Review Article

PsychOpen GOLD

Social Psychological Bulletin
Psychologia Społeczna

Evaluative Conditioning From the Perspective of the Associative-Propositional Evaluation Model

Bertram Gawronski[1], Galen V. Bodenhausen[2]

[1] *University of Texas at Austin, Austin, TX, USA*
[2] *Northwestern University, Evanston, IL, USA*

Corresponding author: Bertram Gawronski (University of Texas at Austin, Austin, TX, USA).
E-mail: gawronski@utexas.edu)

Handling editor: Yoav Bar-Anan (Ben Gurion University, Beer-Sheva, Israel)

Received: 11 December 2017 • **Accepted:** 18 June 2018 • **Published:** 20 September 2018

Citation: Gawronski, B., & Bodenhausen, G. V. (2018). Evaluative conditioning from the perspective of the associative-propositional evaluation model. *Social Psychological Bulletin*, *13*(3), Article e28024. https://doi.org/10.5964/spb.v13i3.28024

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

Abstract

Evaluative conditioning (EC) is defined as the change in the evaluation of a conditioned stimulus (CS) due to its pairing with a positive or negative unconditioned stimulus (US). According to the associative-propositional evaluation (APE) model, EC effects can be the result of two functionally distinct learning mechanisms: associative and propositional learning. The current article reviews the core assumptions of the APE model regarding (1) the defining features of associative and propositional learning, (2) the mental representations resulting from the two learning mechanisms, (3) the processes involved in the behavioral expression of these representations, and (4) the automatic versus controlled nature of the processes underlying EC effects. In addition to reviewing the core assumptions of the APE model, the article reviews relevant evidence to illustrate the theory's main hypotheses, its explanatory and predictive power, as well as empirical challenges for the theory.

Keywords

associative learning, automaticity, dual-process theory, evaluative conditioning, propositional learning

When an object is repeatedly encountered together with a pleasant or unpleasant stimulus, the object tends to acquire the valence of the co-occurring stimulus. For example, a consumer product that is repeatedly paired with pleasant stimuli (e.g., elevating music, a liked celebrity) may be evaluated more positively as a result of the pairings (e.g., Gibson, 2008). Conversely, some negative political campaigns try to undermine voters' support for political opponents by subtly pairing them with unpleasant stimuli (e.g., Weinberger & Westen, 2008). In psychology, the general phenomenon underlying these influences is known as *evaluative conditioning* (EC), which is defined as the change in the evaluation of a conditioned stimulus (CS) due its pairing with a positive or negative unconditioned stimulus (US) (De Houwer, 2007).

In the early years of research on EC, observed changes in CS evaluations have been predominantly explained in terms of associative learning processes, involving the automatic formation of mental associations between co-occurring stimuli (for an overview, see Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010). More recently, however, associative explanations have been challenged by propositional accounts that reject the idea of automatic association formation (e.g., De Houwer, 2009, 2014; Mitchell, De Houwer, & Lovibond, 2009). According to these accounts, EC effects result from the non-automatic generation and truth assessment of mental propositions about the relation between co-occurring stimuli.

Different from an exclusive focus on either associative or propositional learning, the associative-propositional evaluation (APE) model (Gawronski & Bodenhausen, 2006a, 2006b, 2007, 2011, 2014) acknowledges the contribution of both learning mechanisms to EC effects. In the current article, we review the core assumptions of the APE model regarding (1) the defining features of associative and propositional learning, (2) the mental representations resulting from the two learning mechanisms, (3) the processes involved in the behavioral expression of these representations, and (4) the automatic versus controlled nature of the processes underlying EC effects. Although we focus primarily on theoretical assumptions of the APE model, we also review relevant evidence to illustrate the theory's main hypotheses, its explanatory and predictive power, and empirical challenges for the theory.

Associative and Propositional Learning

A central assumption of the APE model is that observed co-occurrences of a CS and a US can influence mental representations via two functionally distinct learning mechanisms (Gawronski & Bodenhausen, 2006a, 2011, 2014). The first mechanism, described as *associative learning*, involves the formation of mental links on the basis of observed spatiotemporal contiguities between objects and events. Resonating with the Hebbian principle of *fire together, wire together*, this learning mechanism is assumed to capture observed regularities in the environment by creating direct mental links between simultaneously activated concepts: "The general idea is an old one, that any two cells or systems of cells that are repeatedly active at the same time will tend to become associated, so that activity in one facilitates activity in the other" (Hebb, 1949, p. 70). Thus, when a CS repeatedly co-occurs with a positive or negative US, the notion of associative learning suggests that the repeated co-activation of their corresponding mental representations creates an associative link between the two.

The second mechanism, described as *propositional learning*, involves the formation of mental representations on the basis of newly acquired information about states of affairs. Two essential features that distinguish propositional learning from associative learning are that propositional learning is sensitive to (1) the particular relation between co-occurring stimuli and (2) the perceived validity of the observed relation. First, whereas representations resulting from associative learning reflect the mere co-occurrence of stimuli regardless of their relation (e.g., mental link between the concepts *Aspirin* and *headaches*), representations resulting from propositional learning capture the specific relation between co-occurring stimuli (e.g., representation of the propositional relation *Aspirin reduces headaches*). Second, whereas associative learning is based on observed co-occurrences regardless of their perceived validity (e.g., mental link between the concepts *smoking* and *cancer*), propositional learning depends on the perceived validity of the encoded relation between stimuli (e.g., denial of the propositional relation *smoking causes cancer*). Thus, although propositional learning can go far beyond mere co-occurrences, it can contribute to EC effects when observed co-occurrences between a CS and a US lead to propositional inferences about evaluative characteristics of the CS (e.g., negative response to a sound that has been paired with electric shocks arising from the propositional inference such as *the sound is always followed by an unpleasant experience*).

An important assumption of the APE model is that the two learning mechanisms are functionally independent, which allows them to operate simultaneously (Gawronski & Bodenhausen, 2006a, 2011). According to the APE model, repeated pairings of a CS and a US can influence mental representations via associative learning, propositional learning, or both. Thus, merely observing a change in the evaluation of a CS as a result of CS-US pairings does not provide any information about whether this change was driven by associative learning, propositional learning, or a joint operation of the two (Gawronski & Bodenhausen, 2011). Although the two learning processes commonly converge, evidence for their independent contributions can be established by experimental manipulations of factors that produce diverging effects on each process, based on their defining characteristics. A contribution of associative learning would be supported by evidence for a mental representation of the CS that reflects the valence of the US even when (1) the CS and the US have a relation suggesting an evaluation of the CS that is opposite to the valence of the US or (2) the evaluation implied by the observed co-occurrence (or relation) is rejected as invalid. Conversely, a contribution of propositional learning would be supported by evidence for a mental representation of the CS that (1) reflects the valence implied by the perceived relation between the CS and the US, even when this relation suggests an evaluation of the CS that is opposite to the valence of the US and (2) is sensitive to the perceived validity of this relation. Although effects of associative and propositional learning may cancel each other out in these cases (e.g., a negative evaluation resulting from a mental link between *Aspirin* and *headaches* counteracting a positive evaluation resulting from the propositional information *Aspirin reduces headaches*), the APE model suggests that the two learning mechanisms can be identified by means of asymmetrical effects on spontaneous and deliberate evaluations. We discuss this hypothesis in more detail when we explain the role of associative and propositional processes in the behavioral expression of mental representations.

In sum, the APE model assumes that EC effects can be the result of associative learning, propositional learning, or a joint operation of the two. Whereas associative learning involves the formation of mental links on the basis of observed spatio-temporal contiguities between objects and events, propositional learning involves the formation of mental representations on the basis of newly acquired information about states of affairs. The central features that distinguish the two kinds of learning mechanisms are that (1) associative learning registers mere co-occurrences, whereas propositional learning is sensitive to the relation between co-occurring stimuli, and (2) associative learning registers new information regardless of its perceived validity, whereas propositional learning depends on the perceived validity of new information.

Mental Representation

Different from the distinction between associative and propositional *processes*, the APE model is sometimes mischaracterized as postulating two distinct *representations* in memory: an associative representation involving mental links between concepts and a propositional representation of subjective beliefs about states of affairs. Although some theories propose two functionally distinct memory systems (e.g., Rydell & McConnell, 2006; Smith & DeCoster, 2000), the APE model assumes that all information is stored in a single associative network (Gawronski & Bodenhausen, 2006a, 2006b, 2011; Gawronski, Brannon, & Bodenhausen, 2017). According to the APE model, every mental proposition is based on patterns of activated associations and, conversely, any pattern of activated associations gives rise to propositional beliefs about states of affairs. From this perspective, it makes little sense to distinguish between *associations* and *propositions* as two distinct forms of mental representation (Gawronski & Bodenhausen, 2006b; Gawronski et al., 2017). In line with this concern, the APE model is more accurately described as a dual-*process* theory in the sense that it distinguishes between two functionally distinct processes by which mental representations can be formed and two functionally distinct processes in the behavioral expression of mental representations. However, it is not a dual-*representation* theory, because it explicitly rejects the idea of two distinct kinds of mental representations (see Gawronski, Sherman, & Trope, 2014).

An important question in this context is how the APE model accounts for the fact that people can have stored knowledge with relational content. For example, most people do not simply associate Aspirin with headaches (i.e., *Aspirin-headaches*); they also have stored knowledge of the relation between the two (i.e., *Aspirin reduces headaches*). Although some EC theorists have rejected associative accounts on the basis of this concern (e.g., De Houwer, 2009, 2014), we argue that this rejection is based on two misunderstandings. First, the distinction between associative and propositional learning refers to two kinds of learning mechanisms, not two kinds of mental representations. Although it is theoretically possible that each learning mechanism is linked to a distinct memory system (e.g., Rydell & McConnell, 2006; Smith & DeCoster, 2000), the distinction between associative and propositional learning does not require two distinct memory systems if one assumes that the outcomes of either learning mechanism are stored in the same associative network (e.g., Gawronski & Bodenhausen, 2006a, Strack & Deutsch, 2004). Second, multi-layer connectionist models involving both excitatory and inhibitory links are perfectly able to represent complex relations between objects and events. Such models typically include a hierarchical structure, in that activated concepts at higher levels specify the relation between activated concepts at lower levels (e.g., Doumas, Hummel, & Sandhofer, 2008; McClelland, McNaughton, & O'Reilly, 1995).[1] For example, in a connectionist model capturing the propositional relation *Aspirin reduces headaches*, the individual concepts *Aspirin* and *headaches* would be represented at a lower level of a multi-layer network, and their relation would be represented via relational concepts at a higher level specifying the relation between the concepts *Aspirin* and *headaches* via excitatory and inhibitory links (e.g., specifying that Aspirin relieves rather than causes headaches).

From the perspective of the APE model, mental representations of this kind qualify as *associative*, because they are based on associative links between nodes and the principle of spreading activation. Of course, such mental representations could also be described as *propositional*, because they capture relational information (e.g., De Houwer, 2009). However, as we explained elsewhere (Gawronski et al., 2017), we deem this disagreement a matter of semantics rather than conflicting theoretical assumptions. In our view, the central question is not whether associative networks are capable of representing complex relations between stimuli (they are!).[2] Rather, the two critical questions are: (1) Do observed co-occurrences create unqualified associative links between the co-occurring stimuli irrespective of their relation (e.g., does exposure to the information *Aspirin relieves headaches* produce a direct associative link between *Aspirin* and *headaches* at a lower level despite the representation of their relation at a higher level)? (2) Does newly acquired information

---

[1] Although connectionist models of relational information can be symbolic (e.g., Doumas et al., 2008) or subsymbolic (e.g., McClelland et al., 1995), our argument regarding the representation of relational information in multi-layer networks applies to connectionist models of both kinds.

[2] Critics may argue that existing connectionist models are unable to capture the complexities of relational processing (Jan De Houwer, personal communication, April 30, 2018). In response to this objection, we deem it important to distinguish between (1) limitations of extant theories in accounting for specific empirical findings in an *a posteriori* manner and (2) basic features of connectionist models that make them unable to account for particular findings in an *a priori* manner. Although existing connectionist models may have empirical limitations, these limitations do not preclude the possibility that they can be addressed in refined models. In our view, arguments about a priori deficits of connectionist models are as implausible as arguments that the human mind is too complex to be captured by scientific theories.

create associative links between co-occurring stimuli irrespective of the perceived validity of that information (e.g., does exposure to the information *Aspirin causes cancer* produce a direct associative link between *Aspirin* and *cancer* even when this information is rejected as false)? These questions do not pertain to the status of associations and propositions as distinct knowledge structures in memory, but to the involvement of associative learning in the formation of mental representations over and above the uncontroversial role of propositional learning.

Another important question in this context concerns the specific components of unqualified links resulting from associative learning. On the one hand, it is possible that associative learning involves the formation of mental links between the CS and the US (i.e., stimulus-stimulus or S-S learning). On the other hand, associative learning may create a mental link between the CS and the evaluative response elicited by the US (i.e., stimulus-response or S-R learning). Although the APE model initially focused exclusively on cases of S-S learning (Gawronski & Bodenhausen, 2006a, 2011), the Hebbian principle of *fire together, wire together* can also accommodate instances of S-R learning. On the one hand, S-S learning can be assumed to occur when a CS is repeatedly paired with the same US. In this case, co-activation of the CS and the US should create a mental link between the two, such that future activation of the CS will spread to the US and, thus, elicit an evaluative response to the CS in line with the valence of the US. On the other hand, S-R learning can be assumed to occur when a CS is repeatedly paired with different USs of the same valence. In this case, co-activation of the CS and a particular evaluative response should create a mental link between the two, such that future activation of the CS will spread to the associated evaluative response. An important difference between the two cases is that changes in the valence of the US that occur after encoding of CS-US pairings should lead to corresponding changes in CS evaluation for S-S learning, but not S-R learning (e.g., Sweldens, Van Osselaer, & Janiszewski, 2010; Walther, Gawronski, Blank, & Langer, 2009).

Similar to the distinction between S-S and S-R representations in associative learning, propositional learning may produce either episodic representations of specific events or abstract representations of the inferred valence of a CS. For example, when a CS repeatedly co-occurs with a US, propositional learning may create a multi-layer representation that captures the observed co-occurrence (e.g., CS and US represented at a lower level of a multi-layer network and their co-occurrence being represented via relational concepts at a higher level). Yet, perceivers may also draw propositional inferences about evaluative characteristics of the CS, leading to the formation of a multi-layer representation that captures the inferred valence of the CS (e.g., CS and inferred valence represented at a lower level of a multi-layer network and their descriptive relation being represented via relational concepts at a higher level). Finally, perceivers may draw propositional inferences about more complex relations between the CS and the US, leading to the formation of multi-layer representations that capture the outcome of these inferences (e.g., CS and US represented at a lower level of a multi-layer network and their causal relation being represented via relational concepts at a higher level). Although the APE model does not include specific hypotheses about the conditions under which each of these inferences occurs, the theory suggests that the outcomes of these inferences are stored in multi-layer associative networks in which relational concepts at higher levels specify the relation between stimulus concepts at lower levels. Importantly, the notion of associative learning suggests that the concepts representing two co-occurring stimuli can have a direct associative link (e.g., direct associative link between the concepts *Aspirin* and *headache*) even when relational concepts at a higher level of the associative network suggest a more specific relation between the two (e.g., *Aspirin reduces headaches*).

In sum, the APE model rejects the distinction between *associations* and *propositions* as distinct forms of mental representation, suggesting that all information is stored in a single associative network. Relational information is assumed to be stored in a hierarchical manner in multi-layer networks, in that relational concepts at higher levels of the network specify the relation between stimulus concepts at lower levels. Thus, from the perspective of the APE model, the two central questions regarding the hypothesized role of associative learning over and above the uncontroversial role of propositional learning are: (1) Do observed co-occurrences create unqualified associative links between the co-occurring stimuli irrespective of their relation? (2) Does newly acquired information create associative links between co-occurring stimuli irrespective of the perceived validity of that information?

Behavioral Expression

A central assumption of the APE model is that stimuli elicit a positive or negative response to the extent that activation of their mental representation leads to a co-activation of evaluative concepts in memory (Gawronski & Bodenhausen, 2006a, 2011, 2014). Specifically, the theory states that principles of similarity matching determine the activation of mental concepts that represent the encountered stimulus, which can spread to other concepts that are associatively linked with the stimulus. To the extent that the associated concepts have a positive or negative valence, their activation is assumed to produce a spontaneous evaluative response that is in line with the valence of these concepts. For example, activation of a CS may spread to a US via a direct associative link between the CS and the US, thereby eliciting a spontaneous evaluative response to the CS that is in line with the valence of the US. According to the APE model, a straightforward method to capture spontaneous evaluations is the use of implicit measures, such as affective priming or the implicit association test (for a review, see Gawronski & De Houwer, 2014).

Similar to the defining characteristics of associative learning, a central feature of *associative activation* is that it is independent of subjective truth or falsity. Specifically, the APE model proposes that the principles of similarity matching determine the activation of associated concepts regardless of whether the activated link is considered valid or invalid (Gawronski & Bodenhausen, 2006a, 2011). For example, if a person has a mental association between the concepts *Muslim* and *terrorist* (e.g., as a result of prior co-activation of the two concepts during media exposure), encountering a Muslim-looking man may activate the concept *terrorist* even if the person rejects the implied connection between Muslims and terrorism (Devine, 1989). According to the APE model, the subjective validity of activated information is determined by a process of *propositional validation*. Specifically, the theory states that activated information is regarded as valid unless the default process of affirming the validity of activated information produces an inconsistency in the currently considered set of beliefs (i.e., the set of currently considered information that is regarded as valid). The central idea underlying this assumption is that, although consistency does not guarantee accuracy, inconsistency is an unambiguous indicator of an erroneous component in one's system of beliefs (Gawronski, 2012). In such cases, the currently considered set of beliefs needs to be updated, which involves a reassessment of the validity of each component.

These assumptions have important implications for understanding dissociations between spontaneous evaluations captured by implicit measures (i.e., implicit evaluations) and deliberate evaluations captured by explicit measures (i.e., explicit evaluations). According to the APE model, spontaneous evaluative responses resulting from associatively activated concepts provide the basis for evaluative beliefs (e.g., a spontaneous negative reaction toward object X providing the basis for corresponding evaluative beliefs such as *I dislike X* or *X is bad*; see Gawronski & Bodenhausen, 2006a, 2011). To the extent that the evaluative belief implied by the spontaneous response is consistent with other currently considered beliefs, it may be endorsed in a deliberate evaluative judgment. If, however, the overall set of currently considered beliefs is inconsistent, the inconsistency has to be resolved to avoid aversive feelings of dissonance (Festinger, 1957). In general, evaluative beliefs regarding a target object may be assessed for their consistency with (1) non-evaluative beliefs about states of affairs and (2) evaluative beliefs regarding other objects (Jones & Gerard, 1967). To the extent that the overall set of currently considered beliefs is inconsistent, consistency may be restored either by rejecting one of the involved beliefs (i.e., reversing its subjective truth value) or by searching for additional information that resolves the inconsistency (e.g., Gawronski, Peters, Brochu, & Strack, 2008; Gawronski & Strack, 2004). According to the APE model, dissociations between spontaneous and deliberate evaluations arise when inconsistency is resolved by rejecting one's spontaneous evaluative reaction as a valid basis for a deliberate evaluative judgment.

In addition to such "bottom-up" effects (i.e., overall set of activated information providing the basis for propositional inferences), the APE model also includes specific assumptions about "top-down" effects (i.e., propositional inferences altering the overall set of activated information). Specifically, the theory states that processes of propositional reasoning can activate new information in the course of validating activated information (Gawronski & Bodenhausen, 2006a, 2011). For example, if people are motivated to believe in the validity of a particular piece of information, they may engage in a selective search for information that supports the validity of that information (Kunda, 1990). In such cases, biased retrieval of information from memory can activate associated concepts of a particular valence, which can alter spontaneous evaluations in a “top-down” fashion (see Galdi, Gawronski, Arcuri, & Friese, 2012; Peters & Gawronski, 2011a).

An important determinant of such top-down effects is whether propositional reasoning involves an affirmation or negation of stored information. According to the APE model, merely negating the validity of activated information (i.e., reversing its truth value) is insufficient to deactivate that information (see also Strack & Deutsch, 2004). In fact, negations often lead to ironic effects, such that the activation level of the underlying associations is enhanced rather than reduced (Wegner, 1994). For example, a mental association between the concepts *old people* and *bad drivers* may give rise to the propositional thought *old people are bad drivers*, which may be rejected as false due to its inconsistency with other momentarily considered beliefs. Although such a negation may lead to a deliberate positive evaluation of old people, it may not deactivate the link between *old people* and *bad drivers*, thereby leading to a spontaneous negative evaluation (e.g., Deutsch, Gawronski, & Strack, 2006; Gawronski, Deutsch, Mbirkou, Seibt, & Strack, 2008). Yet, both spontaneous and deliberate evaluations may show corresponding responses when propositional reasoning involves an affirmation of validity. For example, countering the propositional thought *old people are bad drivers* by affirming the validity of the idea *old people are good drivers* should create or activate a link between *old people* and *good drivers*, thereby leading to positive responses for both spontaneous and deliberate evaluations.

A central aspect of such “bottom-up” and “top-down” effects is that they imply a mutual interaction between associative and propositional processes in the behavioral expression of mental representations. Although spontaneous evaluations are claimed to be the proximal behavioral outcome of associative activation and deliberate evaluations are claimed to be the proximal outcome of propositional validation, either process can function as a distal antecedent of the respective other (Gawronski & Bodenhausen, 2006a, 2011). That is, associative activation can influence deliberate evaluations by providing the basis for propositional inferences. Conversely, propositional inferences can influence spontaneous evaluations by activating new information in the course of validating activated information. According to the APE model, the central determinant of either type of influence is whether propositional reasoning involves an affirmation or negation of relevant information. Whereas the two kinds of proximal influences are assumed to be unconditional, the two kinds of distal influences should occur only when propositional reasoning involves an affirmation of validity, but not when it involves a negation of validity.

In sum, the APE model assumes that principles of feature matching and spread of activation determine the activation of stored information in response to a stimulus (*associative activation*). The perceived validity of activated information is assumed to be determined by the (in)consistency of the overall set of activated information (*propositional validation*). The APE further assumes that spontaneous evaluative reactions depend on the valence of associated concepts that are activated in response to a given stimulus. The reliance on spontaneous evaluative reactions for deliberate evaluative judgments is assumed to depend on the consistency of this reaction with other currently considered beliefs. According to the APE model, associative activation and propositional validation mutually interact, in that (1) the overall set of activated information provides the basis for propositional inferences (bottom-up influence), and (2) propositional inferences can alter the overall set of activated information (top-down influence). The central determinant of either type of influence is whether propositional reasoning involves an affirmation or negation of relevant information.

Behavioral Effects of Associative and Propositional Learning

The APE model is sometimes misunderstood as claiming that mental representations formed via associative learning exclusively influence spontaneous (but not deliberate) evaluations and mental representations formed via propositional learning exclusively influence deliberate (but not spontaneous) evaluations. Different from this mischaracterization, the APE model proposes specific conditions under which (1) deliberate evaluations should be sensitive to the effects of associative learning and (2) spontaneous evaluations should be sensitive to the effects of propositional learning (Gawronski & Bodenhausen, 2006a, 2011). On the one hand, mental representations formed via associative learning should influence deliberate evaluations when the spontaneous evaluative reactions resulting from newly created links between co-occurring stimuli are accepted as a valid basis for evaluative judgments (see Gawronski & Bodenhausen, 2006a, Case 1), but not when these spontaneous evaluative reactions are rejected as false (see Gawronski & Bodenhausen, 2006a, Case 2). On the other hand, mental representations formed via propositional learning should influence spontaneous evaluative reactions when propositional reasoning involves an affirmation of the newly learned information (see Gawronski & Bodenhausen, 2006a, Case 4), but not when propositional reasoning involves a negation of the newly learned information (see Gawronski & Bodenhausen, 2006a, Case 3).

These assumptions have important implications for research on the role of associative and propositional learning in EC. Counter to widespread claims in current theoretical debates, evidence that verbal instructions about CS-US pairings can influence spontaneous evaluations in the absence of directly experienced CS-US pairings (e.g., Gast & De Houwer, 2012; Kurdi & Banaji, 2017) does not support the idea that EC effects exclusively result from propositional learning, as suggested by single-process propositional theories (e.g., De Houwer, 2009, 2014). First, any such conclusion commits the logical fallacy of *affirming the consequent*, in that the conditional *if X, then Y* is used to infer *X* from the observation of *Y* (Gawronski & Bodenhausen, 2015). Although research on verbal instruction effects shows that spontaneous evaluations can be influenced by propositional learning, it does not permit the reverse conclusion that any effect on spontaneous evaluations is a result of propositional learning. Second, as we explained above, the APE model is perfectly consistent with effects of propositional learning on spontaneous evaluations to the extent that propositional reasoning involves an affirmation of the newly learned information (see Gawronski & Bodenhausen, 2006a, Case 4). Yet, such effects do not preclude the possibility that CS-US co-occurrences influence spontaneous evaluations via associative learning (see Gawronski & Bodenhausen, 2006a, Cases 1 and 2).

From the perspective of the APE model, understanding the roles of associative and propositional learning in EC requires experimental designs that pit the outcomes of one learning mechanism against the outcomes of the other (see Gawronski & Bodenhausen, 2006a, Case 5). A general feature of such designs is that they include (1) repeated co-occurrence of a CS with a US and (2) additional information that the valence of the US is not characteristic of the CS. Based on the defining features of the two learning mechanisms, the APE model suggests three cases in which additional information should lead to conflicting outcomes of associative and propositional learning: (1) cases in which observed regularities are rejected as invalid, (2) cases in which the particular relation of a CS and a US suggests a CS evaluation that is opposite to the valence of the US, and (3) cases in which CS evaluations suggested by directly experienced CS-US pairings conflict with CS evaluations suggested by verbal instructions about CS-US pairings. In these cases, the APE model predicts that perceived validity, relational information, and verbal instructions should eliminate mere co-occurrence effects on deliberate evaluations (because propositional reasoning should lead to a rejection of the spontaneous evaluative reactions resulting from the direct link between the co-occurring stimuli formed via associative learning). In contrast, spontaneous evaluations should remain sensitive to the observed co-occurrence between stimuli (because rejection of spontaneous evaluative reactions via propositional reasoning does not deactivate the direct links between the co-occurring stimuli formed via associative learning). Although some studies support these predictions, the overall evidence for conflicting effects of associative and propositional learning is rather mixed (for a review, see Corneille & Stahl, 2018).

Regularity vs. Validity

Peters and Gawronski (2011b) conducted a series of studies that investigated interactive effects of evaluative regularities and their perceived validity on spontaneous and deliberate evaluations. Using a simple impression formation task, participants were presented with evaluative statements about four target individuals. For two of the four targets, 75% of the statements were positive and 25% were negative. For the other two targets, 75% of the statements were negative and 25% were positive. Participants' task was to guess whether each statement was correct or incorrect. Orthogonal to the manipulation of valence proportions, participants received feedback on their individual guesses, such that for two of the targets the majority information was always correct and the minority information was always incorrect. For the remaining two targets, the feedback suggested that the minority information was correct and the majority information was incorrect. Afterwards, spontaneous evaluations were measured with two variants of affective priming (Fazio, Jackson, Dunton, & Williams, 1995; Payne, Cheng, Govorun, & Stewart, 2005); deliberate evaluations were assessed with a self-report measure.

Drawing on the APE model's hypotheses that (1) perceived validity should eliminate mere co-occurrence effects on deliberate evaluations and (2) spontaneous evaluations should remain sensitive to stimulus co-occurrences even when associative links resulting from these co-occurrences are rejected as false (Gawronski & Bodenhausen, 2006a, 2011), Peters and Gawronski (2011b) expected an unqualified main effect of valence proportions on spontaneous evaluations. In contrast, deliberate evaluations were expected to show an interaction of valence proportions and validity feedback, reflecting the actual validity of the observed statements rather than the mere proportions of positive and negative statements. Counter to these predictions, both spontaneous and deliberate evaluations showed a significant interaction, indicating that validity information fully qualified the effects of the observed regularities (see also Moran, Bar-Anan, & Nosek, 2015).

An important aspect of Peters and Gawronski's (2011b) studies is that they focused on the effects of invalidation (or negation) during the *formation* of evaluative representations. This focus deviates from the one in earlier research that focused on the effects of invalidation (or negation) during the *expression* of existing representations (e.g., Deutsch et al., 2006; Gawronski, Deutsch, et al., 2008). Different from the findings by Peters and Gawronski (2011b), these studies consistently found that invalidation (or negation) of prior information qualifies deliberate, but not spontaneous evaluations. In fact, when Peters and Gawronski included a delay between the encoding of evaluative information and its invalidation, they replicated the typical dissociation found in earlier studies. In this case, spontaneous evaluations were indeed less sensitive to invalidation than deliberate evaluations. Together, these findings suggest that dissociations between spontaneous and deliberate evaluations resulting from invalidation (or negation) are due to processes operating during the expression of evaluative representations. However, they pose a challenge to the APE model's hypothesis that mere co-occurrences influence mental representations via associative learning even when these regularities are deemed invalid during encoding.

Co-occurrence vs. Relation

Although Peters and Gawronski's (2011b) findings pose a challenge to the idea that observed regularities can influence mental representations via associative learning regardless of the perceived validity of these regularities, there is some evidence that observed co-occurrences can influence mental representations regardless of the relation between the co-occurring stimuli (for a review, see Corneille & Stahl, 2018). This evidence comes from research showing that CS-US pairings can lead to changes in spontaneous CS evaluations that remain unqualified by the particular relation between the CS and the US. The central finding in this line of research is that information suggesting a "contrastive" relation between a CS and a co-occurring US reverses EC effects on deliberate evaluations without affecting EC effects on spontaneous evaluations. These findings are consistent with the APE model's assumptions that (1) contrastive relations should eliminate mere co-occurrence effects on deliberate evaluations and (2) spontaneous evaluations should remain sensitive to observed co-occurrences even when the propositional meaning of the resulting associative links is qualified by information about a contrastive relation. For example, repeated exposure to the information *Aspirin relieves headaches* may create a direct link between *Aspirin* and *headaches* via associative learning even when propositional learning creates a mental representation of their relation at a higher level of the associative network. Thus, exposure to the stimulus *Aspirin* should activate the negative concept *headaches* via the direct link between two, even when activation spreads to the stored information about their relation. From the perspective of the APE model, entertaining the idea *Aspirin relieves headaches* is not only insufficient to deactivate the direct link between *Aspirin* and *headaches*; it may even enhance this link by contributing to the co-activation of the two concepts.

Although the available evidence for these ideas is somewhat mixed, some studies have found unqualified effects of CS-US co-occurrences on spontaneous evaluations even when deliberate evaluations were fully qualified by information about CS-US relations. For example, in a study by Moran and Bar-Anan (2013), participants were presented with neutral stimuli (CS) that started or stopped either pleasant or unpleasant sounds (US). Afterwards, the authors measured spontaneous evaluations of the CSs with an implicit association test (Greenwald, McGhee, & Schwartz, 1998); deliberate evaluations were assessed with a self-report measure. On the measure of deliberate evaluations, participants showed more favorable judgments of stimuli that started pleasant sounds compared with stimuli that started unpleasant sounds. Conversely, participants showed more favorable judgments of stimuli that stopped unpleasant sounds compared with stimuli that stopped pleasant sounds. In contrast, the measure of spontaneous evaluations reflected the mere co-occurrence of CSs and USs regardless of their relation. That is, participants showed more favorable responses to stimuli that co-occurred with pleasant sounds compared with stimuli that co-occurred with unpleasant sounds, regardless of whether the stimuli started or stopped the sounds.

Hu, Gawronski, and Balas (2017a, Experiments 1 and 2) found a similar pattern using an affective priming task (Fazio et al., 1995) and a manipulation of relational information involving causal relations. Participants were presented with image pairs involving pharmaceutical products (CS) and images of positive or negative health conditions (US). Participants were told that the pharmaceutical products either cause or prevent the depicted health conditions. This manipulation was based on the idea that pharmaceutical products can have positive effects (e.g., curing eczema; causing healthy skin) as well as negative side-effects (e.g., causing eczema; impairing healthy skin). Consistent with Moran and Bar-Anan's (2013) results, Hu et al. found that deliberate evaluations of the pharmaceutical products reflected the relation between the product and the depicted health condition. Specifically, participants showed more favorable judgments of products that caused positive health conditions compared with products that caused negative health conditions. Conversely, participants showed more favorable judgments of products that prevented negative health conditions compared with products that prevented positive health conditions. In contrast, spontaneous evaluations reflected the mere co-occurrence of the products with positive or negative health conditions regardless of their relation. That is, participants showed more favorable responses to products that co-occurred with positive health conditions than products that co-occurred with negative health conditions, regardless of whether the products caused or prevented the health conditions.

Although the reviewed findings support the idea that mere co-occurrences of a CS and a US can influence mental representations via associative learning regardless of their relation, the available evidence for this hypothesis is far from conclusive (for a review, see Corneille & Stahl, 2018). For example, counter to the obtained dissociations between spontaneous and deliberate evaluations, some studies found attenuated EC effects on spontaneous evaluations when relational information suggested a CS evaluation that was opposite to the valence of the US (e.g., Zanon, De Houwer, & Gast, 2012); other studies found reversed EC effects on both spontaneous and deliberate evaluations (e.g., Gawronski, Walther, & Blank, 2005; Hu et al., 2017a, Experiment 3). Although the APE model can be reconciled with the former findings by assuming that participants engaged in propositional inferences about the abstract valence of the CS (instead of storing episodic information about the relation between the CS and the US), it is difficult to reconcile with the latter finding given that effects of co-occurrences are assumed to be unconditional. Thus, a major challenge for future theoretical and empirical work is to identify the conditions under which relational information qualifies effects of mere co-occurrences on spontaneous evaluations.

Experience vs. Instruction

More compelling evidence for the predictions of the APE model comes from research that investigated cases in which CS evaluations suggested by directly experienced CS-US pairings conflict with CS evaluations suggested by verbal instructions about CS-US pairings. In a nutshell, this work found that deliberate evaluations were shaped by verbal instructions about CS-US pairings, whereas spontaneous evaluations were influenced by directly experienced CS-US pairings. These findings are consistent with the APE model's assumption that spontaneous and deliberate evaluations should be differentially sensitive to the outcomes of associative and propositional learning to the extent that (1) a CS repeatedly co-occurs with a US and (2) additional information suggests that the valence of the co-occurring US is not characteristic of the CS.

For example, in a study by Hu, Gawronski, and Balas (2017b), participants were presented with CS-US pairings and then completed measures of spontaneous and deliberate evaluations of the CSs. After completion of the two evaluation measures, half of the participants were presented with CS-US pairings of the opposite valence (i.e., counter-conditioning). The remaining half were informed that the CSs would be presented with USs of the opposite valence without being shown any such pairings (i.e., counter-instructions). Finally, all participants completed the two evaluation measures a second time. Consistent with the predictions of the APE model, both counter-conditioning and counter-instructions effectively reversed initial EC effects on deliberate evaluations. However, only directly experienced CS-US pairings, but not verbal instructions about CS-US pairings, reversed initial EC effects on spontaneous evaluations. These results suggest that actually experienced CS-US pairings influence mental representations in a manner that is different from the effects of propositional inferences. If all EC effects were mediated by propositional inferences (as suggested by single-process propositional theories), counter-conditioning and counter-instructions should have identical effects, which was not the case in Hu et al.'s study.

Learning or Expression?

The mixed evidence regarding competing effects of associative and propositional learning raises the question of how the conflicting findings could be reconciled from the perspective of the APE model. One possibility is to attribute the obtained dissociations between spontaneous and deliberate evaluations to processes during the behavioral expression of an evaluative response. Specifically, dissociations between the two kinds of evaluations may occur when newly acquired propositional information conflicts with the evaluation suggested by a previously formed representation. For example, in Peters and Gawronski's (2011b) study on the effects of observed regularities and perceived validity, the predicted dissociation between spontaneous and deliberate evaluations occurred only when there was a delay between the encoding of evaluative regularities and the presentation of validity information. Similarly, in Hu et al.'s (2017b) study on the effects of counter-conditioning and counter-instructions, both procedures were supposed to reverse the effect of previously experienced CS-US pairings. Thus, in both studies, participants initially formed a mental representation of a target object, and the evaluation suggested by this representation was later qualified by newly acquired information about the validity of previously observed regularities (Peters & Gawronski, 2011b) or forthcoming CS-US pairings (Hu et al., 2017b). Such information may influence the verbal expression of the existing representation on measures of deliberate evaluative judgments, but it may be ineffective in preventing the effect of this representation on measures of spontaneous evaluative reactions. A change in the latter presumably requires a genuine change in the underlying representation, for example via directly experienced CS-US pairings of the opposite valence (see Hu et al., 2017b).

A similar argument may explain the mixed findings regarding the effectiveness of relational information in qualifying EC effects on spontaneous evaluations (e.g., Gawronski et al., 2005; Hu et al., 2017a; Zanon et al., 2012). An important factor that moderated dissociations between spontaneous and deliberate evaluations in Hu et al.'s (2017a) research was whether the information about CS-US relations was acquired together with the CS-US pairings. When relational information was acquired separately before the encoding of the CS-US pairings, it moderated EC effects on deliberate, but not spontaneous, evaluations. However, when relational information was provided simultaneously with the CS-US pairings, it moderated EC effects on both spontaneous and deliberate evaluations. The latter setup is consistent with the one in other studies that found a qualifying effect of relational information on both spontaneous and deliberate evaluations (e.g., Gawronski et al., 2005; Zanon et al., 2012). Thus, a potential explanation of these outcomes is that dissociations between spontaneous and deliberate evaluations occur when participants have to integrate two separately encoded pieces of information (i.e., co-occurrence, relation). To the extent that a retroactive integration of two pieces of stored information requires time and cognitive resources, it may be less likely to occur for spontaneous evaluations, showing mere co-occurrence effects regardless of specific CS-US relations. Yet, if the two pieces of information are integrated during encoding (e.g., when relational information is encoded together with CS-US pairings), the resulting CS representation may directly reflect the valence inferred from the observed relation, leading to a qualifying effect of relational information on both spontaneous and deliberate evaluations.

Although this integrative post-hoc explanation reconciles the reviewed findings with the APE model, it is also consistent with single-process propositional theories that reject the idea of associative learning as a qualitative distinct learning mechanism (e.g., De Houwer, 2009, 2014). To the extent that the observed dissociations between spontaneous and deliberate evaluations are due to expression-related processes, there is no need to postulate a mechanism of associative learning in addition to the uncontroversial idea of propositional learning (see Hu et al., 2017a). Nevertheless, some findings are more difficult to explain without the idea of associative learning based on mere-occurrence (e.g., Gawronski et al., 2005, Experiment 2; Langer, Walther, Gawronski, & Blank, 2009). For example, in a study modeled after the typical paradigm in research on US revaluation (e.g., Walther et al., 2009), Langer et al. (2009) presented participants with positive or negative information about some unknown individuals (i.e., US faces). In a subsequent task, participants received information about whether the US faces like or dislike unknown individuals (i.e., CS faces). Afterwards, half of the participants received novel information about the US faces suggesting a valence that was opposite to the one in the first part of the study (US revaluation). Participants in a control condition received additional information about the US faces that was neutral in valence. Finally, all participants completed a measure of deliberate CS evaluations. Participants in the control condition showed more favorable evaluations of CS faces that were liked by positive US faces compared to the CS faces that were liked by negative US faces. Moreover, participants showed less favorable evaluations of CS faces that were disliked by positive US faces compared to the CS faces that were disliked by negative US faces (see also Gawronski et al., 2005). However, participants in the revaluation condition showed CS evaluations that were in line with the new valence of the US regardless of whether the US faces liked or disliked the CS faces. That is, participants showed more favorable evaluations of CS faces that co-occurred with initially negative, revaluated positive US faces compared to CS faces that co-occurred with initially positive, revaluated negative US faces. This result suggests that participants formed a direct mental link between the CS and the US, which determined their CS evaluations regardless of CS-US relations after US revaluation. This conclusion is consistent with the notion of associative learning. However, it is difficult to reconcile with the categorical rejection of associative learning in single-process propositional theories (e.g., De Houwer, 2009, 2014). According to the latter theories, US revaluation should either reverse the two-way interaction pattern found in Langer et al.'s control condition or leave it unaffected. From a purely propositional view, it seems unclear why CS evaluations should reflect the new valence of the US regardless of CS-US relations. Although Langer et al.'s findings also do not cleanly map onto the core assumptions of the APE model, they do suggest that CS-US co-occurrences can create direct associative links even when the observed co-occurrences and CS-US relations are mentally integrated during encoding. Thus, more work is needed to identify the conditions under which CS-US pairings produce unqualified co-occurrence effects regardless of specific CS-US relations.

Automaticity

The distinction between *associative learning* and *propositional learning* refers to the processes by which observed stimuli shape mental representations; the distinction between *associative activation* and *propositional validation* refers to the processes by which existing mental representations influence judgments and behavior. In either case, the associative-propositional distinction specifies *how* a given process translates inputs into outputs (i.e., operating principles). The distinction between associative and propositional learning refers to two distinct processes that translate input stimuli into representational outputs. The distinction between associative activation and propositional validation refers to two distinct processes in the translation between representational inputs and behavioral outputs (see De Houwer & Moors, 2015).

The question of how a given process translates inputs into outputs has to be distinguished from the question of *when* the process is assumed to operate (i.e., operating conditions). The latter question is central to the automatic-controlled distinction, which subsumes four distinct boundary conditions (see Bargh, 1994): awareness (i.e., does the process operate when there is no awareness?), intentionality (i.e., does the process operate when there is no intention to start the process?), efficiency (e.g., does the process operate when mental resources are reduced?), and controllability (e.g., does the process operate when there is a goal to alter or stop the process?). According to the APE model, there is no perfect mapping between operating principles and operating conditions, such that associative processes would generally operate automatically, whereas propositional processes would generally operate in a controlled fashion (Gawronski & Bodenhausen, 2007, 2009, 2011). Instead, both associative and propositional processes are assumed to have automatic and controlled aspects. Moreover, because different features of automaticity need not co-occur, it is important to distinguish between the unique roles of awareness, intentionality, efficiency, and controllability. Based on the dominant interpretation of EC as a learning effect, the following analysis focuses primarily on automatic and controlled aspects of associative and propositional learning. For a detailed discussion of automatic and controlled aspects of associative activation and propositional validation, we refer interested readers to the chapter by Gawronski and Bodenhausen (2014).

Awareness

The APE model assumes that associative learning is independent of people's awareness of the stimulus contiguities that are responsible for the formation of new associative links (Gawronski & Bodenhausen, 2014). More specifically, the APE model suggests that people do not have to consciously recognize the co-occurrence of a CS with a US during the encoding of their co-occurrence in order to form a mental link between the two. To clarify the empirical implications of this hypothesis, two important aspects deserve a more elaborate discussion.

First, awareness during encoding should not be confused with recollective memory during the expression of an evaluative response. Although measures of recollective memory have been frequently used to investigate the role of awareness in EC, they provide little insights into whether people consciously recognized the co-occurrence of a CS with a US during encoding (Gawronski & Walther, 2012). On the one hand, lack of recollective memory does not guarantee lack of awareness during encoding, because participants may have consciously recognized the co-occurrence of a CS with a US during encoding but forgotten the relevant details over time. On the other hand, accurate memory judgments do not imply awareness of CS-US pairings during encoding, because participants may use their conditioned evaluative response to the CS as a basis for their judgments in the memory task. As explained in more detail by Gawronski and Walther (2012), understanding the role of awareness in EC requires experimental approaches that include (1) direct manipulations of awareness during encoding and (2) online (rather than retrospective) manipulation checks of awareness (see also Sweldens, Corneille, & Yzerbyt, 2014).

Second, awareness of CS-US pairings during encoding should not be confused with attention (Gawronski & Bodenhausen, 2014). Although enhanced attention may increase the likelihood that people consciously recognize systematic co-occurrences between a CS and a US, the two determinants of learning are conceptually distinct in the sense that attention is necessary but insufficient for awareness. That is, people may consciously recognize the co-occurrence of a CS and a US only when they pay attention to the two stimuli, but attention does not guarantee awareness of their systematic co-occurrence. According to the APE model, associative learning is independent of awareness, but it still requires attention to the co-occurring stimuli (e.g., Custers & Aarts, 2011; Field & Moore, 2005).

Whereas associative learning is assumed to be independent of awareness, the APE model assumes that propositional learning requires conscious awareness (Gawronski & Bodenhausen, 2014). Thus, to the extent that people are consciously aware of the co-occurrence of a CS with a US, associative and propositional learning may jointly influence the mental representation of the CS. Yet, lack of awareness should eliminate the effects of propositional learning, leaving associative learning as the sole mechanism that can produce changes in the mental representation of the CS.

Intentionality

According to the APE model, associative learning can be described as unintentional in the sense that the learning process itself does not require the goal to form a new association (Gawronski & Bodenhausen, 2014). However, associative learning can certainly have intentional antecedents, in that people may intentionally expose themselves to repeated co-occurrences to facilitate their acquisition (e.g., intentional exposure to pairs of words in the learning of a foreign language).

Similar considerations apply to propositional learning. Once a propositional thought about an event has been generated (e.g., about the co-occurrence of a CS with a US), the content of this thought may be stored in memory even when people do not have the goal to memorize it. Of course, the goal to memorize the content of a propositional thought may facilitate its storage, but such a goal is not required according to the APE model (Gawronski & Bodenhausen, 2014). In fact, a substantial body of evidence suggests that alternative processing goals (e.g., impression formation) can be more effective in producing a strong memory trace than memorization goals (e.g., Hamilton, Katz, & Leirer, 1980). The same is true for the process of generating a propositional thought, which also does not require an intention to instigate this process (Gawronski & Bodenhausen, 2014). Although propositional thoughts can be the result of intentional inferences about CS-US relations, they can be purely "stimulus-driven" when people notice a systematic relation between stimuli in the absence of a goal to identify such stimulus relations.

Efficiency

According to the APE model, the formation of mental links via associative learning is resource-independent, although attentional distraction may sometimes disrupt associative learning if it undermines attention to and encoding of the co-occurring stimuli (Gawronski & Bodenhausen, 2014). Thus, when investigating the resource-independence of associative learning, it is important to distinguish between different aspects of working memory capacity (Baddeley, 2010). Whereas capacity constraints on episodic memory (e.g., concurrent rehearsal of a complex digit-string) should leave associative learning unaffected, capacity constraints on perceptual processing (e.g., concurrent attention to numbers in a two-back task) may reduce the effects of associative learning to the extent that it undermines the perceptual encoding of the relevant stimulus contiguities (see Custers & Aarts, 2011; Field & Moore, 2005; Pleyers, Corneille, Yzerbyt, & Luminet, 2009).

As for propositional learning, the APE model assumes that the mere consideration of propositional information does not require substantial amounts of cognitive resources, although greater elaboration can certainly strengthen the resulting mental links (Craik & Lockhart, 1972). Nevertheless, the generation of propositional information may require more resources to the extent that the entertained information is more complex (Gawronski & Bodenhausen, 2014). The same is true for complex propositional inferences about CS-US relations that are not directly observable, but have to be inferred from multiple pieces of information. For example, whereas propositional thoughts about the co-occurrence of a CS and a US may spring to mind even when there are little cognitive resources, propositional inferences about causal relations between two stimuli may require more cognitive capacity, such that these inferences may be disrupted under cognitive load.

Controllability

According to the APE model, associative learning is difficult to control in the sense that observed CS-US pairings can create mental links despite the goal of not forming an association between the co-occurring stimuli (e.g., Gawronski, Balas, & Creighton, 2014). Yet, the relative effectiveness of counteractive control is assumed to depend on the employed control strategy (Gawronski & Bodenhausen, 2014). Similar to the distinction between suppression and reappraisal in research on emotion regulation (Gross, 1998), the APE model suggests that *proactive* control strategies are more effective in counteracting the formation of direct associative links on the basis of observed co-occurrence than *reactive* control strategies. For example, whereas suppressing one's evaluative response to the US should be relatively ineffective in preventing the formation of a valence-congruent representation of the CS, categorizing the US in a way that is opposite to its "default" valence might be much more effective (but see Gawronski, Mitchell, & Balas, 2015, for evidence suggesting that reappraisal of US valence may be ineffective in reducing EC effects).

As for the controllability of propositional learning, it is certainly possible to intentionally invalidate self-generated propositional information (Gawronski & Bodenhausen, 2014). Such goal-dependent invalidation is conceptually equivalent to the effects of motivated reasoning, in which people may have a desire to confirm or disconfirm the validity of a given piece of information (Kunda, 1990). As with negation effects in motivated reasoning, however, the effectiveness of intentional invalidation is assumed to be limited to the propositional level, in that merely negating propositional information (i.e., reversing its truth value) is insufficient to counteract the formation of corresponding mental links via associative learning (Deutsch et al., 2006; Gawronski, Deutsch, et al., 2008). For example, repeated pairings of a political candidate with unpleasant stimuli in smear campaigns may create a negative associations with that candidate even when perceivers reject the validity of the observed co-occurrence (e.g., Kosloff, Greenberg, Schmader, Dechesne, & Weise, 2010; Weinberger & Westen, 2008). The process of generating propositional thoughts can also be difficult to control, in that such thoughts may often be the "stimulus-driven" result of noticing a systematic relation between stimuli.

Investigating Automatic Features of EC

Although the APE model's hypotheses about automatic features of associative and propositional learning may seem relatively straightforward, it is important to consider three essential issues in empirical tests of these hypotheses. First, any study on automatic features of associative and propositional learning has to distinguish between learning-related processes during the encoding of CS-US pairings and expression-related processes during the measurement of evaluative responses (Gawronski, Balas, & Hu, 2016; Gawronski & Bodenhausen, 2014). Although this issue is particularly important for understanding the distinct roles of awareness during encoding and recollective memory after encoding, it also applies to research on other features of automaticity. For example, in their work on the controllability of EC, Gawronski, Balas, and Creighton (2014) found that, although EC effects on self-reported evaluations were significantly reduced when participants were instructed to prevent being influenced by CS-US pairings, EC effects on an evaluative priming measure remained unaffected by control instructions. Moreover, although EC effects on self-reported evaluations varied as a function of evaluative priming effects and recollective memory for CS-US pairings, motivation to control the influence of CS-US pairings qualified only the predictive relation of recollective memory. These results suggest that control instructions influenced the use of recollective memories for CS-US pairings in the expression of deliberate CS evaluations (see also Balas & Gawronski, 2012). However, control instructions did not prevent the formation of associative links, which continued to influence both spontaneous and deliberate evaluations of the CS. Thus, when studying automatic features of EC, it is important to distinguish between processes involved in the formation of mental representations and the processes involved in the behavioral expression of these representations (Gawronski et al., 2016).

A second important issue is the difference between moderating and residual effects in research on automatic features of EC. For example, in studies that experimentally manipulate awareness of CS-US pairings (e.g., Stahl, Haaf, & Corneille, 2016), a reduction of EC effects as a result of unawareness would speak to the relative contribution of propositional learning, given that propositional learning is assumed to require awareness. However, such a moderating effect does not provide any evidence for the absence of associative learning. Effects of associative learning would have to be inferred from the size of residual EC effects in the absence of awareness. To the extent that residual EC effects are still statistically significant under conditions of unawareness, the APE model would suggest a joint operation of associative and propositional learning. The contribution of propositional learning would be reflected in the significant moderation of EC effects by awareness. The contribution of associative learning would be reflected in the significant residual EC effect under conditions of unawareness.

Because EC effects of associative learning are presumably much smaller compared to EC effects of propositional learning (see Hofmann et al., 2010), these considerations highlight the importance of statistical power in studies on automatic features of EC. Although the sample of a given study may provide sufficient power to detect a significant moderation of EC effects by manipulations of automatic processing conditions, the same study may have insufficient power to reliably detect a significant residual EC effect under conditions of automatic processing. An illustrative example is a series of studies on the resource-dependence of EC effects by Mierop, Hütter, and Corneille (2017). Across three studies, the authors found a significant reduction of EC effects as a result of cognitive load. Importantly, although there was no evidence for a residual EC effect under cognitive load in each individual study (*N*s = 34, 41, 61), an integrative analysis of the data from all three studies ($N = 136$) did find a significant residual EC effect with an observed power of 93%. These results are consistent with concerns that, although a given study may have sufficient statistical power to detect a significant moderation of EC effects by manipulations of automatic processing conditions, the same study may have insufficient power to detect a significant residual EC effect under conditions of automatic processing.

Finally, a third important issue is that hypotheses about automatic features of EC specify the conditions under which a given a process is assumed to operate (i.e., operating conditions) rather than the defining features of that process (i.e., operating principles). Thus, empirical evidence for or against any of these hypotheses speaks only to the (in)accuracy of assumptions about operating conditions. It does not speak to the question of whether a given process actually contributes to EC effects (Gawronski & Bodenhausen, 2009, 2014). For example, in current debates between proponents of dual-process versus single-process theories, lack of a significant EC effect under conditions of unawareness is sometimes interpreted as evidence against the existence of associative learning (e.g., Mitchell et al., 2009). However, based on the distinction between operating principles and operating conditions, such evidence may simply point to the inaccuracy of the hypothesis that associative learning does not require awareness. Conversely, it has been argued that a significant EC effect under conditions of unawareness provides evidence for the existence of associative learning. Yet, again, such evidence may simply point to the inaccuracy of the hypothesis that propositional learning requires awareness. Conceptually, the question of whether a given process actually contributes to EC effects requires experimental manipulations capturing the operating principles of that process (as discussed in the section *Behavioral Effects of Associative and Propositional Learning*), not hypotheses about its operating conditions (see Gawronski, Sherman, et al., 2014).

Although the literature on automaticity features of EC is too large to be reviewed in this article, it is worth noting that the available evidence is consistent with some assumptions of the APE model and inconsistent with others. Based on a detailed review of relevant evidence by Corneille and Stahl (2018), we conclude that:

1. There is no compelling evidence for residual EC effects under experimental conditions that prevent conscious awareness during encoding of CS-US pairings (e.g., Stahl et al., 2016).[3] This conclusion stands in contrast to the APE model's hypothesis that associative learning can lead to EC effects in the absence of conscious awareness.
2. Although many studies on the resource-dependence of EC effects found no residual effects under mental load (for a notable exception, see Mierop et al., 2017), the majority of these studies used manipulations of attentional load and sample sizes that seem too small to provide sufficient statistical power (e.g., Pleyers et al., 2009). These ambiguities undermine strong conclusions about the accuracy of the APE model's assumption that associative learning can lead to EC effects under conditions of reduced cognitive resources.
3. Although there is evidence that EC effects depends on evaluative processing goals (e.g., Gast & Rothermund, 2011), goals to either promote or prevent the impact of CS-US pairings during encoding have been found to qualify EC effects on deliber-

---

[3] Although some studies suggest that EC effects can be obtained with subliminal presentations of either the CSs (e.g., Gawronski & LeBel, 2008) or the USs (e.g., Rydell, McConnell, Mackie, & Strain, 2006), all successful demonstrations of subliminal EC suffer from methodological limitations (for reviews, see Corneille & Stahl, 2018; Sweldens et al., 2014). Conversely, all studies that do not suffer from these limitations failed to obtain evidence for subliminal EC (e.g., Högden, Hütter, & Unkelbach, 2018; Stahl et al., 2016).

ate, but not spontaneous evaluations (e.g., Gawronski, Balas, et al., 2014). The latter findings are consistent with the APE model's hypotheses regarding the (un)intentionality and (un)controllability of the learning mechanisms underlying EC (see also Hütter & Sweldens, 2018).

Other Questions

The preceding sections reviewed the core assumptions of the APE model that are relevant for understanding the mental underpinnings of EC. Yet, there are a number of additional questions that play a central role in research on EC, but have not been previously addressed by the APE model: (1) Do EC effects reflect statistical CS-US contingencies or CS-US contiguity? (2) Are EC effects resistant to extinction? (3) Do EC effects depend on individual difference factors?

Contingency vs. Contiguity

It is widely assumed that EC effects reflect CS-US contiguities rather than statistical CS-US contingencies (De Houwer, Thomas, & Baeyens, 2001; Walther, Nagengast, & Trasselli, 2005). The core assumptions of the APE model are generally consistent with this hypothesis, but the distinction between associative and propositional learning suggests a potential qualification. Conceptually, the notion of associative learning suggests that repeated co-activation of two concepts creates a mental link between the two. Moreover, the strength of this link should increase with increasing numbers of observed co-occurrences, but it seems unlikely that its strength would decrease when one of the stimuli is occasionally encountered in isolation or with another stimulus. From this perspective, the critical determinant of associative learning should be CS-US contiguity. What matters is the overall frequency of observed co-occurrences, not statistical contingencies of their co-occurrence.

A different conclusion could be drawn for propositional learning, which depends on the generation of propositional thoughts about the relation between co-occurring stimuli. To the extent that a CS sometimes does and sometimes does not co-occur with a particular US, it seems less likely that perceivers infer a systematic relation between the two. For example, if a CS co-occurs with a given US for 50% of one's observations and without the US for the other 50%, a causal relation between the two seems unlikely, at least if one assumes no interfering factor that prevents the occurrence of the US in the presence of the CS. From this perspective, it seems possible that propositional learning is more sensitive to statistical CS-US contingencies rather than CS-US contiguity. Of course, perceivers may sometimes perceive statistical contingencies where there is none, as shown in research on illusory correlations (e.g., Hamilton & Gifford, 1976). Thus, what matters for propositional learning is the perceived statistical contingency between stimuli, which may not always reflect their actual statistical contingency (e.g., Kattner, 2014; Kattner & Ellermeier, 2011).

Extinction

Similar considerations apply to the extinction of EC effects due to unreinforced encounters with a CS in the absence of the US. Consistent with the idea that EC effects reflect CS-US contiguities rather than CS-US contingencies, it is widely assumed that EC effects are resistant to extinction (De Houwer et al., 2001; Walther et al., 2005). Similar to our arguments about the role of CS-US contiguity in associative learning, unreinforced encounters with a CS in the absence of the US may not necessarily reduce the strength of a previously formed mental link. From this perspective, the mental representations resulting from associative learning should lead to EC effects that are largely resistant to extinction.

Again, a different conclusion could be drawn for propositional learning. To the extent that a CS that repeatedly co-occurred with a US is later encountered without the US, perceivers may draw propositional inferences about properties of the CS that reflect the new information, which may lead them to revise their previously formed beliefs about the relation between the CS and the US. Thus, different from the presumed resistance of EC effects resulting from associative learning, EC effects resulting from propositional learning may be systematically reduced by unreinforced encounters with the CS in the absence of the US. This conclusion is consistent with results by Gawronski, Gast, and De Houwer (2015), who found that unreinforced CS presentations reduced EC effects on self-reported evaluations, whereas EC effects on an evaluative priming measure remained unaffected by unreinforced CS presentations.[4]

Individual Differences

A final question concerns the potential role of individual differences in EC. Although previous discussions of the APE model focused primarily on general mechanisms of evaluative learning, its core assumptions suggest that EC effects may be moderated by any individual difference factor that influences either (1) the formation of mental representations via associative or propositional learning or (2) the contribution of associative and propositional processes in the behavioral expression of these representations. With regard to the formation of mental representations, the most likely path by which individual differences may moderate EC effects is by influencing the propositional thoughts that are generated in response to CS-US pairings. For example, individuals with high versus low levels of causal uncertainty (Weary & Edwards, 1994) or high versus low need for structure (Neuberg & Newsom, 1993) may generate different interpretations of ambiguous CS-US

---

[4] EC effects on self-reported evaluations were reduced by unreinforced CS presentations when the CSs had been rated after the initial presentation of CS-US pairings, but not when the CSs had not been rated after the initial presentation of CS-US pairings. According to Gawronski et al. (2015), this asymmetry reflects a differential integration of unreinforced CS presentations into evaluative judgments. Although this idea goes beyond the theoretical assumptions of the APE model, it is consistent with the argument that extinction involves the updating of previously formed beliefs about the relation between the CS and the US.

pairings, with corresponding effects on their mental representation of the CS. With regard to the behavioral expression of mental representations, a moderation by individual differences seems particularly likely for the role of propositional processes in determining the reliance on spontaneous evaluative reactions (Kendrick & Olson, 2012). For example, individuals with high (versus low) need for cognition (Cacioppo & Petty, 1982) or low (versus high) faith in intuition (Epstein, Pacini, Denes-Raj, & Heier, 1996) may be less likely to rely on their spontaneous evaluative reactions for deliberate evaluative judgments. As a result, individuals with these personality characteristics should be less likely to show EC effects of associative learning on explicit measures, but there may be no such individual differences for associatively driven EC effects on implicit measures.

Is the APE Model Falsifiable?

A common question about the APE model is whether the theory is unfalsifiable in the sense that its high level of complexity makes it possible to explain any empirical outcome in a post-hoc fashion without generating testable predictions that could be disconfirmed. Two major sources of this concern are (1) the assumption of mutual interactions between associative and propositional processes and (2) the rejection of a perfect mapping between operating principles (i.e., associative vs. propositional), operating conditions (i.e., automatic vs. controlled), and behavioral outcomes (i.e., spontaneous vs. deliberate evaluations). Although we agree that the high complexity of the APE model requires considering multiple factors in the logical derivation of empirical predictions, the theory includes precise assumptions about (1) the conditions under which particular patterns of process interactions should and should not occur and (2) what behavioral outcomes result from the different interaction patterns (see Gawronski & Bodenhausen, 2006a, 2011). Thus, the theory is falsifiable in the sense that it prohibits behavioral outcomes that are inconsistent with the theory's assumptions about the conditions under which particular process interactions should occur. In fact, although many of the APE model's predictions in this regard have been empirically confirmed, there are some findings that are clearly inconsistent with the predictions of the theory (for a review, see Gawronski & Bodenhausen, 2011). From this perspective, the APE model could be criticized for including assumptions that make it inconsistent with particular empirical findings. However, any such criticism directly contradicts the claim that the APE model is unfalsifiable.

The same considerations apply to more specific criticism that the APE model's mental-process account of EC is unfalsifiable. As should be clear from the current review, the APE model implies a wide range of predictions about the effects of CS-US co-occurrence, perceived validity, CS-US relations, and the conditions under which these effects should and should not occur. Although several of these predictions have been empirically confirmed (e.g., Hu et al., 2017b), others turned out to be inconsistent with the empirical evidence (e.g., Peters & Gawronski, 2011b). Again, the latter findings raise important questions about the validity of various core assumptions of the APE model. However, any such concerns debunk the criticism that the APE model is consistent with any empirical outcome, and thus unfalsifiable.

Based on the available evidence, two assumptions of the APE model should be treated with caution in the absence of further evidence. First, there is currently no compelling evidence that associative learning leads to EC effects in the absence of conscious awareness. Although this conclusion does not necessarily question the contribution of associative learning to EC, it does require a revision of the APE model's assumptions about the conditions under which associative learning occurs. That is, associative learning may create mental links between co-occurring stimuli only when people are aware of the relevant spatio-temporal contiguities. Second, counter to the APE model's assumption that conflicting effects of associative and propositional learning can be identified by means of dissociations between spontaneous and deliberate evaluations, such dissociations may be better understood as the result of expression-related processes rather than learning-related processes. Again, this conclusion does not necessarily question the contribution of associative learning to EC. However, it does require more advanced approaches to identify effects of associative learning. To the extent that alternative approaches fail to provide evidence for the formation of direct associative links during learning, we would agree with the proponents of single-process propositional theories that there is no need to assume a process of associative learning over and above the uncontroversial process of propositional learning. In this case, the APE model may still provide valuable information about the processes involved in the behavioral expression of evaluative representations. However, it would be less informative about the processes involved in the formation of evaluative representations.

Conclusion

The main goal of the current article was to review the core assumptions of the APE model pertaining to the mental processes by which pairings of a CS with a positive or negative US influence evaluations of the CS. According to the APE model, a sufficient understanding of EC effects requires a consideration of (1) associative and propositional learning as two functionally distinct learning mechanisms, (2) the mental representations resulting from the two learning mechanisms, (3) the role of associative activation and propositional validation in the behavioral expression of these representations, and (4) the automatic versus controlled nature of the processes involved in the formation and expression of mental representations. Although the APE model faces non-trivial challenges in explaining mixed findings regarding competing effects of associative and propositional learning, it provides a valuable framework for studying EC effects and the factors that influence the encoding of CS-US pairings and the expression of evaluative responses. By generating strong, falsifiable predictions about the determinants of EC, it has the potential to generate interesting new findings even if the theory itself may ultimately require revisions as a result of unexpected outcomes.

Funding

Preparation of this article was supported by National Science Foundation Grant # 1649900. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Competing Interests

The authors have declared that no competing interests exist.

Acknowledgements

The authors have no support to report.

References

Baddeley, A. (2010). Working memory. *Current Biology*, *20*(4), R136–R140. https://doi.org/10.1016/j.cub.2009.12.014

Balas, R., & Gawronski, B. (2012). On the intentional control of conditioned evaluative responses. *Learning and Motivation*, *43*(3), 89–98. https://doi.org/10.1016/j.lmot.2012.06.003

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (pp. 1-40). Hillsdale, NJ, USA: Lawrence Erlbaum Associates.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*(1), 116–131. https://doi.org/10.1037/0022-3514.42.1.116

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

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

Custers, R., & Aarts, H. (2011). Learning of predictive relations between events depends on attention, not on awareness. *Consciousness and Cognition*, *20*(2), 368–378. https://doi.org/10.1016/j.concog.2010.05.011

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. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior*, *37*(1), 1–20. https://doi.org/10.3758/LB.37.1.1

De Houwer, J. (2014). Why a propositional single-process model of associative learning deserves to be defended. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (pp. 530-541). New York, NY, USA: Guilford Press.

De Houwer, J., & Moors, A. (2015). Levels of analysis in social psychology. In B. Gawronski & G. V. Bodenhausen (Eds.), *Theory and explanation in social psychology* (pp. 24-40). New York, NY, USA: Guilford Press.

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*(6), 853–869. https://doi.org/10.1037/0033-2909.127.6.853

Deutsch, R., Gawronski, B., & Strack, F. (2006). At the boundaries of automaticity: Negation as reflective operation. *Journal of Personality and Social Psychology*, *91*(3), 385–405. https://doi.org/10.1037/0022-3514.91.3.385

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*(1), 5–18. https://doi.org/10.1037/0022-3514.56.1.5

Doumas, L. A. A., Hummel, J. E., & Sandhofer, C. M. (2008). A theory of the discovery and predication of relational concepts. *Psychological Review*, *115*(1), 1–43. https://doi.org/10.1037/0033-295X.115.1.1

Epstein, S., Pacini, M., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology*, *71*(2), 390–405. https://doi.org/10.1037/0022-3514.71.2.390

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*(6), 1013–1027. https://doi.org/10.1037/0022-3514.69.6.1013

Festinger, L. (1957). *A theory of cognitive dissonance.* Evanston, IL, USA: Row, Peterson & Company.

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

Galdi, S., Gawronski, B., Arcuri, L., & Friese, M. (2012). Selective exposure in decided and undecided individuals: Differential relations to automatic associations and conscious beliefs. *Personality and Social Psychology Bulletin*, *38*(5), 559–569. https://doi.org/10.1177/0146167211435981

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. (2012). Back to the future of dissonance theory: Cognitive consistency as a core motive. *Social Cognition*, *30*(6), 652–668. https://doi.org/10.1521/soco.2012.30.6.652

Gawronski, B., Balas, R., & Creighton, L. A. (2014). Can the formation of conditioned attitudes be intentionally controlled? *Personality and Social Psychology Bulletin*, *40*(4), 419–432. https://doi.org/10.1177/0146167213513907

Gawronski, B., Balas, R., & Hu, X. (2016). What do I know about it versus how do I feel about it? A theoretical analysis of response processes in evaluative conditioning. *Social Cognition*, *34*(5), 357–381. https://doi.org/10.1521/soco.2016.34.5.357

Gawronski, B., & Bodenhausen, G. V. (2006a). Associative and propositional 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. (2006b). Associative and propositional processes in evaluation: Conceptual, empirical, and meta-theoretical issues: Reply to Albarracín, Hart, and McCulloch (2006), Kruglanski and Dechesne (2006), and Petty and Briñol (2006). *Psychological Bulletin*, *132*(5), 745–750. https://doi.org/10.1037/0033-2909.132.5.745

Gawronski, B., & Bodenhausen, G. V. (2007). Unraveling the processes underlying evaluation: Attitudes from the perspective of the APE Model. *Social Cognition*, *25*(5), 687–717. https://doi.org/10.1521/soco.2007.25.5.687

Gawronski, B., & Bodenhausen, G. V. (2009). Operating principles versus operating conditions in the distinction between associative and propositional processes. *Behavioral and Brain Sciences*, *32*(2), 207–208. https://doi.org/10.1017/S0140525X09000958

Gawronski, B., & Bodenhausen, G. V. (2011). The associative-propositional evaluation model: Theory, evidence, and open questions. *Advances in Experimental Social Psychology*, *44*, 59–127. https://doi.org/10.1016/B978-0-12-385522-0.00002-0

Gawronski, B., & Bodenhausen, G. V. (2014). The associative-propositional evaluation model: Operating principles and operating conditions of evaluation. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (pp. 188-203). New York, NY, USA: Guilford Press.

Gawronski, B., & Bodenhausen, G. V. (2015). Theory evaluation. In B. Gawronski & G. V. Bodenhausen (Eds.), *Theory and explanation in social psychology* (pp. 3-23). New York, NY, USA: Guilford Press.

Gawronski, B., Brannon, S. M., & Bodenhausen, G. V. (2017). The associative-propositional duality in the representation, formation, and expression of attitudes. In R. Deutsch, B. Gawronski, & W. Hofmann (Eds.), *Reflective and impulsive determinants of human behavior* (pp. 103-118). New York, NY, USA: Psychology Press.

Gawronski, B., Peters, K. R., Brochu, P. M., & Strack, F. (2008). Understanding the relations between different forms of racial prejudice: A cognitive consistency perspective. *Personality and Social Psychology Bulletin*, *34*(5), 648–665. https://doi.org/10.1177/0146167207313729

Gawronski, B., & De Houwer, J. (2014). Implicit measures in social and personality psychology. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (2nd ed., pp. 283-310). New York, NY, USA: Cambridge University Press.

Gawronski, B., Deutsch, R., Mbirkou, S., Seibt, B., & Strack, F. (2008). When “just say no” is not enough: Affirmation versus negation training and the reduction of automatic stereotype activation. *Journal of Experimental Social Psychology*, *44*(2), 370–377. https://doi.org/10.1016/j.jesp.2006.12.004

Gawronski, B., Gast, A., & De Houwer, J. (2015). Is evaluative conditioning really resistant to extinction? Evidence for changes in evaluative judgments without changes in evaluative representations. *Cognition and Emotion*, *29*(5), 816–830. https://doi.org/10.1080/02699931.2014.947919

Gawronski, B., & LeBel, E. P. (2008). Understanding patterns of attitude change: When implicit measures show change, but explicit measures do not. *Journal of Experimental Social Psychology*, *44*(5), 1355–1361. https://doi.org/10.1016/j.jesp.2008.04.005

Gawronski, B., Mitchell, D. G. V., & Balas, R. (2015). Is evaluative conditioning really uncontrollable? A comparative test of three emotion-focused strategies to prevent the acquisition of conditioned preferences. *Emotion*, *15*(5), 556–568. https://doi.org/10.1037/emo0000078

Gawronski, B., Sherman, J. W., & Trope, Y. (2014). Two of what? A conceptual analysis of dual-process theories. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (pp. 3-19). New York, NY, USA: Guilford Press.

Gawronski, B., & Strack, F. (2004). On the propositional nature of cognitive consistency: Dissonance changes explicit, but not implicit attitudes. *Journal of Experimental Social Psychology*, *40*(4), 535–542. https://doi.org/10.1016/j.jesp.2003.10.005

Gawronski, B., & Walther, E. (2012). What do memory data tell us about the role of contingency awareness in evaluative conditioning? *Journal of Experimental Social Psychology*, *48*(3), 617–623. https://doi.org/10.1016/j.jesp.2012.01.002

Gawronski, B., Walther, E., & Blank, H. (2005). Cognitive consistency and the formation of interpersonal attitudes: Cognitive balance affects the encoding of social information. *Journal of Experimental Social Psychology*, *41*(6), 618–626. https://doi.org/10.1016/j.jesp.2004.10.005

Gibson, B. (2008). Can evaluative conditioning change attitudes toward mature brands? New evidence from the Implicit Association Test. *The Journal of Consumer Research*, *35*(1), 178–188. https://doi.org/10.1086/527341

Greenwald, A. G., McGhee, D. E., & Schwartz, J. K. L. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Gross, J. J. (1998). Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology*, *74*(1), 224–237. https://doi.org/10.1037/0022-3514.74.1.224

Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology*, *12*(4), 392–407. https://doi.org/10.1016/S0022-1031(76)80006-6

Hamilton, D. L., Katz, L. B., & Leirer, V. O. (1980). Cognitive representation of personality impressions: Organizational processes in first impression formation. *Journal of Personality and Social Psychology*, *39*(6), 1050–1063. https://doi.org/10.1037/h0077711

Hebb, D. O. (1949). *The organization of behavior*. New York, NY, USA: Wiley & Sons.

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

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

Hu, X., Gawronski, B., & Balas, R. (2017a). 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

Hu, X., Gawronski, B., & Balas, R. (2017b). Propositional versus dual-process accounts of evaluative conditioning: II. The effectiveness of counter-conditioning and counter-instructions in changing implicit and explicit evaluations. *Social Psychological & Personality Science*, *8*(8), 858–866. https://doi.org/10.1177/1948550617691094

Hütter, M., & Sweldens, S. (2018). Dissociating controllable and uncontrollable effects of affective stimuli on attitudes and consumption. *Journal of Consumer Research*, *45*, 320–349. https://doi.org/10.1093/jcr/ucx124

Jones, E. E., & Gerard, H. B. (1967). *Foundations of social psychology*. New York, NY, USA: Wiley & Sons.

Kattner, F. (2014). Reconsidering the (in)sensitivity of evaluative conditioning to reinforcement density and CS–US contingency. *Learning and Motivation*, *45*, 15–29. https://doi.org/10.1016/j.lmot.2013.09.002

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

Kendrick, R. V., & Olson, M. A. (2012). When feeling right leads to being right in the reporting of implicitly-formed attitudes, or how I learned to stop worrying and trust my gut. *Journal of Experimental Social Psychology*, *48*(6), 1316–1321. https://doi.org/10.1016/j.jesp.2012.05.008

Kosloff, S., Greenberg, J., Schmader, T., Dechesne, M., & Weise, D. (2010). Smearing the opposition: Implicit and explicit stigmatization of the 2008 U.S. presidential candidates and the current U.S. president. *Journal of Experimental Psychology. General*, *139*(3), 383–398. https://doi.org/10.1037/a0018809

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480

Kurdi, B., & Banaji, M. R. (2017). Repeated evaluative pairings and evaluative statements: How effectively do they shift implicit attitudes? *Journal of Experimental Psychology. General*, *146*(2), 194–213. https://doi.org/10.1037/xge0000239

Langer, T., Walther, E., Gawronski, B., & Blank, H. (2009). When linking is stronger than thinking: Associative transfer of valence disrupts the emergence of cognitive balance after attitude change. *Journal of Experimental Social Psychology*, *45*(6), 1232–1237. https://doi.org/10.1016/j.jesp.2009.07.005

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

Mierop, A., Hütter, M., & Corneille, O. (2017). Resource availability and explicit memory largely determine evaluative conditioning effects in a paradigm claimed to be conducive to implicit attitude acquisition. *Social Psychological & Personality Science*, *8*(7), 758–767. https://doi.org/10.1177/1948550616687093

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

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. (2015). Processing goals moderate the effect of co-occurrence on automatic evaluation. *Journal of Experimental Social Psychology*, *60*, 157–162. https://doi.org/10.1016/j.jesp.2015.05.009

Neuberg, S. L., & Newsom, J. T. (1993). Personal need for structure: Individual differences in the desire for simpler structure. *Journal of Personality and Social Psychology*, *65*(1), 113–131. https://doi.org/10.1037/0022-3514.65.1.113

Payne, B. K., Cheng, S. 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

Peters, K. R., & Gawronski, B. (2011a). Mutual influences between the implicit and explicit self-concepts: The role of memory activation and motivated reasoning. *Journal of Experimental Social Psychology*, *47*(2), 436–442. https://doi.org/10.1016/j.jesp.2010.11.015

Peters, K. R., & Gawronski, B. (2011b). Are we puppets on a string? Comparing the impact of contingency and validity on implicit and explicit evaluations. *Personality and Social Psychology Bulletin*, *37*(4), 557–569. https://doi.org/10.1177/0146167211400423

Pleyers, G., Corneille, O., Yzerbyt, V., & Luminet, O. (2009). Evaluative conditioning may incur attentional costs. *Journal of Experimental Psychology: Animal Behavior Processes*, *35*(2), 279–285. https://doi.org/10.1037/a0013429

Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology*, *91*(6), 995–1008. https://doi.org/10.1037/0022-3514.91.6.995

Rydell, R. J., McConnell, A. R., Mackie, D. M., & Strain, L. M. (2006). Of two minds: Forming and changing valence-inconsistent implicit and explicit attitudes. *Psychological Science*, *17*(11), 954–958. https://doi.org/10.1111/j.1467-9280.2006.01811.x

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*(2), 108–131. https://doi.org/10.1207/S15327957PSPR0402_01

Stahl, C., Haaf, J., & Corneille, O. (2016). Subliminal evaluative conditioning? Above-chance CS identification may be necessary and insufficient for attitude learning. *Journal of Experimental Psychology. General*, *145*(9), 1107–1131. https://doi.org/10.1037/xge0000191

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*(3), 220–247. https://doi.org/10.1207/s15327957pspr0803_1

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

Sweldens, S., Van Osselaer, S., & Janiszewski, C. (2010). Evaluative conditioning procedures and the resilience of conditioned brand attitudes. *The Journal of Consumer Research*, *37*(3), 473–489. https://doi.org/10.1086/653656

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

Walther, E., Nagengast, B., & Trasselli, C. (2005). Evaluative conditioning in social psychology: Facts and speculations. *Cognition and Emotion*, *19*(2), 175–196. https://doi.org/10.1080/02699930441000274

Weary, G., & Edwards, J. A. (1994). Individual differences in causal uncertainty. *Journal of Personality and Social Psychology*, *67*(2), 308–318. https://doi.org/10.1037/0022-3514.67.2.308

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review*, *101*(1), 34–52. https://doi.org/10.1037/0033-295X.101.1.34

Weinberger, J., & Westen, D. (2008). RATS, we should have used Clinton: Subliminal priming in political campaigns. *Political Psychology*, *29*(5), 631–651. https://doi.org/10.1111/j.1467-9221.2008.00658.x

Zanon, R., De Houwer, J., & Gast, A. (2012). Context effects in evaluative conditioning of implicit evaluations. *Learning and Motivation*, *43*(3), 155–165. https://doi.org/10.1016/j.lmot.2012.02.003

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

PsychOpen GOLD