# Evaluative conditioning: a review and a model

Netherlands Journal of Psychology (March 2007) 63:31–41
DOI 10.1007/BF03061060

ARTICLE

# Evaluative conditioning: a review and a model

**Remco C. Havermans · Anita Jansen**

**Keywords** affect · configural learning · evaluative conditioning · generalisation · Pavlovian conditioning

Evaluative conditioning refers to the transfer of affective value to an initially neutral stimulus by pairing the neutral stimulus with an affective stimulus (positive or negative). In the first part of this article the associative nature of evaluative conditioning is discussed and existing models of evaluative learning are described. It is argued that these models of evaluative conditioning do not provide a complete account of evaluative conditioning and generate non-specific predictions. The second part of the article concerns the development of a more specific model of evaluative conditioning, describing such conditioning in terms of Pearce's 1987 model of stimulus generalisation and configural learning. By viewing the transfer of affect as a generalisation and configural learning process, this model can account for most demonstrations of evaluative conditioning and generates more precise and unambiguous predictions. (*Netherlands Journal of Psychology*, 63, 38-49).

As preferences and aversions are important determinants of human behaviour, much research has been devoted to understanding the nature of human likes and dislikes. Some preferences and aversions are innate, such as the preference for sweet tastes and the aversion for bitter tastes. However, most preferences and aversions are acquired with experience (Capaldi, 1996). Associative learning has been proposed as the primary mechanism underlying such development of likes and dislikes (De Houwer, Thomas, & Baeyens, 2001). In 1975, Levey and Martin demonstrated associatively acquired affect. In this seminal study they first let participants categorise pictures of paintings as liked, neutral, or disliked. In the subsequent conditioning phase, neutral pictures were consistently paired with either a liked, disliked, or another neutral picture. At test, where participants had to evaluate the pictures that had been presented during conditioning, it was found that the neutral pictures paired with the disliked stimuli were evaluated as more negative. Similarly, the neutral pictures paired with the liked pictures were now evaluated as more positive. Martin and Levey (1978) coined the term evaluative conditioning (EC) in referring to this apparent associative learning of preferences and aversions.

EC appears to be a reliable effect. Numerous picture-picture studies have successfully demonstrated transfer of affect, using pictures of fountains and sculptures (e.g., Hammerl & Grabitz, 1993), or pictures of human faces (e.g., Baeyens, Eelen, Van den Bergh, & Crombez, 1989). Apart from being a reliable effect, EC also appears to be a general phenomenon. The transfer of affect has been shown to occur using different flavours as stimuli (e.g., Zellner, Rozin, Aron, & Kulish, 1983; Baeyens, Eelen, Van den Bergh, & Crombez, 1990; Havermans & Jansen, 2007), or haptic stimuli (Hammerl & Grabitz, 2000). Several studies have also demonstrated cross-modal EC (see e.g., Todrank, Byrnes, Wrzesniewski, & Rozin, 1995). It should be noted though that cross-modal EC appears to be somewhat less robust than EC using stimuli of the same perceptual modality, as demonstrated by several failures to obtain cross-modal EC (see Rozin, Wrzesniewski, & Byrnes, 1998).

Remco C. Havermans, and, (✉)
*Maastricht University, Maastricht*Correspondence to: Remco Havermans, Experimental Psychology, Faculty of Psychology, Maastricht University, PO Box 616, NL 6200 MD Maastricht.
E-mail: r.havermans@psychology.unimaas.nlSubmitted 10 November 2006; revision accepted 23 April 2007.

Although EC has been studied extensively over the last three decades there is still considerable debate concerning the associative nature of the learning of likes and dislikes. In this article, we summarise different demonstrations of EC and discuss to what degree these findings can be conceptualised as a form of Pavlovian stimulus-stimulus learning. The foremost accounts of EC are evaluated in the light of this discussion. It is argued that none of the present models of EC provide a satisfactory account of EC. In the second part of this article we discuss the development of a new model to account for the variety of EC effects.

## The associative nature of evaluative conditioning

In general, it is suggested that EC comprises the learning of an association between a neutral stimulus and an affective stimulus. EC thus appears similar to Pavlovian stimulus-stimulus learning. In Pavlovian conditioning (PC), one learns an association between a neutral conditioned stimulus (CS; e.g., a tone, or a light) and a biologically relevant unconditioned stimulus (US; e.g., food, or an electric shock). Due to the formation of such an association, the mere presentation of a CS comes to elicit conditioned responding in anticipation of the US (e.g., appetitive behaviour when the US is food, or fear-motivated behaviour when the US is an electric shock) (Pearce & Bouton, 2001). In EC, the neutral stimulus is usually referred to as the CS and the affective stimulus as the US. As De Houwer et al. (2001) note, there are both similarities and discrepancies between findings in PC and EC.

### Similarities between EC and PC

In PC an association between the CS and US gradually increases in strength to an asymptotic value with an increasing number of trials, which can be represented by a typically negatively accelerated learning curve. Some EC studies have found such a significant gradual change in affective value of the CS with an increasing number of trials (e.g., Baeyens, Eelen, Crombez, & Van den Bergh, 1992). However, De Houwer et al. (2001) note that there have also been failures to demonstrate such a gradual acquisition of affect. This can be explained by arguing that the formation of an evaluative association occurs exceptionally rapidly. Indeed, significant transfer of affect can be shown after just a single pairing of the CS and the affective US (Stuart, Shimp, & Engle, 1987).

There are also several functional characteristics of EC that parallel PC. One of these characteristics concerns the demonstration that EC is sensitive to a counterconditioning procedure. In demonstrating counterconditioning of acquired affective value, Baeyens et al. (1989) conducted a picture-picture EC study. After baseline ratings of the pictures, neutral CSs were paired with either positive, or negative USs. After the acquisition phase the CSs were rated as more positive (when paired with a liked picture) or more negative (when paired with a negative US), indicative of EC. After the post-acquisition assessment of likes and dislikes, some CSs were now paired with a US with the opposite affective value than during the acquisition phase. The other CSs were either presented alone or not presented at all. After this phase of the experiment participants had to rerate all the CSs. The rating of the CSs that had been paired with the US of opposite affective value returned to baseline level, whereas evaluation of the other CSs did not change in comparison with the post-acquisition ratings.

Hammerl and Grabitz (1996) demonstrated evaluative sensory preconditioning in two experiments, also using a picture-picture paradigm. In the preconditioning phase, participants received several sequential pairings of two neutral CSs. In the second phase, $CS_2$ was paired with a liked US. At test, transfer of positive affect was not only apparent for $CS_2$ but also for $CS_1$ that had not been paired with the US. This finding corresponds well with the demonstration of Pavlovian sensory preconditioning (see e.g., Rescorla & Durlach, 1981).

As in PC, EC also appears sensitive to stimulus preexposure effects. For instance, Stuart et al. (1987) demonstrated that CS preexposure attenuates EC. More recently, De Houwer, Baeyens, Vansteenwegen, and Eelen (2000) also found some indication that preexposure to the CS may attenuate the transfer of affect to this CS in a subsequent EC procedure. Similarly, EC is sensitive to US preexposure. Hammerl, Bloch, and Silverthorn (1997) found that repeated preexposure of the US in a picture-picture study attenuated EC. This is in accordance with the finding that US preexposure leads to reduced conditioned responding in PC (see Randich & Lolordo, 1979).

Another effect that parallels findings in PC is that EC is sensitive to US revaluation. Baeyens, Eelen, Van den Bergh, and Crombez (1992) conducted a picture-picture study using photos of human faces. After an initial acquisition phase, the USs were revalued by accompanying the liked picture with negative adjectives and the disliked picture with positive adjectives. At test, the ratings of the CSs showed a shift in a similar direction to the shift in affective value of the USs due to revaluation. This effect was still present at a one-month follow-up.

### Discrepancies between EC and PC

Although EC appears to closely resemble PC, there are some notable differences between EC and PC. For example, contiguity between the CS and US is important in demonstrating PC. This is also certainly true for EC, but as opposed to PC, where the most learning is apparent when the CS directly precedes the US (i.e., a delayed conditioning procedure), EC works particularly well when the CS and the affective US are presented in compound (i.e., a simultaneous conditioning procedure) (see Rozin et al., 1998). Further, in contrast to general findings in PC, EC can be demonstrated using a backward conditioning procedure, where the US precedes the presentation of the CS (Martin & Levey, 1978; Stuart et al., 1987). It should be noted though that this effect has not always been obtained (see Hammerl & Grabitz, 1993).

A more notable discrepancy with PC is the finding that statistical contingency does not play an important role in the transfer of affect. Baeyens, Hermans, and Eelen (1993) manipulated the degree of statistical contingency between the CS and the US in a picture-picture study. Transfer of affect, indicative of EC, did not differ between a group in which the CS was always paired with the US, a group in which the CS was paired with the US in half of the trials, and a group in which the CS was paired with the US in only a third of the trials. Closely related to this discrepancy is the repeated finding that for EC to take place, one does not need to be aware of the contingency between the CS and the US. Although there is still considerable debate about the implicit nature of EC (see Field, 2000; Lovibond & Shanks, 2002), the general conclusion is that EC does not require contingency awareness as opposed to PC (see De Houwer et al., 2001).

Another important discrepancy with PC is that EC appears to be highly resistant to an extinction procedure. Whereas in PC the learned response can be extinguished by presenting the CS without the US, such postacquisition nonreinforced exposure to the CS typically does not affect the acquired affective value of the CS (e.g., Baeyens, Crombez, Hendrickx, & Eelen, 1995; De Houwer et al., 2000). This does not necessarily imply that EC is different from PC. Pavlovian conditioned responses cannot be extinguished in the sense that they are unlearned either. When a Pavlovian CS is presented outside the extinction treatment, conditioned responding typically recovers (see Bouton, 1993). It has been argued that such a renewal effect can also account for the apparent resistance of EC to extinction. Whereas the extinction of Pavlovian conditioned responding is measured during extinction treatment, the extinction of EC is usually assessed after an extinction procedure. The assessment of the affective value of the CS at test after extinction may be regarded as a novel situation, or context, allowing for the renewal of EC (Lipp, Oughton, & LeLievre, 2003). Indeed, Lipp and colleagues (2003) demonstrated extinction of EC when assessing affective value of the CS during the extinction treatment. It should be noted though that the extinction of EC was less complete and far less rapid than the simultaneous extinction of Pavlovian conditioned skin conductance responding. Therefore, one can still argue that EC is exceptionally, though not completely, resistant to extinction.

Similar to the resistance to extinction, EC also appears to be resistant to modulation. In Pavlovian discrimination learning, responding to a CS can come under modulatory control of a feature stimulus when the CS is reinforced in the presence of, or when preceded by the feature stimulus (feature positive discrimination). Similarly, responding to the CS can come under modulatory control when it is reinforced only in the absence of the feature stimulus (feature negative discrimination). The feature stimulus is said to set the occasion for the CS to be reinforced (in feature positive discrimination training), or not reinforced (in the case of feature negative discrimination training) (Holland, 1983; Rescorla, 1985). Baeyens and colleagues (Baeyens, Crombez, De Houwer, & Eelen, 1996; Baeyens, Hendrickx, Crombez, & Hermans, 1998) have repeatedly failed to demonstrate such occasion setting in EC.

### Conclusions

In many respects EC appears to be similar to PC. Nonetheless, considering the specific characteristics of EC, particularly its resistance to extinction and the unimportance of contingency (awareness), EC is usually regarded as a form of associative stimulus-stimulus learning which qualitatively differs from PC (but see Davey, 1994a). Since EC differs from PC, EC cannot be accounted for by well-defined models of associative learning, such as the Rescorla-Wagner model (Rescorla & Wagner, 1972). Therefore, different models of EC have been proposed to account for both the similarities and discrepancies with Pavlovian learning.

## Models of EC

### The conceptual categorisation model of EC

Davey (1994b) has argued that EC does not reflect associative learning at all, but rather conceptual categorisation. According to Davey, a CS usually has some features in common with the US it is paired with. Due to this pairing, the common features become more salient and thus the CS is categorised as being more similar to the US. Field and Davey (1997) demonstrated how the pairing of stimuli promotes conceptual categorisation. In this study, participants were presented with pictures of faces of Martians and Venusians. The researchers constructed these faces and the exemplar Martian and Venusian face differed on six specific features. In the experiment, ambiguous alien faces (sharing as many Martian as Venusian features) were paired with more prototypical Martians or Venusians. At test, the ambiguous aliens were evaluated as more Venusian-like when paired with a prototypical Martian, whereas the ambiguous aliens that had been paired with the prototypical Venusian were rated as more Martian-like. Although these results seem to be at odds with the general finding in EC that the evaluative shift is in the direction of the US, Field and Davey state that these results nonetheless show how pairing of stimuli leads to categorisation and thus argue that such conceptual categorisation may also play an important role in demonstrating EC.

Field and Davey (1999) point out that in several EC studies, CS-US pairs were constructed on the basis of perceptual similarity, hence promoting conceptual categorisation. Further, in a picture-picture study, they found no evidence for EC when the paired CSs and USs were perceived as dissimilar. Therefore, they argue that most demonstrations of EC are not the result of associative learning, but the result of an artefact due to the experimental paradigm (see also Shanks & Dickinson, 1990). De Houwer et al. (2001) however argue that in many EC studies, CSs were assigned to USs on a random basis, or stimuli were counterbalanced, thus controlling for perceptual similarity effects. Baeyens, De Houwer, Vansteenwegen, and Eelen (1998) further argue that the conceptual categorisation account of EC cannot explain cross-modal EC. It is highly unlikely that the CS and the US are perceived as being similar in this situation. Furthermore, it cannot explain why US revaluation should specifically affect the evaluation of the CS that has been paired with the revalued US (see De Houwer et al., 2001).

## The holistic representation model of EC

Martin and Levey (1994) do not discard EC as an artefact, but like Davey (1994b) they do not regard EC as the formation of an association. Martin and Levey describe EC as the automatic formation of a holistic representation containing elements of both the CS and the US. When a neutral and an affective stimulus are presented in compound, they are automatically fused or integrated into a single representation. According to these authors, this fusion process constitutes one of the most primitive and basic forms of learning. In EC, the CS is thought to activate the holistic CS-US representation and thus the affective value of the US, accounting for the observed evaluative shift. It accounts for most findings in EC, but encounters some difficulty in explaining sensory preconditioning as demonstrated by Hammerl and Grabitz (1996). Since the formation of a holistic representation depends on the co-occurrence of the CS and the US, a CS that has never been paired with the US should not be able to acquire affective value according to the holistic representation account.

## The referential learning model of EC

Baeyens and colleagues (Baeyens, Eelen, Crombez, & Van den Bergh, 1992; Baeyens & De Houwer, 1995) do describe EC in terms of the learning of an association, albeit different from Pavlovian learning. According to Baeyens, Eelen, Crombez, et al. (1992), human associative learning comprises two distinct learning systems: a signal learning system and a referential learning system. The signal learning system controls PC. In PC one has to be aware of the contingency between the CS and the US to be able to adequately anticipate the US when the CS is presented. However, in EC, controlled by the referential learning system, the CS merely has to refer or activate the representation of the US for effective transfer of affect. The referential learning model thus requires far less information processing resources and hence does not require much conscious processing.

The referential learning model is very similar to the account by Martin and Levey (1994), but as opposed to their account of EC, Baeyens and De Houwer (1995) do not regard EC as the basis of more complex stimulus learning such as PC. The referential learning model also explains most of the results of EC studies. However, Stevenson, Boakes, and Wilson (2000) point out that this model only allows affective value to transfer. Stevenson and colleagues (Stevenson, Prescott, & Boakes, 1995; Stevenson, Boakes, & Prescott, 1998; Stevenson, Boakes, & Wilson, 2000) have repeatedly demonstrated that not only affect, but also other salient attributes of the US can transfer. For example, in one of their experiments, Stevenson et al. (2000) found that liking for a sucrose-paired odour was greater than liking for a citric-paired odour. Next to the transfer of affect, the sucrose-paired odour was also rated as more sweet and the citric-paired odour as more sour at test. De Houwer et al. (2001) acknowledge that the referential learning model cannot account for this result, unless it is presumed that the referential learning system controls all but Pavlovian anticipatory responding.

In summary, the present models of EC cannot account for all demonstrations of EC and are difficult to distinguish as they render roughly the same set of predictions. De Houwer et al. (2001, p. 866) conclude that: '… these models tell us little about the specific processes that underlie EC and thus do not permit the formulation of precise hypotheses. It is therefore imperative that existing theories are re-examined and refined or that new models are proposed.'

## Stimulus generalisation and configural learning as a model of EC

In the first part of this article, we reviewed and summarised the different findings concerning EC and discussed to what degree EC can be conceptualised as a form of Pavlovian stimulus-stimulus learning (for a more comprehensive review concerning this topic we refer to De Houwer et al., 2001). The present models of EC as described above all provide a relatively complete account of EC, but they cannot explain all the findings and do not generate very precise predictions. Typically, these models either fail to account for the apparent importance of perceptual similarity between stimuli in demonstrating EC, or fail to explain EC when paired stimuli are perceptually dissimilar. In reviewing the findings in EC studies we argue in the following sections that EC is perhaps best understood in terms of stimulus generalisation and configural learning.

In 1987, Pearce devised a model of stimulus generalisation, describing the role of generalisation and configural learning in the context of PC. In accordance with Atkinson and Estes (1963), Pearce presumes that an individual possesses a buffer containing representations of all the stimuli present in the context to which that individual is exposed. According to Pearce, when a stimulus is presented it activates a representation of that particular stimulus in the buffer. Through a process of spreading activation it can also activate representations of stimuli that are not presented but share features in common with the presented stimulus. So when a CS X has acquired excitatory associative strength due to pairings with a US (e.g., food), some conditioned responding will be elicited by a novel CS Y when this stimulus shares features in common with X. This principle of stimulus generalisation as formulated by Pearce can also be applied to EC. According to the present model, the hedonic shift observed in EC is due to generalisation of affective value from the affective US to the CS. Similar to the generalisation of associative strength, we argue that the amount of generalised affective value is determined by the perceptual similarity of the neutral stimulus with the affective US. Suppose a stimulus X has a certain affective value. The degree of generalised affect from X to a stimulus Y then depends on the perceptual similarity of X and Y and, of course, the affective value of stimulus X. Equation 1 represents this relationship and it directly

$$a_Y = {}_X S_Y \times A_X .$$

**Equation 1**

corresponds with the equation of generalised associative strength described by Pearce (1987).

Parameter $a_Y$ represents the total amount of generalised affective value to stimulus Y. The value of $a_Y$ depends on the similarity between stimuli X and Y, represented by the parameter ${}_X S_Y$, which can also be described as the degree to which stimulus Y is able to directly activate a representation of X. Affective value of stimulus X is represented by the parameter $A_X$ and can take on a value anywhere between –1 (negative valence) and 1 (positive valence).

The $S$ parameter in Equation 1 requires some more specification. In accordance with Pearce (1987, 2002) we argue that the similarity ($S$) between two stimuli depends on the number of common features and the total number of activated representations by each stimulus. Pearce (1987, 1994) formulates similarity as displayed in Equation 2.

$${}_X S_Y = \frac{n_C}{n_X} \times \frac{n_C}{n_Y}.$$

**Equation 2**

Equation 2 describes how $S$ depends on the ratio of the number of common features $n_C$ (not only including the perceptual features of the stimuli, but also the context in which these stimuli are presented) and the total number of activated representations by stimulus X, as well as the ratio of the number of common features and the total number of activated representations by stimulus Y. If one assumes that stimulus X and Y are equally intense and thus activate the same number of common representations, Equation 2 can be rewritten as Equation 3.

$${}_X S_Y = \frac{n_C^{\,2}}{(n_X \times n_Y)}$$

**Equation 3**

Although formulated somewhat differently than by Davey (1994b), we also argue that the perceptual similarity between stimuli plays a key role in demonstrating EC. Using a picture-picture paradigm, Field and Davey (1999) demonstrated the importance of perceptual

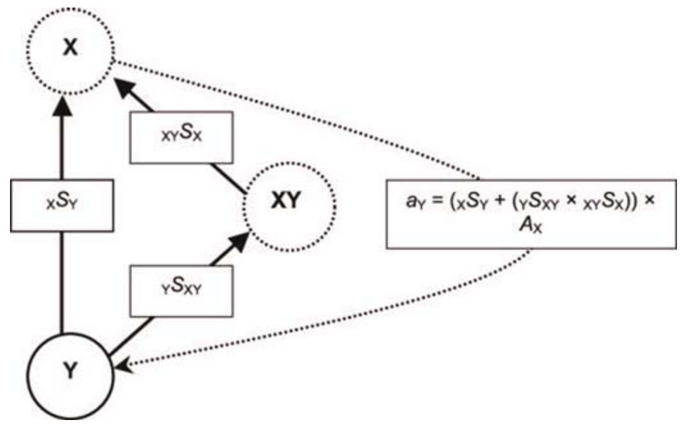


**Figure 1** A schematic representation of how an initially neutral stimulus Y acquires affective value from an affective stimulus X. Due to pairing of X and Y a configural XY representation is formed. When Y is presented, affective value of X ($A_X$) will generalise to Y ($a_Y$) to the extent that Y is capable of directly activating a representation of X ($_XS_Y$) and activating the representation of X via the configural XY representation ($_YS_{XY} \times {}_{XY}S_X$).

similarity in EC. Pictures of human faces were evaluated after which neutral pictures (CSs) were paired with either liked or disliked pictures (the USs). At test, participants had to rerate the pictures. Shifts in affective value were found, but only for the CSs that were categorised as being similar to the US they had been paired with. Moreover, such transfer of affect was also found in a control group that had received explicitly unpaired presentations of the CSs and the USs and in another control group of participants who simply had to rate the pictures twice. Note that the participants in these control groups were not informed beforehand that pictures could or could not be paired.

Although perceptual similarity certainly plays an important role in EC, De Houwer et al. (2001) emphasise that perceptual similarity is not a prerequisite for demonstrating EC. The stimulus generalisation model of EC as presented thus far, also cannot provide an account of EC when the CS and the US are perceptually dissimilar. However, the generalisation of affective value can be promoted through the formation of a configural representation of the CS and the affective US. The potentially important role of configural representations in human associative learning has already been noted by several researchers (e.g., Shanks, Charles, Darby, & Azmi, 1998; Shanks, Darby, & Charles, 1998; Stevenson et al., 2000).

According to Pearce (1987), a configural representation is formed when two or more stimuli are presented in compound. Pearce (2002) has described how such configural representations may play a role in animal flavour-flavour learning, a phenomenon that bears close resemblance to the EC paradigm. In animal flavour-flavour learning, animals receive pairings of artificial flavours paired with an inherently preferred flavour (e.g., a sweet flavour). This leads to the acquisition of a conditioned flavour preference for the previously neutral flavour. As in EC, conditioned flavour preferences are acquired very rapidly and are highly resistant to an extinction procedure (see Capaldi, 1996). Pearce states that whenever a neutral flavour is paired with the preferred flavour, the co-activation of the representation of the neutral and preferred flavour leads to the formation of a configural representation, comprising elements of both flavours. When the neutral flavour is presented at test, a representation of the preferred flavour is activated indirectly through the activation of the configural representation. The degree to which the initially neutral flavour is able to activate the configural representation rapidly increases, requiring but a few trials.

Figure 1 displays how this notion of configural learning can also be applied to EC. As described by the figure, when an affective stimulus X is paired with a neutral stimulus Y, the co-activation of the representation of X and Y leads to the formation of an XY configural representation. When Y is presented at test, it activates the representation of X directly to the extent $_XS_Y$, and indirectly through the activation of XY to the extent $_YS_{XY} \times {}_{XY}S_X$. As such, the stimulus generalisation model does not preclude the possibility of cross-modal EC, in contrast to Davey's model (1994b). However, it does predict that cross-modal EC is more difficult to demonstrate. Further, it predicts that effective EC requires but a few CS-US pairings.

Summarised, affective value can generalise both directly and indirectly from multiple affective sources. Therefore, Equation 1 can be extended to the more general Equation 4, which again directly corresponds with Pearce's 1987 notion of generalisation of associative strength in the context of PC. Equation 4 provides a formula for determining the degree of generalised

$$a_Y = \sum{}_n S_Y \times A_n.$$

**Equation 4**

affective value to stimulus Y from *n* stimuli similar to Y having some affective value (positive, or negative).

The present stimulus generalisation model of EC explains the irrelevance of contingency in EC by stating that EC partly depends on the formation of a configural representation. This latter process of configuring is thought to occur automatically relying solely on the co-activation of the representations of the presented stimuli. Pearce's 1994 connectionist model of configural learning describes the formation of a configural representation as an automatic and immediate non-iterative process, as opposed to the gradual formation of a Pavlovian association between stimuli. Differing levels of contingency should thus exert no effect on the demonstration of EC and one need not be aware of the specific contingency between the CS and the US.

The model also predicts that simultaneous presentations of the CS and the US should be most effective in demonstrating EC. Although Rozin et al. (1998) note that this seems to be the case, EC has been frequently demonstrated using forward pairings of the CS and the US. In its present form the stimulus generalisation model of EC does not permit the formation of a configural representation when the CS and the US are presented sequentially. As Pearce (2002) notes though, the model can accommodate the formation of a configural representation when stimuli are presented in a serial compound if it is assumed that activation of the representation of a stimulus is not terminated instantly with the offset of the presented stimulus, but gradually decays.

In summary, the present model predicts that perceptual similarity plays a key role in demonstrating EC. It also predicts that contiguity, but not contingency, is important in demonstrating EC. Apart from these initial predictions, the stimulus generalisation and configural learning model of EC can account for most typical EC findings.

## Application to extinction

As the model describes EC in terms of stimulus generalisation and configural learning, the model predicts that the acquired affective value should be extremely resistant to extinction. Once a configural representation is formed, nonreinforced exposure to the CS should not affect its ability to activate the configural CS-US representation, and hence generalisation of affective value should be relatively unaffected by an extinction procedure.

## Application to modulation

The model predicts that EC should resist modulation as Baeyens and colleagues (Baeyens et al., 1996; Baeyens et al., 1998) demonstrated. In one of their experiments (see Baeyens et al., 1996) the participants tasted a flavour (A) with a certain colour (X) paired with Tween (a taste which was generally rated as highly aversive by the participants). When the flavour was presented without the specific colour, the flavour was not paired with Tween. Participants thus received two trial types: a type in which flavour A was not paired with Tween (A0 trials) and a trial type in which the flavour-colour compound XA was paired with Tween (XA+ trials). Colour X did not modulate evaluative responding at all. At test, evaluative shifts were observed for the flavour A whether or not presented in compound with the colour. The model explains this pattern of findings by suggesting that during discrimination training an XA-$US_{Tween}$ configural representation was formed. Due to X being of a different sensory modality than both A and the US, the similarity between X and XA-$US_{Tween}$ would be much smaller than the similarity between A and XA-$US_{Tween}$. Generalisation of affective value from the US would thus be limited primarily to stimulus A and the XA compound as was demonstrated by Baeyens and colleagues (1996).

## Application to counterconditioning

Baeyens and colleagues (1989) found that EC is sensitive to a counterconditioning procedure. Although at face value this finding appears to imply that EC reflects some form of associative learning, the present nonassociative model of EC can accommodate this pattern of results. According to the model, in the initial acquisition phase a configural representation is formed containing both elements of the CS and the affective $US_1$. When in the following phase the CS is now paired with another affective stimulus ($US_2$) that has an affective value directly opposite to $US_1$, presentation of the CS at test will activate a representation of both USs. Assuming that the CS activates the USs to a similar extent, the net generalised affective value to the CS will approach zero.

## Application to US revaluation

Baeyens, Eelen, Van den Bergh, et al. (1992) showed that EC is sensitive to US revaluation. The present model can accommodate the observed US revaluation effect without referring to the formation of an association. The model predicts that when an affective US is revalued, the CS that has previously been paired with this US will show an evaluative shift in the same direction as the US, because at test it activates the representation of the revalued US through the activation of the previously acquired configural CS-US representation.

Application to sensory preconditioning

As described above, the finding of evaluative sensory preconditioning (Hammerl & Grabitz, 1996) posed serious implications for the holistic representation account of EC proposed by Martin and Levey (1994). Although the present model of EC is very similar to the notion of the formation of a holistic representation put forth by Martin and Levey, it does not preclude sensory preconditioning. In the first phase of the paradigm, it is predicted that a configural $CS_1$-$CS_2$ representation is formed. In the subsequent conditioning phase, a $CS_2$-US configural representation will be formed. When $CS_1$ is presented at test it will be able to activate the representation of the US directly through its perceptual similarity with the US and indirectly through the activation of the configural representations.

Application to stimulus preexposure

Another effect in EC that appears to demonstrate the associative nature of EC is the detrimental effect of stimulus preexposure on EC (e.g., Stuart et al., 1987; Hammerl et al., 1997). The present model cannot explain the detrimental effects of CS (or US) preexposure on EC, unless it is assumed that such preexposure leads to perceptual habituation to the preexposed stimulus, hence limiting the formation of a configural representation during the subsequent conditioning phase. Hammerl et al. made a somewhat similar proposal in explaining the effects of US exposure in EC. In concordance with Randich and Lolordo (1979), Hammerl and colleagues state that repeated US exposure leads to habituation to the US and as a consequence the US will lose some of its affective value. As such, both preexposure and postexposure of the US should attenuate EC.

Although the present model of EC encounters some difficulty in explaining stimulus preexposure effects on EC, it should be noted that these effects have not yet been studied extensively. Moreover, the proposed detrimental effect of CS preexposure on EC has not been found to be a reliable effect (see De Houwer et al., 2000; Stevenson et al., 2000).

**Tests of the model**

As described above, the stimulus generalisation and configural learning model of EC accommodates most EC findings. However, the question arises how to test the present model. The present model of EC states that any evaluative shift results from stimulus generalisation and is not a result of predicitive learning, that is PC. Therefore, it should be possible to dissociate these accounts of evaluative conditioning.

Discrimination learning

One of the primary assumptions of the present stimulus generalisation model is that similarity between stimuli is an important factor in demonstrating EC. In EC, a CS will show an evaluative shift in the direction of the affective value of the activated representation of the affective US it has been paired with. If a CS is thus paired with both a positive and negative stimulus, the evaluative shift at test will be determined by the degree in which the representations of these different USs are activated. When both USs are activated to the same extent and have directly opposing affective values, no EC will be apparent at test.

However, consider the following discrimination. A neutral stimulus A is paired with a negative US, another stimulus B is paired with a positive US. Next to these paired presentations, A is also presented in compound with another CS C and paired with the positive US, and B is also presented in compound with C and paired with a negative US. This then leads to an A – $US_1^-$/B – $US_2^+$/ AC – $US_2^+$/BC – $US_1^-$ discrimination. How would such a discrimination affect the transfer of affect to A, B, and C separately?

The referential model states that an evaluative association depends solely on the co-occurrence of the CS and the US. As each CS in this case co-occurs with both a positive and negative US, no significant EC is expected to be observed according to this model. The conceptual categorisation model renders the same prediction when one assumes that all individual stimuli are equally similar to one another and the positive and negative USs have directly opposing affective values. The holistic representation model does not specify when or how a particular holistic representation will be retrieved and thus does not render a specific prediction in this case. In contrast to these existing models of EC, the stimulus generalisation model predicts that one can 'solve' the discrimination. It predicts that A shows a negative shift, B a positive shift, while there is no substantial evaluative shift for stimulus C. Presentation of A is more likely to lead to the activation of the representation of the negative US, as A is more similar to the configural A – $US_1^-$ stimulus than the configural AC – $US_2^+$ representation. In contrast, B will more strongly activate a representation of the positive US, as stimulus B is more similar to the configural B – $US_2^+$ stimulus than the BC – $US_1^-$ representation. Presentation of C will activate the representation of both the positive and negative US to a similar extent, thus limiting EC. Whether one truly can 'solve' such a discrimination remains to be investigated.

### Compound conditioning

Another prediction of the present model concerns compound conditioning. In PC, when two CSs are presented in compound and paired with a US, this compound stimulus will be just as effective at eliciting conditioned responding as any other single CS paired with this US. The present model, however, predicts that when a compound comprising two neutral CSs (e.g., A and B) is paired with an affective US Y, the transfer of affective value to this compound will be less than the transfer of affective value from Y to a CS comprising a single neutral stimulus X. When the AB compound is presented at test, this will activate the AB-Y configural representation that in turn activates the A, B and Y representations. As the similarity between the AB-Y representation and Y is smaller than the similarity between the X-Y representation and Y, Y will be less activated by presentation of the AB compound as compared with X and thus, the transfer of affective value from Y to the AB compound will be less than the transfer from Y to X. Additionally, the AB compound also activates representations of its neutral elements A and B separately, and this further limits the transfer of affective value. None of the other models of EC make such a prediction, but again, whether compound conditioning is indeed different in EC in comparison with PC awaits further research.

### Learned irrelevance

In PC, when a CS is explicitly unpaired with the US, this typically retards the formation of an association when the CS is later made predictive of the US. This effect has been termed learned irrelevance and can be explained by arguing that during the uncorrelated presentation of the CS and the US, the context in which the CS and US are both presented acquires associative strength, hence blocking the later acquisition of a CS-US association when the CS is made predictive of the US (see Balsam & Tomie, 1985). If EC can be explained in terms of PC, then one would expect to be able to demonstrate such an effect within an EC paradigm. However, the present model predicts that – given that the CS and US are to some degree perceptually similar – even uncorrelated pairings of the CS and US may already lead to a shift in the subsequent evaluation of the CS. Therefore, prior uncorrelated presentations of the CS and an affective US promote rather than attenuate subsequent evaluative learning when the CS and affective US are presented contingently.

### Context specificity

Pavlovian conditioned responding can be context specific. For example, extinguished conditioned responding can be renewed when the CS is presented in a context that differs from the context in which the extinction treatment took place. This implies that extinction reflects inhibitory conditioning and that this second learned inhibitory meaning of the CS is context specific. Similar context specificity applies to Pavlovian excitatory conditioned responding when inhibition to the CS has been learned prior to excitatory conditioning training (Nelson, 2002). Generally though, initially acquired Pavlovian conditioned responding generalises easily across different contexts (Bouton, 1993). This, however, should not be the case with EC. A CS and an affective US are always paired within a specific context. Therefore, according to the present model at least, features of this context will be incorporated in the configural CS-US representation. As such, presenting the CS in another context should always lead to generalisation decrement and hence a loss of an initially acquired affective shift.

## Concluding comments

The primary aim of the present article is to develop a new model of EC. The model accommodates most EC findings and generates specific hypotheses concerning EC effects. It can be argued that previous models of EC either fail to explain the importance of perceptual similarity in demonstrating EC (e.g., Baeyens, Eelen, Crombez, et al., 1992; Martin & Levey, 1994), or fail to explain the possibility of the transfer of affect when stimuli are perceptually dissimilar (Davey, 1994b). By describing EC in terms of generalisation of affective value and emphasising the role of configural learning in such generalisation, the importance of perceptual similarity is incorporated in the present model while not precluding the possibility of EC when stimuli are perceptually dissimilar. As such, it can be argued that the present model provides a more comprehensive account of EC than previous models have.

One may argue that perceptual similarity lies in the eye of the beholder. Judging similarity between two stimuli may differ between species and even between individuals of the same species. Describing perceptual similarity in terms of the number of common elements between stimuli thus appears to be an oversimplification. Nonetheless, it allows for the simple manipulation of perceptual similarity. Increasing the number of common elements between two stimuli should make these stimuli more alike for each individual member of each type of species. Therefore, within an individual subject, increasing the number of common elements between the CS and the affective US should always lead to more effective evaluative learning.

As noted above, the present stimulus generalisation and configural learning model of EC is an adaptation of the 1987 Pearce model of stimulus generalisation in the context of Pavlovian learning. Since this model can be so easily modified to account for both the generalisation of associative strength and affective value, there is no reason to assume that stimulus generalisation should be restricted to the generalisation of these two potential stimulus attributes. So although the generalisation model is presented here as a model of EC, it can also explain the simultaneous transfer of other salient stimulus attributes.

Although the stimulus generalisation and configural learning account of EC provides a more complete and specific account of EC, as any other model it requires rigorous testing. What it has in common with other models of EC is that it describes EC as being distinct from PC. This means that in specific cases it should be possible to dissociate EC from PC. In this respect, the present model renders more specific predictions than previous models of EC. We agree with De Houwer et al. (2001) that future EC studies should be explicitly aimed at testing the boundary conditions of EC. Such studies should provide results with which one could evaluate the merits of the present model.

## Authors note

Remco Havermans' main field of interest concerns the associatively learned nature of appetitive behaviour, such as the motivation to eat and the acquired motivation to (mis)use drugs or alcohol. Anita Jansen's fields of interest mainly concern eating behaviour in general and mechanisms underlying the maintenance of eating disorders.

## References

Atkinson, R.C., & Estes, W.K. (1963). Stimulus sampling theory. In R.D. Luce, R.B. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 3, pp.121-268). New York: Wiley.

Baeyens, F., Crombez, G., De Houwer, J., & Eelen, P. (1996). 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. (1995). Parameters of human evaluative flavor-flavor conditioning. *Learning and Motivation, 26,* 141-160.

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., Vansteenwegen, D., & Eelen, P. (1998). Evaluative conditioning is a form of associative learning: On the artifactual nature of Field and Davey's (1997) artifactual account of evaluative learning. *Learning and Motivation, 29,* 461-474.

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., & 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. (1990). Flavor-flavor and color-flavor conditioning in humans. *Learning and Motivation, 21,* 434-455.

Baeyens, F., Eelen, P., Van den Bergh, O., & Crombez, G. (1992). The content of learning in human evaluative conditioning: Acquired valence is sensitive to US-revaluation. *Learning and Motivation, 23,* 200-224.

Baeyens, F., Hendrickx, H., Crombez, G., & Hermans, D. (1998). Neither extended sequential nor simultaneous feature positive training result in modulation of evaluative flavor conditioning in humans. *Appetite, 31,* 185-204.

Baeyens, F., Hermans, D., & Eelen, P. (1993). The role of CS-US contingency in human evaluative conditioning. *Behaviour Research and Therapy, 31,* 731-737.

Balsam, P.D., & Tomie, A. (1985). *Context and learning*. Hillsdale, NJ: Erlbaum.

Bouton, M.E. (1993). Context, time and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin, 114,* 80-99.

Capaldi, E.D. (1996). Conditioned food preferences. In E. D. Capaldi (Ed.), *Why we eat what we eat: The psychology of eating* (pp. 53-80). Washington, DC: American Psychological Association.

Davey, G.C.L. (1994a). Is evaluative conditioning a qualitatively distinct form of classical conditioning? *Behaviour Research and Therapy, 32,* 291-299.

Davey, G.C.L. (1994b). Defining the important questions to ask about evaluative conditioning: A reply to Martin and Levey (1994). *Behaviour Research and Therapy, 32,* 307-310.

De Houwer, J., Baeyens, F., Vansteenwegen, D., & Eelen, P. (2000). Evaluative conditioning in the picture-picture paradigm with random assignment of conditioned stimuli to unconditioned stimuli. *Journal of Experimental Psychology: Animal Behavior Processes, 26,* 237-242.

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Associative learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin, 127,* 853-869.

Field, A.P. (2000). I like it but I'm not sure why: Can evaluative conditioning occur without awareness? *Consciousness and Cognition, 9,* 13-36.

Field, A.P., & Davey, G.C.L. (1997). Conceptual conditioning: Evidence for an artifactual account of evaluative learning. *Learning and Motivation, 28,* 446-464.

Field, A.P., & Davey, G.C.L. (1999). Reevaluating evaluative conditioning: A nonassociative explanation of conditioning effects in the visual evaluative conditioning paradigm. *Journal of Experimental Psychology: Animal Behavior Processes, 25,* 211-224.

Hammerl, M., Bloch, M., & Silverthorne, C.P. (1997). Effects of US alone presentations on human evaluative conditioning. *Learning and Motivation, 28,* 491-509.

Hammerl, M., & Grabitz, H.-J. (1993). Human evaluative conditioning: Order of stimulus presentation. *Integrative Physiological and Behavioral Science, 28,* 191-194.

Hammerl, M., & Grabitz, H.-J. (1996). Human evaluative conditioning without experiencing a valued event. *Learning and Motivation, 27,* 278-293.

Hammerl, M., & Grabitz, H.-J. (2000). Affective-evaluative learning in humans: A form of associative learning or only an artifact? *Learning and Motivation, 31,* 345-363.

Havermans, R.C., & Jansen, A. (2007). Increasing children's liking of the taste of vegetables through flavour-flavour learning. *Appetite, 48,* 259-262.

Holland, P.C. (1983). Occasion setting in Pavlovian feature positive discriminations. In M.L. Commons, R.J. Herrnstein, & A.R. Wagner (Eds.), *Quantitative analyses of behavior: Discrimination processes* (Vol. 4, pp. 183-206). New York: Ballinger.

Levey, A.B., & Martin, I. (1975). Classical conditioning of human 'evaluative' responses. *Behaviour Research and Therapy, 13,* 221-226.

Lipp, O., Ougthon, N., & LeLievre, J. (2003). Evaluative learning in human Pavlovian conditioning: Extinct, but still there? *Learning and Motivation, 34,* 219-239.

Lovibond, P.F., & Shanks, D.R. (2002). The role of awareness in Pavlovian conditioning: Empirical evidence and theoretical implications. *Journal of Experimental Psychology: Animal Behavior Processes, 28,* 3-26.

Martin, I., & Levey, A.B. (1978). Evaluative conditioning. *Advances in Behaviour Research and Therapy, 1,* 57-102.

Martin, I., & Levey, A.B. (1994). The evaluative response: Primitive but necessary. *Behaviour Research and Therapy, 32,* 301-305.

Nelson, J.B. (2002). Context specificity of excitation and inhibition in ambiguous stimuli. *Learning and Motivation, 33,* 284-310.

Pearce, J.M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review, 94,* 61-73.

Pearce, J.M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review, 101,* 587-607.

Pearce, J.M. (2002). Evaluation and development of a connectionist theory of configural learning. *Animal Learning & Behavior, 30,* 73-95.

Pearce, J.M., & Bouton, M.E. (2001). Theories of associative learning in animals. *Annual Review of Psychology, 52,* 111-139.

Randich, A., & Lolordo, V.M. (1979). Associative and nonassociative theories of the UCS preexposure phenomenon: Implications for Pavlovian conditioning. *Psychological Bulletin, 86,* 523-548.

Rescorla, R.A. (1985). Conditioned inhibition and facilitation. In R.R. Miller & N.E. Spear (Eds.), *Information processing in animals: Conditioned inhibition* (pp. 299-236). Hillsdale: Erlbaum.

Rescorla, R.A., & Durlach, P.J. (1981). Within-event learning in Pavlovian conditioning. In N.E. Spear & R.R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 81-112). Hillsdale: Erlbaum.

Rescorla, R.A., & Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. Black & W.F. Prokasy (Eds.), *Classical Conditioning II* (pp. 64-99). New York: Apple-Century-Crofts.

Rozin, P., Wrzesniewski, A., & Byrnes, D. (1998). The elusiveness of evaluative conditioning. *Learning and Motivation, 29,* 397-415.

Shanks, D.R., Charles, D., Darby, R.J., & Azmi, A. (1998). Configural processes in human associative learning. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24,* 1353-1378.

Shanks, D.R., Darby, R.J., & Charles, D. (1998). Resistance to interference in human associative learning: Evidence of configural conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 24,* 136-150.

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.

Stevenson, R.J., Boakes, R.A., & Prescott, J. (1998). Changes in flavor sweetness resulting from implicit learning of a simultaneous flavor-sweetness association: An example of learned synesthesia. *Learning and Motivation, 29,* 113-132.

Stevenson, R.J., Boakes, R.A., & Wilson, J.P. (2000). Resistance to extinction of conditioned odor perceptions: Evaluative conditioning is not unique. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 423-440.

Stevenson, R.J., Prescott, J., & Boakes, R.A. (1995). The acquisition of taste properties by flavors. *Learning and Motivation, 26,* 433-455.

Stuart, E.W., Shimp, T.A., & Engle, R.W. (1987). Classical conditioning of consumer attitudes: Four experiments in an advertising context. *Journal of Consumer Research, 14,* 334-349.

Todrank, J., Byrnes, D., Wrzesniewski, A., & Rozin, P. (1995). Odors can change preferences for people in photographs: A cross-modal evaluative conditioning study with olfactory USs and visual CSs. *Learning and Motivation, 26,* 116-140.

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.