Review Article

PsychOpen GOLD

Social Psychological Bulletin
Psychologia Społeczna

A Model of Attribute Conditioning

Christian Unkelbach[1], Sabine Förderer[1]

[1] *Universität zu Köln, Köln, Gernany*

Corresponding author: Christian Unkelbach (Social Cognition Center Cologne, Universität zu Köln, Richard-Strauss-Str. 2, 50931 Köln, Germany. E-mail: christian.unkelbach@uni-koeln.de)

Handling editor: Robert Balas (Polish Academy of Sciences, Warsaw, Poland)

Received: 4 March 2018 • **Accepted:** 14 July 2018 • **Published:** 20 September 2018

Citation: Unkelbach, C., & Förderer, S. (2018). A model of attribute conditioning. *Social Psychological Bulletin*, *13*(3), Article e28568. https://doi.org/10.5964/spb.v13i3.28568

Related: This article is part of SPB's special issue on "Evaluative Conditioning – Theoretical Accounts", Social Psychological Bulletin, 13(3).

Abstract

We present a model of attribute conditioning, the phenomenon that people's assessment of stimuli's specific attributes (e.g., a person's characteristics) changes due to pairings with other stimuli possessing these specific attributes (e.g., another "athletic" person). These changes in attribute assessments go beyond evaluation changes due to these pairings (i.e., evaluative conditioning effects). We provide a short historical overview of the phenomenon and the available data. Then we present a potential mental model of the effect: We assume attribute conditioning to be a form of stimulus-stimulus learning. CS-US pairings establish an enduring referential link between CS and US. We present an associative as well as a distributed memory variant of this referential link. Based on this model, we provide the answers to the specific questions that guide the present special issue. Finally, we discuss the relation of evaluative and attribute conditioning.

Keywords

attitude formation, evaluative learning, stimulus-stimulus learning, referential link

When an initially neutral stimulus (e.g., a face) is paired with another positive stimulus (e.g., a smiling face) or negative stimulus (e.g., a frowning face), people tend to evaluate the initially neutral face differently afterwards. Typically, people rate the face to be more likeable in the former case, and to be less likeable in the latter case. This change in liking due to the pairings of an initially neutral stimulus (a conditioned stimulus, CS) with a positive or negative stimulus (an unconditioned stimulus, US) is called *evaluative conditioning* (EC; De Houwer, 2007; see Hofmann, De Houwer, Perugini, Baeyens, and Crombez, 2010, for a quantitative overview). However, pairings also change the assessment of other CS attributes as well. For example, when an initially neutral person is repeatedly paired with an athletic person or an unathletic person, the neutral person is also rated as more athletic or unathletic, respectively. This phenomenon is called *attribute conditioning* (AC; Förderer & Unkelbach, 2015).

In the following, we first present a historical overview of the AC phenomenon. Then, we discuss potential mental processes that may explain the AC phenomenon, based on the accounts that are typically recruited to explain EC effects and compare their explanatory power with regards to the available data for AC effects. Building on this analysis, we then describe a mental process model that accounts for the available data. Based on this model, we provide answers to the questions that have been guiding the present issue on evaluative conditioning. Finally, we embed our model within the broader context of the AC-EC relation. In the remainder, we refer to EC effects as evaluations, and to AC effects as assessments, to facilitate a distinction between EC and AC.

A Short Historical Overview

Again, the basic EC effect is a change in CS evaluations. However, it is rather surprising that the research focus has been almost exclusively on valence evaluations (e.g., likeability), especially given the ascribed explanatory breadth of EC, for example, for the acquisition of consumer preferences. In other words, if a neutral face is paired with a clearly positive IAPS picture of a person celebrating a victory on the tennis court (e.g., IAPS 8350; Lang, Bradley, & Cuthbert, 1997), why should people evaluate the CS face as only more likeable? Intuitively, one might assume that other attributes besides "likeability" change as well; for example, people may assess the CS as more athletic, too.

The answer to why changes in other attributes due to CS-US pairings have been largely neglected may be a methodological challenge. How much a person "likes" a given stimulus has substantial influences on almost any other judgments about the stimulus. Such "halo" effects (Gräf & Unkelbach, 2016; Thorndike, 1920) will and should change people's assessments of other attributes. First, global likeability impression changes other attribute assessments in the direction of the global impression (e.g., Nisbett & Wilson, 1977); that is, the likeable CS face should also be rated higher on other positive attributes such as intelligence, humor, or morality. Conversely, the unlikeable CS face should be rated lower on these attributes. Second, likeability is confounded with many other attributes; that is, people might judge the tennis court winner as likeable just because she is athletic (see Förderer & Unkelbach, 2011). To isolate specific influences of pairings on specific CS attributes is therefore not straightforward.

Putting aside the methodological problems, one of the very first and frequently cited papers on valence acquisition by pairings (i.e., EC effects) by Arthur Staats and Carolyn Staats (Staats & Staats, 1958; “Attitudes established by classical conditioning”) was preceded by a paper that, in principle, anticipated the AC phenomenon. The authors paired different adjectives (e.g., “pretty”, “sweet”, and “healthy”) with nonsense syllables and predicted that across different pairings of CSs (here: nonsense syllables), the evaluative meaning would be learned as the common component of the attribute words (USs) (Staats & Staats, 1957; “Meaning established by classical conditioning”). However, they also assumed that participants not only associate the rather unspecific connotation of these words with the CSs (i.e., a positive evaluation), but also the specific denotative meaning of the US words. Thus, they postulated an attribute conditioning effect, but never directly tested it.

Given the prominence of these two papers by Staats and Staats, it is surprising that the idea to condition specific attributes rather than global evaluations was not picked up for a rather long time. To the best of our knowledge, the first authors to investigate the notion that pairings might influence other attribute assessments were Kim, Allen, and Kardes (1996). They presented pairings of a white pizza-box with the logo “L Pizza House” (CS) with the picture of a race car (US). In comparison to a control group, in which the pizza-box was never systematically paired with the race car, participants rated the brand “L Pizza House” higher on the attribute “fast delivery” compared to the control group. A second experiment replicated this differential rating on another specific attribute: Participants observed the presentation of “L Brand Tissue” (CS) with pictures of kittens (US). These pairings also increased the ratings on the attribute “softness” compared to a control group in which the pictures were presented unsystematically. Yet, these data suffered from the discussed methodological problem: Participants also provided substantially higher general positive evaluations of the brands paired with a race car and kittens. Thus, the reported specific effects might be simple halo effects from the general positive assessment (i.e., an EC effect) on the relevant ratings of the brands (i.e., speed for pizza delivery; softness for a facial tissue), rather than a genuine assessment change due to the pairings.

The first systematic investigation of what we call AC was provided by De Houwer, Baeyens, Randell, Eelen, & Meersmans (2005); they showed that it is possible to change participants’ ratings of babies and Chinese Kanji by pairing these stimuli with pictures of clearly male and female babies. However, they only found the basic effect in four out of eight experiments on explicit ratings (i.e., “definitely a boy --- definitely a girl”) and only in one experiment in a semantic priming task.

Other researchers also found effects on specific attributes. For example, Olson, Kendrick, and Fazio (2009) tried to change participants’ assessment of various Pokémon’s speed and size. They argued that the evaluative dimension is particularly conducive to learning; however, attribute learning should be possible under certain circumstances. In a first experiment, they paired two Pokémon with pictures indicative of speed (e.g., a snail vs. a motorcycle) and size (e.g., an ant vs. an ocean-liner). However, they did not find any effects of these pairings on ratings of speed and size, while they found a standard EC effect for Pokémon paired with clearly positive and negative USs. In their second experiment, they primed participants with words related to size prior to the pairings; in an ostensibly unrelated task, participants performed a lexical decision task for 40 trials before the conditioning phase; before each word/non-word target, they were subliminally presented with one of 20 words related to size. The control condition presented non-word letter strings instead. They found a difference in size ratings of the two Pokémon when size was primed before conditioning, but not in the control group for which size was not primed as a concept.

Finally, Landau and Leed (2012) changed participants' "fame" ratings in two experiments by pairing pictures of non-famous faces with famous faces, although the authors did not frame these pairings as conditioning, but as a memory paradigm and did not control for general evaluations of the non-famous faces.

This short literature overview may explain why AC effects, in comparison to EC effects, have garnered less research interest despite their potential high relevance for applications and theories of EC: Changes in assessments of specific attributes (i.e., AC effects) seem to be smaller than typical EC effects (De Houwer et al., 2005; Kim et al., 1996), they seem to depend on specific circumstances (Olson et al., 2009), and researchers face the methodological challenge of distinguishing them from halo effects based on EC (e.g., Landau & Leed, 2012).[1]

In the following, however, we will summarize empirical evidence showing that given the same methodological scrutiny as EC, AC is also a highly reliable and stable effect. Further, the specific circumstances are easily created and often inherent in the pairing paradigms. And finally, the methodological challenges can be solved statistically and experimentally; that is, we will argue that AC effects do not depend on halo effects or inferences from EC effects.

Addressing the Methodological Challenges of Attribute Conditioning

As discussed, most if not all person attributes have an evaluative connotation and one must therefore tease apart EC and AC effects. To separate EC from AC effects, Förderer and Unkelbach (2011) tried to change participants' assessment in four experiments of initially neutral male faces (Exp. 1-3) as well as geometric shapes and non-words (Exp. 4) on the dimension of athleticism. They paired these CSs with pictures that showed athletic persons (e.g., a man playing beach volleyball) or pictures of people doing non-athletic activities (e.g., a man typing on a computer). To account for halo influences from EC effects, both types of USs were selected to be equally likeable, based on pre-ratings. In addition, the experiments collected both likeability and athleticism ratings. All experiments also included

[1] Please note that we omit a paper by Gawronski and Mitchell (2014), who simultaneously changed participants' assessment of valence and arousal. Arousal, however, is not a specific attribute in the present sense of a stimulus attribute.

indirect measures: a semantic version of the affect misattribution procedure (AMP) by Payne, Cheng, Govorun, and Stewart (2005) and a categorical priming task. The experiments found changes in participants' athleticism assessment of the initially neutral faces due to the pairings on direct ratings, while likeability remained largely unaffected. Further, these effects remained stable when statistically controlling for participants' idiosyncratic likeability ratings of the CSs, and they were also found on the categorization frequencies in the AMP adaption (Exp. 1, 2, and 4) and response latency differences in the categorical priming task (Exp. 3). Thus, these experiments established changes in a specific attribute beyond changes in general evaluations.

This evidence rested on the statistical control of EC effects. A more experimental approach to establish AC beyond changes in evaluation was provided by Förderer and Unkelbach (2014). First, they showed differential AC effects for a variety of attributes (i.e., humorous, sexy, educated, athletic, and soft) with only evaluatively positive USs. Second, they used celebrities as USs with more than one clear attribute (e.g., people judge Angelina Jolie high on the attributes "attractive" and "familial"). Similar to Olson et al. (2009), they primed one attribute prior to the pairings and found that only the primed attribute assessment changed. Thus, given the same US, experimentally changing the accessibility of a given attribute changed CS ratings. If AC effects resulted from a general positive assessment due to the CS being paired with a positive US (i.e., a halo effect), *all* positive attributes and not only specific attributes should change, and a CS paired with Angelina Jolie should become similarly attractive and familial. However, contradicting a halo effect explanation, the data showed highly specific effects on the respective attributes.

AC effects are also not restricted to personality traits. Glaser and Walther (2013) successfully changed participants' assessment of initially neutral stimuli on the dimensions of "size" and "softness". They also collected likeability ratings, as "bigger" might imply "better", and "softer" is potentially more positive than "harder". Similar to Förderer and Unkelbach (2011), however, they found independent contributions of the pairings to participants' likeability, size (Exp. 1), and softness ratings (Exp. 2).

Given this evidence, pairing a CS with a US seems to influence specific CS attributes, beyond a general shift in evaluations. On this functional level, one may define AC then as people's changed assessment of initially neutral stimuli's (CSs) attributes due to repeated pairing with stimuli possessing these attributes above and beyond changed evaluations. In the following, we discuss data that inform a potential *mental* model of how these changes come about (see De Houwer, 2011), based on mental accounts of EC effects.

Mental Processes in EC and Their Explanatory Potential for AC

Again, on a functional level, AC and EC are highly similar. The question is how this similarity extends to the assumed mental processes. For EC effects, Hofmann et al. (2010) meta-analytically investigated mental process accounts. Thus, we discuss AC effects with regards to the accounts presented by Hofmann et al. (2010) to explain EC effects on a mental level, namely stimulus-response learning (S-R), referential or stimulus-stimulus learning (S-S), propositional learning, conceptual categorization, implicit misattribution, and holistic representation formation.

AC as S-R Learning?

Starting with Staats and Staats (1957), and continuing within consumer research (Stuart, Shimp, & Engle, 1987), attribute learning may be a form of classical conditioning; that is, participants have *responses* to the US and because the CS signals the US (see Kim et al., 1996), the CS elicits similar *responses*. If this kind of stimulus-response (S-R) learning underlies AC effects, AC should follow predictions from the Rescorla-Wagner theory of classical conditioning (Rescorla & Wagner, 1972). In particular, AC effects should show *blocking* and *extinction*. Kamin (1969) described the blocking phenomenon: If a specific CS (i.e., CS 1) has been paired with a US, and then a second CS (i.e., CS 2) is paired in combination with the first CS (i.e., CS 1 together with CS 2) with the same US, CS 2 will be blocked; that is, its assessment should not change. The reason is that CS 1 already serves as a signal for the US, and the CS 2 is redundant and therefore not learned as a signal for the US. Similarly, AC effects should show extinction; that is, when the CS occurs frequently without the US, it should not elicit the conditioned response any longer, as it does not reliably signal the US any longer.

Förderer and Unkelbach (2015) tested both blocking (Exp. 2 and 3) and extinction (Exp. 4 and 5) of AC effects, and found no evidence for these central aspects of S-R learning. First, AC effects were insensitive to blocking: Participants' assessments of a second CS, paired with a US in a compound with a previously paired first CS, did not significantly differ from the first CS. Both CS 1 and CS 2 showed changes on direct ratings and the adapted AMP measure. Second, AC effects were insensitive to extinction: After eight pairings of CSs with USs, half of the CSs were also presented alone for another eight trials. The observed AC effects did not differ between CSs that were presented with a US (i.e., pairings) and those that were presented again alone. Again, these effects were present both on explicit ratings and the adapted AMP measure. Given these data, it seems unlikely that AC effects are based on a type of signal or expectancy learning in accordance with the theory by Rescorla and Wagner (1972).

AC as S-S Learning?

CS-US pairings might link or associate the mental CS representation with the US representation, a possibility Baeyens, Eelen, Crombez, and Van den Bergh (1992) have termed "referential learning", or stimulus-stimulus (S-S) learning. Thus, because a neutral face is linked with an athletic/non-athletic target, the assessment of the CS's athleticism might change because the representation of the US is also activated at the time of assessment. This S-S link has two variants. First, the CS may be directly linked with the mental representation of the attribute; and thus, when presented alone, also activates the attribute and thereby influences the attribute assessment. Second, the CS may be linked with the mental representation of the US, which includes the attribute.

Förderer and Unkelbach (2016) tested predictions from the referential learning account in three experiments using a US revaluation paradigm (Mackintosh, 1983; Walther, Gawronski, Blank, & Langer, 2009). Revaluation implies that *after* the CS-US pairings, USs attributes change; for example, after pairing an athletic US with a neutral CS, the US is shown to lose its athletic prowess and becomes unathletic. First, if CS-US pairings link CS-attribute representations, then changing US attributes should not influence CS assessments. However, if CS-US pairings link the CS and the US, then changing US attributes should change CS assessments.

Concretely, Förderer and Unkelbach (2016) paired pencil drawings of athletic and unathletic people (US) with pictures of male faces (CS). After the pairings, participants observed changes of the US attributes: Athletic USs became corpulent and unathletic USs became muscular. Before these US revaluations, participants showed standard AC effects on direct ratings and the adapted AMP; after the revaluation, these AC effects vanished or even fully reversed. Even more, after showing the US to possess a fully novel attribute (i.e., musicality), the CS assessment changed on the attribute "musical" without further pairings. However, this change was only significant for explicit ratings, but not for an AMP adapted to assess musicality. These data are in line with referential learning that links CS and US, rather than a linkage between CS and the specific attribute.

AC as Propositional Learning?

Propositional learning is an alternative to referential learning (see Mitchell, De Houwer, & Lovibond, 2009). Accordingly, evaluations and assessments change due to the propositions about the CS and US elicited by the pairings. As propositional learning may in principle account for effects of referential learning, it is a viable alternative. A way to distinguish between referential and propositional learning are relational qualifiers of the CS-US pairings (e.g., Fiedler & Unkelbach, 2011; Moran & Bar-Anan, 2013). Given propositional learning, a link of the type "is similar to" should lead to *evaluation* shifts in the direction of the US, but a "is different from" link should lead to shifts away from the US. However, mere referential learning would always predict shifts in the direction of the US.

For AC, however, the situation is more complicated and depends on the specific propositional relation. For example, if a CS "dislikes" an athletic US, the CS should not become unathletic. There is no justification to judge the CS as unathletic based on this proposition without additional assumptions. On the other hand, if a CS "is different from" an athletic US, the CS *should* become unathletic. In an unpublished study, Högden, Hütter, & Unkelbach (2018) tested and found in three experiments that relational qualifiers indeed change the AC effect, but exactly for those qualifiers that should not lead to changes, namely when CSs "like" and "dislike" athletic and unathletic USs. Alternatively, mere referential links may also be excitatory or inhibitory (e.g., Holland & Sherwood, 2008; Unkelbach & Fiedler, 2016); it is thus at present unclear whether propositional learning may explain AC effects or mere referential learning with inhibitory links.

AC as Conceptual Recategorization?

Building on work by Field and Davey (1997, 1999), CS-US pairings might make features of the CS salient that were not salient without the presence of the US. Concretely, pairing an initially neutral face with a beach volley player might make the face's sharp cheeks salient, indicating a lean figure, leading to higher ratings on the attribute "athleticism". Conversely, pairing the same face with a person comfortably watching a movie on a couch might make the comfortable expression of the neutral face salient, indicating a relaxed state and lower ratings on the attribute "athleticism". However, the conceptual recategorization account cannot explain the revaluation effects observed by Förderer and Unkelbach, and it is thus unlikely that this account explains AC effects.

AC as Implicit Misattribution?

Jones, Fazio, and Olson (2009) proposed for EC effects that a US elicits a positive or negative reaction within participants, which is misattributed to the CS. This account is unlikely to account for AC effects for two reasons. First, implicit misattribution cannot explain US revaluation effects. Second, AC effects are not diminished for completely meaningless CS (e.g., a geometric shape) in comparison to a meaningful CS (e.g., a human face); both may be conditioned as "athletic" (Förderer & Unkelbach, 2012). However, the former is a much less likely source for an "athletic" reaction. Finally, misattribution should be a function of confusability between the CS and the US (Jones et al., 2009, p. 935). Accordingly, the more similar CS and US are, the stronger the predicted misattribution effects (Hofmann et al., 2010, p. 392). Förderer and Unkelbach (2015; Exp. 1) systematically varied CS and US similarity and found no effect of stimulus category similarities (e.g., pairing humans with human traits and foods with food traits vs. humans with food traits and food with human traits).

AC as a Holistic Representation?

Levey and Martin (1975) suggested that repeated pairings lead to an enduring holistic representation of CS and US. If the CS is presented after the pairings for evaluation, the US is also activated and CS evaluations shift in the direction of the US. This account is in line with AC's resistance to extinction (Förderer & Unkelbach, 2015), and may account for the observed effects so far. The only data points that would be incompatible with a holistic account explanation are the relational qualifiers by Högden and Unkelbach, which are so far unpublished.

Given the available data's presented (in-)compatibilities with existing accounts of EC, we may now present a potential mental model of AC.

A Model of AC Effects

To reiterate, on a functional level, AC is the change of CS attribute assessments due to pairings with a US possessing these attributes. On a process level, S-R learning, conceptual categorization, and implicit misattribution are unlikely candidates given the empirical situation. Rather, S-S learning, propositional learning, and holistic account formation may explain AC effects. Further, S-S learning and the holistic account are almost indistinguishable empirically. This is clearly apparent from Hofmann and colleagues' analysis (see Hofmann et al., 2010, Table 6, p. 412); across 15 coded moderators of EC effects, there is not a single diverging prediction. The following model is thus, in principle, compatible with S-S learning (and the holistic account) and propositional learning accounts; that is, at present, the model does not distinguish between those.

The model's central assumption is that AC is due to a stable referential link between CS and US that is created by CS-US pairings (see Figure 1a, top panel). The CS is thereby only indirectly linked to the conditioned attribute. The left part shows the US represented with three attributes (e.g., funny, athletic, competitive). The middle part shows a CS-US pairing; the middle part exemplifies a silent assumption, namely that across pairings, a specific US attribute (here: Attribute 2, for example: "athletic") becomes salient or accessible, which was done experimentally by Olson et al. (2009) and Förderer and Unkelbach (2014). That is, across the learning trials of USs high and low on this specific attribute, the specific attribute becomes salient. This assumption is also present in EC; that is, participants need a focus on valence (Gast & Rothermund, 2011). The right part then shows the judgment phase; a link is now established between CS and US due to the pairings, and presenting the CS alone now co-activates the US with its salient, but also with all other attributes, and thus changes the CS assessment. This model accounts for the effects of priming specific attributes (Förderer & Unkelbach, 2014; Olson et al., 2009), both at learning and at judgment. In particular, US attributes might change, become salient, or may be added or deleted at the assessment stage, which should then also influence CS assessments. Further, if one assumes that US valence is the most salient and accessible stimulus attribute, the present model is an EC model as well in terms of referential or propositional learning.

To be sure, Figure 1a is only one way to mentally represent a referential link; it is grounded in an associative model of memory (e.g., Rosch, 1975) with a spreading activation assumption (e.g., Anderson, 1983). However, associative spreading activation accounts of memory have substantial problems and many data sets do not comply with spreading activation predictions (e.g., Ratcliff & McKoon, 1981). Alternatively, one may think of the model in a non-associative way (see Figure 1b); here the referential link is depicted in a distributed memory model that does not rely on associations (McClelland, McNaughton, & O'Reilly, 1995). In Figure 1b, stimuli and their attributes are coded in pattern vectors of units. Thus, there is no single unit in the model that carries meaning, but meaning emerges from the pattern of the full vector. The sole additional assumption for this model to arrive at the same predictions as an associative model is that the pairings create a shared context for the CS and the US, which leads to a similarity in the stimuli's representations. Figure 1b's left panel shows the hypothetical representation of a US. For ease of presentation, we assume that the pattern of the top three units codes the US identity, and the pattern of the middle four units codes the US's attributes. The pattern of the bottom three units codes context. The middle panel then shows what happens during the pairings: both the CS and the US now share a common pattern based on the pairing context, which is functionally the same as establishing a CS-US link. The right panel shows the CS assessment stage. That is, memory is prompted with the CS (e.g., by showing the CS face), and due to the shared pattern from the context, the response from memory will include the pattern for the US attributes as well, although the response from memory will not be as strong as when prompted with the US.

Thus, Figures 1a and 1b illustrate two ways to integrate a referential link into two memory architectures (i.e., associative memory and distributed memory). The central point for both variants is that pairings link CS and US and the AC effect occurs due to this link.

The model, in both variants, makes a number of empirical predictions which we will summarize in the following. First, AC effects should show no dissociations between "explicit" and "implicit" measures. As there is only one stable referential link assumed,

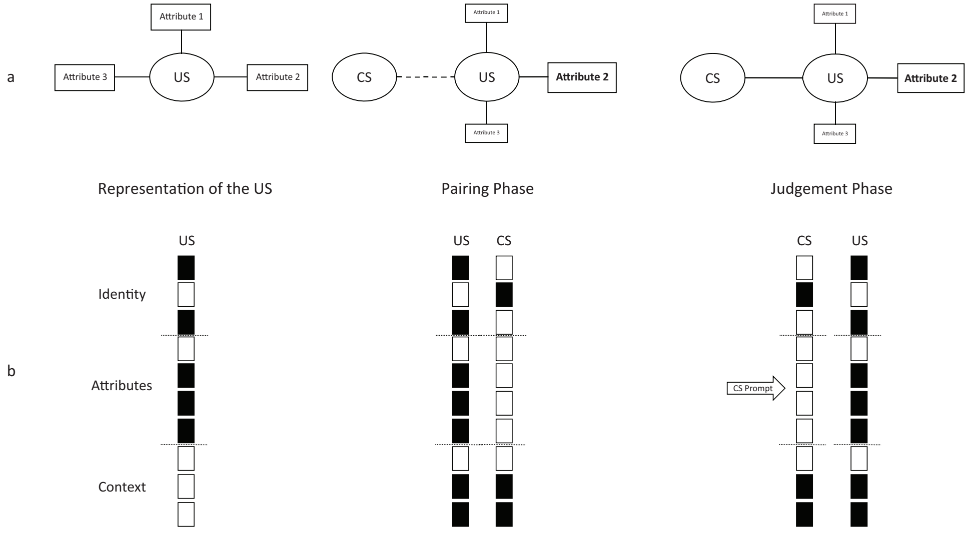

Figure 1. The top row shows an associative model of the AC effect. A given US (left side) is represented with three attributes. The middle panel shows the pairings. These pairings make a specific attribute salient (here: Attribute 2) and create a link between CS and US. During judgment, presenting the CS co-activates the US and thereby influences the assessment of CS attributes. The bottom row shows a distributed memory variant of the AC effect. Stimuli are represented by a vector of 10 units. The pattern of three assumed sub-vectors code stimulus identity, stimulus attributes, and context. The CS-US pairings create a shared pattern for the context, which is functionally equivalent to establishing a CS-US link. The right side shows the judgment phase. As the CS and US vectors now share a context pattern, presenting the CS (i.e., prompting memory with the CS) also returns information about the US attributes, which influences the CS assessment.

measures should converge when assessing the effect. Second, AC effects should be sensitive to US revaluations. As the assumed link is between CS and US, and not between CS and attribute, changing US representations should also change CS attribute assessments. Third, AC should be sensitive to awareness and attention manipulations as far as they influence the quality of the referential link representation (see Moors, 2016). Concretely, AC effects should be stronger for supraliminal in comparison to subliminal and AC effects should be stronger for undivided compared to divided attention. Finally, as AC effects should be insensitive to blocking and extinction, as the link is not based on a form of signal learning.

The model is presently silent with regards to the role of relational qualifiers; in other words, in its current form the precise mental representation of the referential link is not specified. It could be a "mere" referential link, which would imply US attribute main effects independent of the introduced relation. It could be an excitatory/inhibitory link, which would imply interactions of US attributes and relations on CS attribute assessments for all relations (e.g., CS is similar or dissimilar to US; CS likes or dislikes US; CS starts or stop US). Or it could be a propositional link, which would imply US attribute main effects for some relations (e.g., CS likes or dislikes US), but interactions for others (e.g., CS is similar or dissimilar to US). Further empirical work will provide answers to this question (Högden et al., 2018).

The model is also silent about the operating conditions of AC, such as efficiency, goal dependence, or controllability (see Corneille & Stahl, 2018). As noted by Gawronski and Bodenhausen (2009), it is difficult to arrive at solid conclusions regarding underlying models based on their differential effects of operating condition. For example, one may not preclude "mere" referential links in comparison to propositional links just because subliminal stimulus presentation yielded no AC effect. Both mere referential and propositional links may require a certain level of awareness (see Moors, 2016). Thus, including model assumptions in terms of operating conditions does not seem a fruitful endeavor.

Potential AC-EC Model Relations

The previous section presented a mental model of AC. As the model derives from mental process models of EC, and AC and EC obviously share variance, it is an obvious question how EC and AC relate. This relation may *a priori* take three forms. First, AC and EC may be orthogonal. Second, AC and EC may be concurrent. And third, AC and EC may be identical. To illustrate these forms, let us take the concrete example of an initially neutral target who is paired with a likeable (i.e., positive) or unlikeable (i.e., negative) US with either the attribute "athletic" (slightly positive) or "unathletic" (slightly negative).

Orthogonal Relation

On a functional level, orthogonality implies that participants' responses realize all four possibilities: the target may be become likeable and athletic, likeable and unathletic, unlikeable and athletic, and unlikeable and unathletic. As attributes are typically not valencefree, two of the possibilities are valence-congruent (i.e., likeable and athletic, unlikeable and unathletic) and should show slightly stronger effects on the DVs due to mutual halo effects. On a mental process level, an orthogonal AC-EC relation suggests (although does not necessitate) different processes; for example, EC effects might be due to implicit misattribution, while AC effects are due to the referential CS-US link established here. Strong evidence for the different processes would require additional experimental work depending on the assumed processes; for example, one may employ US revaluation experiments in which the US becomes (un)likeable or (un)athletic after the pairings; if CS assessment and evaluation changes are due to different processes (e.g., referential learning and misattribution, respectively), the post-US revaluation scores should show differential patterns for athleticism assessments and likeability evaluations.

On a theoretical level, orthogonal AC-EC effects with a process dissociation would imply that on a mental level, learning is different for a stimulus' denotative meaning (i.e., what is it?) and connotative meaning (i.e., do I like it?).

Concurrent Relation

On a functional level, a concurrent relation predicts the same as orthogonality. Participants' responses should also realize all four combinations of likeability and athleticism. On a mental process level, a concurrent relation suggests the same processes, but different information within the process; for example, both EC effects might be due to referential CS-US links. In other words, valence would be another attribute associated with the stimulus. Again, to substantiate such a relation, additional experimental variations would be necessary. Different from the orthogonal relation, though, both athleticism assessments and likeability evaluations should be equally sensitive to the process-relevant manipulations (e.g., US revaluation).

On a theoretical level, concurrent AC-EC effects would imply that on a mental level, learning processes are similar for denotative meaning and connotative meaning, but the learned content is differentiated (here: likeability and athleticism).

Identity Relation

If AC and EC are the same, predictions on a functional level differ from the previous two relations. Participants' responses should not fully realize all four possibilities, but ratings should follow a linear trend from both evaluation and assessment being high (i.e., likeable and athletic) to both being low (i.e., unlikeable and unathletic). In other words, this would imply an additive (i.e., two main effects on both dependent variables) instead of an interaction pattern. On a process level, the process must be the same, independent of which process is factually assumed.

Finally, on a theoretical level, identity would imply that attribute assessments are based on evaluative learning or that evaluative learning is based on attribute assessments. People would derive denotative meaning from the stimulus connotation or derive connotation from the stimulus denotation. These possibilities could be separated empirically by statistical means (i.e., regression analyses) or experimentally by manipulating goals, needs, or other motivational states. Connotative meaning is sensitive to the latter, denotative meaning is not. Thus, if connotation is based on denotation, the hungry or satiated participant should assess a bowl of porridge as calorie-high, but the hungry participant should like it better. If denotation is based on connotation, the hungry participant should like it better, but also assess the calorie-content higher.

Summary

At present, there is no data to clearly distinguish between these potential relations, although the orthogonal relation is most consistent with the available AC data (e.g., Förderer & Unkelbach, 2012) and results from other EC-related findings, such as Gawronski and Mitchell's (2014) independent conditioning of valence and arousal. In addition, it is tempting to map such an orthogonal relation onto other existing models, such as the associative-propositional model of attitude acquisition by Gawronski and Bodenhausen (2006), with AC and its denotative meaning mapping onto the propositional part, and EC and its connotative meaning mapping onto the associative part. However, as clearly stated above, there is no strong evidence with regards to the propositional or "mere associative" nature of the proposed link, and thus, we refrain from making this relation.

Answering the Special Issue's Central Questions for AC

Having presented our current model of AC effects, both in its associative or distributed variant, the empirical clarifications, and the potential AC-EC relations, we may now provide concrete answers to the questions that guide the organization of this special issue.

How Do CS-US Pairings Change CS assessments?

As Figure 1 suggests, participants' responses change because the CS presentation at the time of assessment co-activates the US representation and, thereby, assessments change in the direction of the US's attributes. This co-activation happens due to an enduring referential link, which may be implemented in an associative memory architecture (Figure 1a) or in a distributed memory model (Figure 1b). At present, we make no assumptions about the nature of this link (e.g., "merely associative" vs. "propositional"). For both variants, neither an explicit retrieval process or an inferential process is assumed or necessary.

What Is Stored in Memory?

AC effects are due to referential links, either associative or distributive, between CS and US. To set this assumption apart from other models, we may clarify what is not stored. First, we do not assume that the CS changes with regards to its attributes. This would imply newly formed links between attributes and CSs or changes in the CSs' attribute pattern. Second, the present model does not assume response learning. This would imply that a CS does not "possess" the hypothetical attributes (as in Figure 1), but that the CS elicits a

PsychOpen GOLD

respective response in participants. Such a learning assumption is also testable by changing the pairing sequences (e.g., from forward to backward conditioning; see Kim et al., 1996), but AC effects have been remarkably insensitive to variations in the pairing procedures.

Assuming that CS representations *per se* remain unchanged also solves the problem of how almost every attribute may be conditioning to every stimulus; for example, how brands may become trustworthy. A brand is an abstract concept that cannot be "trustworthy" in the same way a human being may be trustworthy. Similarly, Förderer and Unkelbach (2011) conditioned athleticism to abstract shapes. Assuming that the CS presentation does not initially change allows that even abstract shapes may acquire valence (i.e., EC effects) and specific attributes (i.e., AC effects).

Are There Important Processes That Occur Over Time Between Initial Acquisition and Evaluative Response?

The model does not so far include any consolidation processes. Yet, it allows for testable predictions. If either the link or the context pattern from Figure 1 is lost (e.g., forgetting), AC effects should disappear. Thus, any influences that distort the mental representations of CS or US, either during encoding or during judgments, should also reduce AC effects. Similarly, one might assume that CS patterns change or new CS-attribute links are established over time when CS and US are frequently co-activated (e.g., when CSs are judged repeatedly). However, this process is not included in the present model and has not been tested empirically.

Does EC Depend on Automatic or Deliberate Processes During Acquisition?

The model does not distinguish between different processes in acquisition. The model, however, is in both variants clear that the determinant of the AC effect is the quality of the mental representation (see Moors, 2016). Representation quality is best implemented in the distributed version, as one might randomly change single units in the patterns to simulate noise or low representational quality (e.g., for short presentation times in priming tasks). In extreme cases, it follows that if the US attributes are not encoded or the CS-US link is not established, no AC effect will occur. The quality of the representation will depend on factors such as attention, depth of encoding, or strength of the CS-US relation. Factors such as presentation duration or number of learning trials should therefore influence AC effects.

Does Stimulus Cooccurrence Have Different Effects on Deliberate Versus Automatic Measurements?

As stated above, the model does not assume differential links or pathways from CS to US, and therefore, the model does not theoretically predict dissociations between direct or "explicit" (i.e., explicit ratings of attributes) or more indirect or "implicit" measures (e.g., AMPs or categorical priming tasks). For EC research, this differentiation is often made based on the influences of learning conditions, such as awareness (e.g., Högden et al., 2018; Sweldens, Corneille, & Yzerbyt, 2014) or attention (e.g., Blask, Walther, & Frings, 2017; Field & Moore, 2005; see Stahl & Corneille, 2018, for a general review of these operating conditions and their impact on evaluative learning). However, again, there needs to a be sufficient quality of presentation. That is, if participants are not aware of the pairings or do not pay attention to the pairings, no referential link will be established; quality of representation should thus influence “explicit” and “implicit” measures to the same degree (see Hu, Gawronski, & Balas 2017; for an example of such convergence in EC).

Empirically, to the best of our knowledge, there has also never been a dissociation reported between direct and indirect measures on AC effects. Observed differences between direct and indirect measures (e.g., Förderer & Unkelbach, 2016, Exp. 3) are most likely due to reliability issues of the employed measures or the quality of the mental representation, rather than differential processes; for example, a typical AMP trial only presents the to-be-assessed stimulus for only 100 ms, resulting in much more impoverished representation than the unlimited presentation of the target for an explicit rating (see Moors, 2016).

What Is the Role of Awareness?

Awareness has been discussed extensively as a central moderator for the acquisition of attitudes (e.g., Hütter, Sweldens, Stahl, Unkelbach, & Klauer, 2012; Stahl, Unkelbach, & Corneille, 2009; see Sweldens et al., 2014, for a review). Most of the time in EC research, awareness is called “contingency awareness”, albeit what is factually implemented as a measure is often a test of whether participants may remember which CS was paired with a which US (e.g., Walther & Nagengast, 2006). Contingency itself is rarely manipulated in social psychological learning research (see Kattner & Ellermeier, 2011, for an exception). The present model does not specify the role of awareness for changes in participants’ CS assessments. As long as the mental link between CS and US is established, independent of people’s ability to report that link in any form, AC effects should emerge.

Empirically, however, there is a substantial effect of awareness on AC effects. For example, De Houwer et al. (2005) found AC effects of pairing gender-neutral babies (Exp. 6) and Kanji (Exp. 7 and 8) with clearly male and female babies only for those CS for which participants could select the correct corresponding US out of six potential USs. Similarly, Förderer and Unkelbach (2011; Exp. 1-4) reported significant correlations between participants’ ability to correctly select the correct US and the size of the AC effects, both on direct and indirect measures. Thus, for the quality of representation to be sufficiently high, participants might need to be “aware” of the pairings and able to correctly remember them. On the empirical level, this predicts no AC effects for CS stimuli for which participants cannot report the paired US or the paired US attribute (see Stahl et al., 2009). This prediction is empirically and not theoretically derived, though.

Other forms of awareness, for example, the awareness of the experimenter’s hypotheses (i.e., demand effects), seem to play no role (Förderer & Unkelbach, 2016). The main empirical argument here is that the effect sizes for AC effects are comparable for direct and indirect measures. Demand effects should typically lead to stronger effects on “explicit” compared to “implicit” dependent variables, which is not the case empirically beyond effect size losses due to measurement unreliability (see also Olson et al., 2009).

How Do Relational Qualifiers Present at Acquisition Influence EC?

Relational qualifiers do specify the CS-US link. EC research has used, for example, verbal qualifiers such as "is a friend/enemy of" (Fiedler & Unkelbach, 2011) or "hates/loves" (Förderer & Unkelbach, 2012). But also functional relations such as a CS "stopping" or "starting" the US (e.g., an unpleasant noise; Moran, Bar-Anan, & Nosek, 2016) of functional foci on similarities and differences have been used (e.g., Corneille, Yzerbyt, Pleyers, & Mussweiler, 2009; Unkelbach & Fiedler, 2016).

As discussed above, relational qualifiers are therefore the main tool to distinguish between different forms of the proposed referential link. To repeat, a propositional link should distinguish between a US that "is different from" a CS (effect: AC reversal), and a US "who hates" a CS (effect: standard AC). An excitatory/inhibitory link should not distinguish between the former to relations (i.e., different/hates), but should distinguish between relations of "different/hates" (effect: AC reversal or null effect) and "similar/loves" (effect: standard AC). A "mere" referential link should show standard AV effects for all cases. Presently, the precise role of relational qualifiers in EC is an open empirical question and not specified in the model.

Is AC Inevitable When Stimuli Cooccur?

The central point in Figure 1 is the formation of a CS-US link. Co-occurrence or pairings of stimuli are obviously events that foster the establishment of such mental links and might be the default option (see Fiedler & Unkelbach, 2011). However, AC is not inevitable as one might think of many factors that might hinder or prevent the establishment of the relevant link, such as lack of attention or awareness.

Does the Model Assume Specific Factors That Would Moderate AC?

A central moderator for the emergence of specific AC effects is the accessibility or salience of the respective attribute in the US representation. This is supported by the priming results from Olson et al. (2009) and from Förderer and Unkelbach (2014). Thus, central moderators besides priming the to-be-conditioned attribute are factors that make a specific attribute more or less salient or accessible. This is the step in Figure 1 from the left to the middle panels. By selecting USs from both ends of the relevant dimensions (e.g., "likeable-unlikeable"; "healthy-unhealthy"), attribute salience typically emerges without further experimental influences. However, one may think of explicit instructions to focus on specific attributes or other ways to change attribute accessibility or attribute salience (Gast & Rothermund, 2011).

How Does Verbal Information About the CSUS Co-Occurrence (Instruction) Change CS Evaluation?

In EC research, there is evidence that symbolic experiences are as effective as factual experiences (e.g., Gast & De Houwer, 2012). For example, instructing participants that something has happened might be as effective as the factual experience that something has happened (e.g., Van Dessel, De Houwer, Gast, & Tucker-Smith, 2015). The present AC model also assumes no privileged status of factual vs. symbolic experiences. One might speculate that factual experiences lead to deeper encoding, but without such additional assumptions, instructions or symbolic information should be as effective as the actual pairing for AC effects to emerge, as the link might follow from direct or indirect/symbolic experiences.

Does the Model Predict Sensitivity to Statistical CSUS Contingency or CSUS Contiguity?

The differentiation between contingency and contiguity is most relevant for S-R conceptions of learning (i.e., when the CS is a signal for the US). As the model is based on S-S learning and the AC effects seem to be insensitive to blocking and extinction (Förderer & Unkelbach, 2015), true CS-US contingency is not necessary. Rather, mere CS-US contiguity (i.e., CS-US pairings) should suffice to establish AC effects.

Does the Model Predict That AC Would Be Sensitive to Later Presentations of CS or US Alone (i.e., Show Extinction)?

Similar to the previous question, the model is an S-S model and thus does not predict extinction. This has been shown empirically by Förderer and Unkelbach (2015; Exp. 4 and 5).

Does the Model Predict Any Individual Differences at Any Stage (Acquisition, Storage, Expression)?

Vogel, Hütter, and Gebauer (2017) showed that EC effects are stronger for people high in the personality factor agreeableness and neuroticism. The agreeableness effect was not predicted, while the neuroticism effect is grounded in Larsen and Diener's (1987) observation that people high in neuroticism are more sensitive to stimulus valence and should thus show stronger EC effects. A comparable hypothesis cannot be derived for AC and the model makes no predictions about personality or inter-individual differences.

Does the Model Specify (and How) Differences Between Types of Paired Stimuli? Are There Any Specific Categories of Stimuli That Are Assumed to Produce Different AC Outcomes?

In classical conditioning, some CSs serve as a better signal for a given US than others, an observation often explained as genetic preparedness (e.g., Seligman, 1970). The referential nature of the present model makes exactly the opposite prediction. The nature of the CS, the US, or the attribute should not matter. It should, in principle, be possible to link any CS with any US and thereby, any CS with any US attribute. For example, Förderer and Unkelbach (2015; Exp. 1) systematically varied CS and US similarity and found no effect of stimulus category similarities (e.g., pairing humans with human traits and foods with food traits vs. humans with food traits and food with human traits).

Are There Any Important Predictions, Not Mentioned Previously, That Would Be Central for Testing the Model?

In Lewin's work (e.g., Lewin, 1936), valence is a function of the needs and goals within the organism and the affordances of the environment, that is, the organism's life space. A restaurant might have positive valence for a hungry person, but not for a satiated person. An attractive man might be a potential mating rival for another heterosexual man and thus be evaluated negatively; for a homosexual man, the attractive man might be a positive stimulus as a potentially desirable partner. Thus, valence in its original conceptualization is the outcome of an interaction; as a consequence, EC effects should also strongly depend on such interactions between the organism and the environment.

Conversely, specific attributes should be far more independent from specific needs and goals. In a naive empiricistic view, one might assume that attributes exist (e.g., the size of an object; Glaser & Walther, 2013) without a person assessing the attribute. Concretely, people should by and large agree on a target's attractiveness, but may differ in how much they like the target. Thus, a central prediction of the model is the stability of AC effects across needs, goals, bodily states, moods, and personality.

Conclusions and the AC-EC Relation

The idea that specific attributes might be learned was already presented by Staats and Staats (1957); yet, there has been relatively little research compared to evaluative learning on the learning of specific attributes. We presented a mental processes model of such AC effects, which assumes an enduring referential link between CS and US. While key predictions of this model have been confirmed (e.g., sensitivity to US revaluation, insensitivity to blocking and extinction), especially the nature of the referential link (e.g., "mere" referential vs. propositional) requires further empirical work. In addition, the precise functional and process relations of AC and EC effects need further investigations. However, the sketched research program had the foremost goal to establish AC as a true phenomenon beyond halo effects from evaluative learning and present a testable model of AC effects. We believe this endeavor has so far been successful and specifying the model and its relation to EC further will inform AC research in particular, and theories of attitude formation in general.

Funding

The authors have no funding to report.

Competing Interests

The authors have declared that no competing interests exist.

Acknowledgements

The authors have no support to report.

References

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 261–295. https://doi.org/10.1016/S0022-5371(83)90201-3

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*(2), 133–142. https://doi.org/10.1016/0005-7967(92)90136-5

Blask, K., Walther, E., & Frings, C. (2017). When congruence breeds preference: The influence of selective attention processes on evaluative conditioning. *Cognition and Emotion*, *31*(6), 1127–1139. https://doi.org/10.1080/02699931.2016.1197100

Corneille, O., & Stahl, C. (2018). Associative attitude learning: A closer look at evidence and how it relates to attitude models. *Personality and Social Psychology Review.* Advance online publication. https://doi.org/10.1177/1088868318763261

Corneille, O., Yzerbyt, V. Y., Pleyers, G., & Mussweiler, T. (2009). Beyond awareness and resources: Evaluative conditioning may be sensitive to processing goals. *Journal of Experimental Social Psychology*, *45*(1), 279–282. https://doi.org/10.1016/j.jesp.2008.08.020

De Houwer, J. (2007). A conceptual and theoretical analysis of evaluative conditioning. *The Spanish Journal of Psychology*, *10*(2), 230–241. https://doi.org/10.1017/S1138741600006491

De Houwer, J. (2011). Why the cognitive approach in psychology would profit from a functional approach and vice versa. *Perspectives on Psychological Science*, *6*(2), 202–209. https://doi.org/10.1177/1745691611400238

De Houwer, J., Baeyens, F., Randell, T., Eelen, P., & Meersmans, T. (2005). Beyond evaluative conditioning? Searching for associative transfer of nonevaluative stimulus properties. *Cognition and Emotion*, *19*(2), 283–306. https://doi.org/10.1080/02699930441000328

Fiedler, K., & Unkelbach, C. (2011). Evaluative conditioning depends on higher order encoding processes. *Cognition and Emotion*, *25*(4), 639–656. https://doi.org/10.1080/02699931.2010.513497

Field, A. P., & Davey, G. C. L. (1997). Conceptual conditioning: Evidence for an artificial account of evaluative learning. *Learning and Motivation*, *28*(3), 446–464. https://doi.org/10.1006/lmot.1997.0980

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*(2), 211–224. https://doi.org/10.1037/0097-7403.25.2.211

Field, A. P., & Moore, A. C. (2005). Dissociating the effects of attention and contingency awareness on evaluative conditioning effects in the visual paradigm. *Cognition and Emotion*, *19*(2), 217–243. https://doi.org/10.1080/02699930441000292

Förderer, S., & Unkelbach, C. (2011). Beyond evaluative conditioning! Evidence for transfer of non-evaluative attributes. *Social Psychological & Personality Science*, *2*, 479–486.

Förderer, S., & Unkelbach, C. (2012). Hating the cute kitten or loving the aggressive pit-bull: EC effects depend on CS–US relations. *Cognition and Emotion*, *26*(3), 534–540. https://doi.org/10.1080/02699931.2011.588687

Förderer, S., & Unkelbach, C. (2014). The moderating role of attribute accessibility in conditioning multiple specific attributes. *European Journal of Social Psychology*, *44*(1), 69–81. https://doi.org/10.1002/ejsp.1994

Förderer, S., & Unkelbach, C. (2015). Attribute conditioning: Changing attribute-assessments through mere pairings. *Quarterly Journal of Experimental Psychology*, *68*(1), 144–164. https://doi.org/10.1080/17470218.2014.939667

Förderer, S., & Unkelbach, C. (2016). Changing US attributes after CS-US pairings changes CS-attribute-assessments: Evidence for CS-US associations in attribute conditioning. *Personality and Social Psychology Bulletin*, *42*(3), 350–365. https://doi.org/10.1177/0146167215626705

Gast, A., & De Houwer, J. (2012). Evaluative conditioning without directly experienced pairings of the conditioned and the unconditioned stimuli. *Quarterly Journal of Experimental Psychology*, *65*(9), 1657–1674. https://doi.org/10.1080/17470218.2012.665061

Gast, A., & Rothermund, K. (2011). What you see is what will change: Evaluative conditioning effects depend on a focus on valence. *Cognition and Emotion*, *25*(1), 89–110. https://doi.org/10.1080/02699931003696380

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and prepositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*(5), 692–731. https://doi.org/10.1037/0033-2909.132.5.692

Gawronski, B., & Bodenhausen, G. V. (2009). Operating principles versus operating conditions in the distinction between associative and propositional processes. *Behavioral and Brain Sciences*, *32*, 207–208.

Gawronski, B., & Mitchell, D. G. V. (2014). Simultaneous conditioning of valence and arousal. *Cognition and Emotion*, *28*(4), 577–595. https://doi.org/10.1080/02699931.2013.843506

Glaser, T., & Walther, E. (2013). Two in one conditioning? The role of valence in concept learning. *Learning and Motivation*, *44*(3), 174–183. https://doi.org/10.1016/j.lmot.2012.12.003

Gräf, M., & Unkelbach, C. (2016). Halo effects in trait assessment depend on information valence: Why being honest makes you industrious, but lying does not make you lazy. *Personality and Social Psychology Bulletin*, *42*(3), 290–310. https://doi.org/10.1177/0146167215627137

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin*, *136*(3), 390–421. https://doi.org/10.1037/a0018916

Högden, F., Hütter, M., & Unkelbach, C. (2018). Does evaluative conditioning depend on awareness? Evidence from a continuous flash suppression paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. https://doi.org/10.1037/xlm0000533

Holland, P. C., & Sherwood, A. (2008). Formation of excitatory and inhibitory associations between absent events. *Journal of Experimental Psychology: Animal Behavior Processes*, *34*(3), 324–335. https://doi.org/10.1037/0097-7403.34.3.324

Hu, X., Gawronski, B., & Balas, R. (2017). Propositional versus dual-process accounts of evaluative conditioning: I. The effects of co-occurrence and relational information on implicit and explicit evaluations. *Personality and Social Psychology Bulletin*, *43*(1), 17–32. https://doi.org/10.1177/0146167216673351

Hütter, M., Sweldens, S., Stahl, C., Unkelbach, C., & Klauer, K. C. (2012). Dissociating contingency awareness and conditioned attitudes: Evidence of contingency-unaware evaluative conditioning. *Journal of Experimental Psychology: General*, *141*(3), 539–557. https://doi.org/10.1037/a0026477

Jones, C. R., Fazio, R. H., & Olson, M. A. (2009). Implicit misattribution as a mechanism underlying evaluative conditioning. *Journal of Personality and Social Psychology*, *96*(5), 933–948. https://doi.org/10.1037/a0014747

Kamin, L. J. (1969). Selective association and conditioning. In N. J. Mackintosh & W. K. Honig (Eds.), *Fundamental issues in associative learning* (pp. 42–64). Halifax, NS, Canada: Dalhousie University Press.

Kattner, F., & Ellermeier, W. (2011). Does evaluative learning rely on the perception of contingency? Manipulating contingency and US density during evaluative conditioning. *Experimental Psychology*, *58*(5), 391–399. https://doi.org/10.1027/1618-3169/a000108

Kim, J., Allen, C. T., & Kardes, F. R. (1996). An investigation of the mediational mechanisms underlying attitudinal conditioning. *JMR, Journal of Marketing Research*, *33*(3), 318–328. https://doi.org/10.2307/3152128

Landau, J. D., & Leed, S. A. (2012). The illusion of fame: How the nonfamous become famous. *The American Journal of Psychology*, *125*(3), 351–360. https://doi.org/10.5406/amerjpsyc.125.3.0351

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International affective picture system (IAPS): Technical manual and affective ratings. Gainesville, FL, USA: NIMH Center for the Study of Emotion and Attention.

Larsen, R. J., & Diener, E. (1987). Affect intensity as an individual difference characteristic: A review. *Journal of Research in Personality*, *21*(1), 1–39. https://doi.org/10.1016/0092-6566(87)90023-7

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

Lewin, K. (1936). The present state of psychology. In *Principles of topological psychology* (F. Heider & G. M. Heider, Trans.) (pp. 3-7). New York, NY, USA: McGraw-Hill.

Mackintosh, N. J. (1983). *Conditioning and associative learning*. New York, NY, USA: Oxford University Press.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419–457. https://doi.org/10.1037/0033-295X.102.3.419

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, *32*(2), 183–198. https://doi.org/10.1017/S0140525X09000855

Moors, A. (2016). Automaticity: Componential, causal, and mechanistic explanations. *Annual Review of Psychology*, *67*(1), 263–287. https://doi.org/10.1146/annurev-psych-122414-033550

Moran, T., & Bar-Anan, Y. (2013). The effect of object–valence relations on automatic evaluation. *Cognition and Emotion*, *27*(4), 743–752. https://doi.org/10.1080/02699931.2012.732040

Moran, T., Bar-Anan, Y., & Nosek, B. A. (2016). The assimilative effect of co-occurrence on evaluation above and beyond the effect of relational qualifiers. *Social Cognition*, *34*(5), 435–461. https://doi.org/10.1521/soco.2016.34.5.435

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, *35*(4), 250–256. https://doi.org/10.1037/0022-3514.35.4.250

Olson, M., Kendrick, R., & Fazio, R. (2009). Implicit learning of evaluative vs. non-evaluative covariations: The role of dimension accessibility. *Journal of Experimental Social Psychology*, *45*(2), 398–403. https://doi.org/10.1016/j.jesp.2008.10.007

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, *89*(3), 277–293. https://doi.org/10.1037/0022-3514.89.3.277

Ratcliff, R., & McKoon, G. (1981). Does activation really spread? *Psychological Review*, *88*(5), 454–462. https://doi.org/10.1037/0033-295X.88.5.454

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

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*(3), 192–233. https://doi.org/10.1037/0096-3445.104.3.192

Seligman, M. E. (1970). On the generality of the laws of learning. *Psychological Review*, *77*(5), 406–418. https://doi.org/10.1037/h0029790

Staats, A. W., & Staats, C. K. (1958). Attitudes established by classical conditioning. *Journal of Abnormal and Psychology*, *57*(1), 37–40. https://doi.org/10.1037/h0042782

Staats, C. K., & Staats, A. W. (1957). Meaning established by classical conditioning. *Journal of Experimental Psychology*, *54*(1), 74–80. https://doi.org/10.1037/h0047716

Stahl, C., Unkelbach, C., & Corneille, O. (2009). On the respective contributions of awareness of unconditioned stimulus valence and unconditioned stimulus identity in attitude formation through evaluative conditioning. *Journal of Personality and Social Psychology*, *97*(3), 404–420. https://doi.org/10.1037/a0016196

Stuart, E. W., Shimp, T. A., & Engle, R. W. (1987). Classical conditioning of consumer attitudes: Four experiments in an advertising context. *The Journal of Consumer Research*, *14*(3), 334–349. https://doi.org/10.1086/209117

Sweldens, S., Corneille, O., & Yzerbyt, V. (2014). The role of awareness in attitude formation through evaluative conditioning. *Personality and Social Psychology Review*, *18*(2), 187–209. https://doi.org/10.1177/1088868314527832

Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology*, *4*(1), 25–29. https://doi.org/10.1037/h0071663

Unkelbach, C., & Fiedler, K. (2016). Contrastive CS-US relations reverse evaluative conditioning effects. *Social Cognition*, *34*(5), 413–434. https://doi.org/10.1521/soco.2016.34.5.413

Van Dessel, P., De Houwer, J., Gast, A., & Tucker-Smith, C. (2015). Instruction-based approach-avoidance effects: Changing stimulus evaluation via the mere instruction to approach or avoid stimuli. *Experimental Psychology*, *62*(3), 161–169. https://doi.org/10.1027/1618-3169/a000282

Vogel, T., Hütter, M., & Gebauer, J. E. (2017). Is evaluative conditioning moderated by Big Five personality traits? *Social Psychological and Personality Science*. https://doi.org/10.1177/1948550617740193

Walther, E., & Nagengast, B. (2006). Evaluative conditioning and the awareness issue: Assessing contingency awareness with the four-picture recognition test. *Journal of Experimental Psychology: Animal Behavior Processes*, *32*(4), 454–459. https://doi.org/10.1037/0097-7403.32.4.454

Walther, E., Gawronski, B., Blank, H., & Langer, T. (2009). Changing likes and dislikes through the backdoor: The US-revaluation effect. *Cognition and Emotion*, *23*(5), 889–917. https://doi.org/10.1080/02699930802212423

PsychOpen GOLD is a publishing service by
Leibniz Institute for Psychology Information (ZPID),
Trier, Germany. https://www.leibniz-psychology.org