 235-250.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Dawson, M.E., & Shell, A.M. (1987). Human autonomic and skeletal classical conditioning: The role of conscious factors. In G. Davey (Ed.), *Cognitive processes and Pavlovian conditioning in humans* (pp. 27-57). New York: Wiley.
- De Houwer, J., Baeyens, F., & Eelen, P. (1994). Verbal evaluative conditioning with undetected US presentations. *Behaviour Research and Therapy*, 32, 629-633.
- De Houwer, J., Hendrickx, H., & Baeyens, F. (in press). Evaluative learning with "subliminally" presented stimuli. *Consciousness and Cognition*.
- Davey, G.C. (1994). Is evaluative conditioning a qualitatively distinct form of classical conditioning? *Behaviour Research and Therapy*, 32, 291-299.
- Eifert, G.H., Craill, L., Carey, E., & O'Connor, C. (1988). Affect modification through evaluative conditioning with music. *Behaviour Research and Therapy*, 26, 321-330.
- Garcia, J., Rusinak, K.W., & Brett, I.P. (1977). Conditioning food-illness aversions in wild animals: Caveant canonici. In H. Davis & H.M.B. Hurwitz (Eds.), *Operant-Pavlovian interactions*. NJ: Erlbaum.
- Grings, W.W. (1983). Comments on A. Levey and I. Martin's "Cognitions, evaluations and conditioning". *Advances in Behaviour Research and Therapy*, 4, 205-207.
- Hamm, A.O., & Vaitl, D. (1996). Affective learning: Awareness and aversion. *Psychophysiology*, 33, 698-710.
- Hammerl, M., & Grabitz, H.-J. (1996). Human evaluative conditioning without experiencing a valued event. *Learning and Motivation*, 27, 278-293.
- Hendrickx, H., De Houwer, J., Baeyens, F., Eelen, P., & Van Avermaet, E. (1997). Hidden covariation detection might be very hidden indeed. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 23, 201-220.
- Konorski, J. (1967). *Integrative activity of the brain*. Chicago: University of Chicago Press.
- Ledoux, J.E., Romanski, L., & Xagoraris, A. (1989). Indelibility of subcortical emotional memories. *Journal of Cognitive Neuroscience*, 1, 238-243.
- Levey, A.B., & Martin, I. (1987). Evaluative conditioning: A case for hedonic transfer. In H.J. Eysenck & I. Martin (Eds.), *Theoretical foundations of behavior therapy*. New York: Plenum Press.
- Mackintosh, N.J. (1983). *Conditioning and associative learning*. New York: Oxford University Press.
- Martin, I., & Levey, A.B. (1987). Learning what will happen next: Conditioning, evaluation and cognitive processes. In G. Davey (Ed.), *Cognitive processes and Pavlovian conditioning in humans* (pp. 57-82). New York: Wiley.
- Rescorla, R.A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, 43, 151-160.

- Rozin, P., & Zellner, D. (1985). The role of Pavlovian conditioning in the acquisition of food likes and dislikes. In N.S. Braveman & P. Bronstein (Eds.), *Experimental assessments and clinical applications of conditioned food aversions* (pp. 189-202). New York: New York Academy of Sciences.
- Merckelbach, H., de Jong, P., Arntz, A., & Schouten, E. (1993). The role of evaluative learning and disgust sensitivity in the etiology and treatment of spider phobia. Special Issue: Phobia: Etiological, cognitive, and physiological aspects. *Advances in Behaviour Research and Therapy*, 15, 243-255.
- Merckelbach, H., de Jong, P., & Van den Hout, M. (1994). Twee vormen van klassieke conditionering: twee vormen van therapie? *Gedragstherapie*, 27, 309-324.
- Shanks, D.R., & Dickinson, A. (1990). Contingency awareness in evaluative conditioning: A comment on Baeyens, Eelen, and Van den Bergh. *Cognition and Emotion*, 4, 19-30.
- Shanks, D.R., & St. John, M.F. (1994). Characteristics of dissociable human learning systems. *Behavioural and Brain Sciences*, 17, 367-447.
- Wagner, A.R., & Brandon, S.E. (1989). Evolution of a structured connectionist model of Pavlovian conditioning (AESOP). In S.B. Klein & R.R. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory*. Hillsdale, NJ: Erlbaum.
- White, K., & Davey, G. (1989). Sensory preconditioning and UCS inflation in human "fear" conditioning. *Behaviour Research and Therapy*, 27, 161-166.
- Zellner, D.A., Rozin, P., Aron, M., & Kulish, C. (1983). Conditioned enhancement of human's liking for flavor by pairing with sweetness. *Learning and Motivation*, 14, 338-350.

Received, March 1998

Revision received, August 1998

Accepted, September 1998