Social and Personality Psychology Compass 8/8 (2014): 448–462, 10.1111/spc3.12124

Implicit and Explicit Evaluation: A Brief Review of the Associative–Propositional Evaluation Model

Bertram Gawronski[1]* and Galen V. Bodenhausen[2]
[1]*University of Texas at Austin*
[2]*Northwestern University*

Abstract

A central theme in contemporary psychology is the distinction between implicit and explicit evaluation. Research has shown various dissociations between the two kinds of evaluations, including different antecedents, different consequences, and discrepant evaluations of the same object. The current article provides a brief review of the associative–propositional evaluation (APE) model, which accounts for these dissociations by conceptualizing implicit and explicit evaluations as the behavioral outcomes of two functionally distinct, yet mutually interacting, mental processes. Whereas implicit evaluations are assumed to be the outcome of associative processes, explicit evaluations are conceptualized as the outcome of propositional processes. Associative processes determine the activation of mental contents on the basis of feature similarity and spatiotemporal contiguity; propositional processes involve the validation of activated mental contents on the basis of cognitive consistency. The APE model includes specific assumptions about mutual interactions between the two processes, implying precise predictions about converging versus diverging patterns of implicit and explicit evaluation.

Conflicts between the "head" and the "heart" can be rather common in everyday life. We may feel romantically attracted to a particular person despite firmly believing that this person is not a good match; and the sight of a high-calorie dessert may elicit an impulse to indulge although we know that it is unhealthy and detrimental for our goal to lose weight. Conversely, we may experience feelings of apprehension and discomfort when encountering members of stigmatized groups even though we intellectually abhor prejudice and wish to express solidarity with minorities; and a small spider may elicit an unpleasant fright response although we know that it is entirely harmless. Over the last two decades, psychologists have gained valuable insights into the causes and consequences of such evaluative conflicts by comparing verbal judgments on traditional self-report measures (e.g., attitude scales, likeability ratings) to spontaneous responses on performance-based paradigms (e.g., sequential priming tasks, implicit association test).[1] Conceptually, evaluative judgments on the former type of measures can be described as *explicit evaluations* in the sense that their evaluative meaning is explicit in the observed response (e.g., participants explicitly report their agreement or disagreement with an evaluative statement about an object). Conversely, spontaneous responses on the latter type of measures can be described as *implicit evaluations* in the sense that their evaluative meaning is implicit in the observed response (e.g., evaluative responses are inferred from participants' latencies in responding to positive and negative words that are preceded by brief presentations of an object).

The main goal of the current article is to provide a brief review of the associative–propositional evaluation (APE) model, which explains dissociations between implicit and explicit evaluations in terms of two functionally distinct mental processes (Gawronski & Bodenhausen, 2006, 2007, 2011). By specifying the mutual interplay between these processes, the APE model provides an overarching framework that specifies when implicit and explicit

evaluations should be related and when they should be unrelated. In addition, the APE model offers precise predictions about the conditions under which a given factor should lead to (a) changes in implicit but not explicit evaluations, (b) changes in explicit but not implicit evaluations, or (c) corresponding changes in implicit and explicit evaluations. Over the past years, research drawing on the core assumptions of the APE model has provided valuable insights in a wide range of areas, including romantic relationships (e.g., Eastwick, Eagly, Finkel, & Johnson, 2011), prejudice and stereotyping (e.g., Gawronski, Peters, Brochu, & Strack, 2008), social justice (e.g., Van den Bos & Maas, 2009), consumer behavior (e.g., Redker & Gibson, 2009), alcohol consumption (e.g., Moss & Albery, 2009), psychopathology (e.g., Ouimet, Gawronski, & Dozois, 2009), political decision-making (Galdi, Arcuri, & Gawronski, 2008), eating behavior (e.g., Hollands, Prestwich, & Marteau, 2011), and media effects (Strick, Holland, Van Baaren, Van Knippenberg, & Dijksterhuis, 2013). In the current article, we review the core assumptions of the APE model and its predictions regarding changes in implicit and explicit evaluations. In addition, we discuss a few questionable assumptions about implicit and explicit evaluations and the nature of their underlying mental entities.

Associative and Propositional Processes

A central assumption of the APE model is that implicit evaluations reflect the behavioral outcome of *associative processes*, whereas explicit evaluations are the behavioral outcome of *propositional processes*. Associative processes are defined as the *activation* of mental associations in memory, which we assume to be driven by the principles of feature matching and spatiotemporal contiguity. Propositional processes are defined as the *validation* of the information implied by activated associations, which we assume to be guided by the principles of cognitive consistency. Although associative and propositional processes are functionally distinct, the two processes are assumed to mutually interact, such that associative processes may influence propositional processes, and vice versa.

Associative activation

According to the APE model, associative processes are essential for understanding any kind of evaluation, because they determine which mental contents are activated in response to an object. A central factor in this process is the similarity between the features of input stimuli and existing representations. Specifically, we assume that principles of feature matching influence the activation of mental concepts that represent the encountered stimulus (e.g., features of a face activating the concept *African American*), which may spread to other concepts that are associatively linked with the activated concepts in memory (e.g., activation of the concept *African American* spreading to the associated stereotypical attribute *hostile*). To the extent that the associated concepts have, in the aggregate, a positive or negative connotation, their activation is assumed to elicit a spontaneous affective response that is in line with the net valence of these concepts (*implicit evaluation*).

An important aspect of feature matching is that stimuli do not have to be perceptually identical across time and contexts to elicit the same evaluative response. Instead, configurations of input stimuli that pass a critical threshold of similarity are sufficient to activate the same mental representation (Smith, 1996). Thus, even unknown stimuli may elicit spontaneous affective responses to the extent that they resemble a previously encountered stimulus with a stored evaluative representation. In line with this contention, Gawronski and Quinn (2013) showed that faces of unknown individuals can elicit spontaneous positive or negative reactions by virtue of their resemblance to known faces. In fact, implicit evaluations of unknown faces were indistinguishable from implicit evaluations of the known faces they resembled, suggesting that

unfamiliar faces are assimilated to existing representation of known faces (see also Duckworth, Bargh, Garcia, & Chaiken, 2002).

Another important aspect of associative activation is that it is not an all-or-none process, such that encountering a given object would activate each and every concept that is associated with that object in memory. Instead, objects tend to activate only a limited subset of associated concepts. Which subset is activated in response to an object is assumed to be constrained by the overall configuration of input stimuli, including both the target object and the context in which it is encountered. For example, encountering an African American man in a jazz bar may activate the stereotypical attribute *musical*, whereas the same African American man may activate the stereotypical attribute *criminal* if he is encountered in a dark alley. Hence, a stimulus may elicit distinct implicit evaluations depending on the particular context in which it is encountered (e.g., Barden, Maddux, Petty, & Brewer, 2004; Wittenbrink, Judd, & Park, 2001; for a review, see Gawronski & Sritharan, 2010).

Such context effects are not limited to environmental cues with a clear semantic relation to the mental concepts that are associated with a stimulus (e.g., semantic relation between *jazz bar* and the stereotypical attribute *musical*); they may also involve incidental visual cues that simply happened to be present during the formation of evaluative associations (e.g., perceptual features of a room). Consistent with this assumption, Gawronski, Rydell, Vervliet, and De Houwer (2010) have shown that expectancy-violating counterattitudinal experiences enhance attention to incidental features of the environmental context, thereby leading to an integration of these context cues into the mental representation of the counterattitudinal experience. As a result, counterattitudinal experiences influence subsequent implicit evaluations only in the context in which these experiences occurred, whereas initial attitudinal experiences continue to influence implicit evaluations in any other context (e.g., Rydell & Gawronski, 2009; for a review, see Gawronski & Cesario, 2013).

Propositional validation

A central aspect of associative activation is that it is independent of subjective truth or falsity. Specifically, we assume that mentally associated concepts can be activated regardless of whether the relation implied by the activated link is considered valid or invalid. For example, encountering a Muslim-looking man may activate the concept *terrorism* even if a person rejects the implied connection between Muslims and terrorism (Devine, 1989). According to the APE model, the subjective validity of activated links is determined by a process of propositional validation. Specifically, we assume that activated information is regarded as valid unless the default process of affirming the validity of activated information produces an inconsistent set of beliefs. The central idea underlying these assumptions 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 momentarily considered set of information needs to be updated, which requires a reassessment of the validity of each component.

With regard to evaluative responses, we assume that the affective reactions resulting from associatively activated concepts are translated into the format of a propositional statement (e.g., a negative affective reaction toward object X is transformed into propositional statements such as "X is bad" or "I dislike X"). To the extent that the propositional evaluation implied by one's spontaneous affective response is consistent with other momentarily considered propositions, it may be endorsed in a verbal evaluative judgment (*explicit evaluation*). If, however, the overall set of momentarily considered propositions is inconsistent, the inconsistency has to be resolved to avoid aversive feelings of

dissonance (Festinger, 1957). In general, propositional evaluations of a given object may be assessed for their consistency with (a) non-evaluative propositions about states of affairs and (b) propositional evaluations of other attitude objects (Jones & Gerard, 1967). To the extent that a set of momentarily considered propositions is inconsistent, consistency may be restored either by rejecting one of the involved propositions (i.e., reversing the subjective truth value of that proposition) or by searching for an additional proposition that resolves the inconsistency (Gawronski & Strack, 2004).

To illustrate the role of consistency in the process of propositional validation, consider a case in which the propositional implication of a negative affective reaction to a social group (e.g., *I dislike African Americans*) is inconsistent with the joint implication of a propositional evaluation of another attitude object (e.g., *evaluations of disadvantaged groups are wrong*) and a non-evaluative proposition about states of affairs (e.g., *African Americans are a disadvantaged group*) (Figure 1, panel A).[2] In this case, the inconsistency between the three propositions may lead to a rejection of the negative affective reaction as a basis for an evaluative judgment (e.g., *I like African Americans*) (Figure 1, panel B). However, consistency may also be restored by rejecting either the propositional evaluation of another relevant attitude object (e.g., *Negative evaluations of disadvantaged groups are okay*) (Figure 1, panel C) or the non-evaluative proposition about states of affairs (e.g., *African Americans are not a disadvantaged group*) (Figure 1, panel D). According to the APE model, implicit and explicit evaluations should be unrelated when inconsistency leads to a rejection of the negative affective reaction as a basis for an evaluative judgment. In contrast, the two kinds of evaluations should be positively correlated when inconsistency is resolved by rejecting any of the other two propositions (e.g., Brochu, Gawronski, & Esses, 2011; Gawronski, Peters, et al., 2008).

In addition to such "bottom-up" effects of associative on propositional processes, the APE model also includes precise assumptions about "top-down" effects of propositional on associative processes. Specifically, we assume that processes of propositional reasoning can influence associative processes by activating new information in the course of validating activated information. For example, if people are motivated to believe in the validity of a particular proposition, they may engage in a selective search for information that supports the validity of that proposition (Kunda, 1990). In such cases, biased retrieval of confirmatory information can activate associated concepts of a particular valence, which produces correspondence between implicit and explicit evaluations in a "top-down" fashion (e.g., Peters & Gawronski, 2011; see also Galdi, Gawronski, Arcuri, & Friese, 2012).

An important factor in such top-down effects is whether propositional reasoning involves an affirmation or negation of the relevant information. Specifically, we argue that merely negating a particular proposition (i.e., reversing its truth value) is insufficient to deactivate the associated concepts underlying this proposition. In fact, negations often lead to ironic effects, such that the activation level of the underlying association is enhanced rather than reduced (e.g., Gawronski, Deutsch, Mbirkou, Seibt, & Strack, 2008). For example, negating the proposition *old people are bad drivers* may enhance the associative link between the concepts *old people* and *bad drivers*, thereby leading to a dissociation between implicit and explicit evaluations (Deutsch, Gawronski, & Strack, 2006). This situation is different if processes of propositional reasoning involve an affirmation of new information. For example, affirming the proposition *old people are good drivers* may strengthen the association between the concepts *old people* and *good drivers*, thereby increasing the correspondence between implicit and explicit evaluations. Thus, if a person is motivated to hold a positive impression of an attitude object (e.g., person or social group) but experiences a negative affective reaction toward that object, the individual may engage in a directed memory search to retrieve positive information about the object, which should promote a positive evaluation for both explicit and implicit evaluations. If, however,

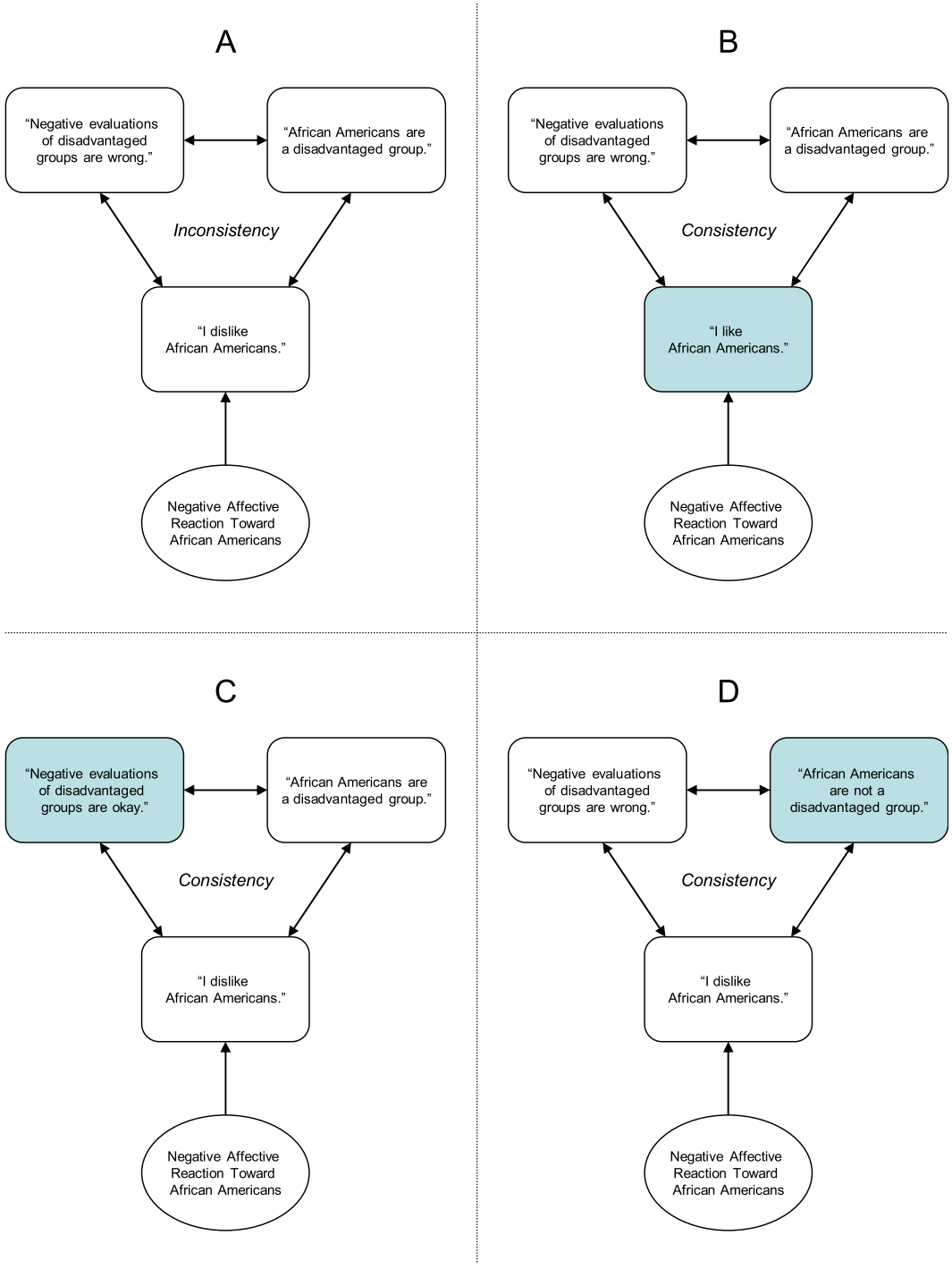

Figure 1 Interplay between affective reactions (circles) and propositional beliefs (squares) in racial prejudice against African Americans. Panel A depicts the case of an inconsistent belief system; panels B, C, and D depict consistent belief systems, involving either a rejection of affective reactions for evaluative judgments (panel B) or a reliance on affective reactions for evaluative judgments (panels C and D). Figure adapted from Gawronski, Peters, et al. (2008); reprinted with permission.

the positive impression is maintained by merely negating the negative evaluation implied by the affective response (i.e., without retrieving supportive positive information), explicit and implicit evaluations should show a dissociation, such that explicit evaluations reflect the desired positive evaluation, whereas implicit evaluations should reflect the original negative response. The same predictions apply to cases when someone is motivated to hold a negative impression of an attitude object but experiences an undesired positive reaction toward that object.

Associative and propositional learning

The preceding sections focused primarily on the role of associative and propositional processes during the activation of existing representations and their use in verbally reported judgments. However, it is also important to understand the psychological processes through which such representations are initially established. In the APE model, we distinguish between two conceptually distinct processes of forming evaluative representations, depending on whether they are based on associative or propositional principles. Drawing on our general definition of associative processes, *associative learning* can be described as the formation of associative links between mental concepts on the basis of spatiotemporal contiguities in the environment. The central assumption underlying this definition is that observed co-occurrences between objects and events result in a co-activation of their corresponding mental concepts, which in turn creates an associative link between the two. Repeatedly observing the same co-occurrences strengthens this link, which facilitates the spread of activation from one concept to the other upon encountering one of the two associated stimuli. An illustrative example of associative learning is the formation of a mental link between a conditioned stimulus (CS) and a positive or negative unconditioned stimulus (US) due to repeated pairings of the two stimuli. Such associatively formed links can be a source of evaluative conditioning effects (for a meta-analysis, see Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010) to the extent that subsequent presentations of the CS spread to the representation of the US, thereby eliciting an evaluative response to the CS that is in line with the valence of the US (e.g., Walther, Gawronski, Blank, & Langer, 2009).

In contrast to the associative principle of mere co-activation, *propositional learning* is defined as the formation of evaluative representations on the basis of propositional information that is regarded as valid. This definition is based on our conceptualization of propositional processes as being concerned with the validity of momentarily activated information. Propositional learning may be based on new information that is presented in the format of propositional statements (e.g., persuasive arguments asserting the quality of a consumer product). Alternatively, propositional learning can be based on propositional inferences about observed stimulus events in the environment (e.g., co-occurrences between stimulus events can trigger propositional inferences about their causal relation). Whereas the former case involves the acquisition of externally provided propositional information, the latter case involves the acquisition of self-generated propositional information. Yet, in both cases, the new information has to pass a process of propositional validation. This validity assessment is assumed to be equivalent to the one involved in the endorsement of evaluative judgments, such that new propositional information may be regarded as either true or false depending on its consistency with other momentarily considered propositions.

Although associative and propositional learning represent distinct mechanisms of forming evaluative representations, their outcomes are assumed to interact in a manner that is similar to the mutual interactions in the endorsement of evaluative judgments. First, associatively formed representations may provide the input for propositional inferences, implying a "bottom-up" effect of associative on propositional processes. Thus, whether or not the evaluation implied by an associatively formed representation is regarded as valid depends on

the consistency of this evaluation with other momentarily considered propositions (e.g., Gawronski & LeBel, 2008). Second, propositional processes may influence associative processes in a "top-down" fashion when externally provided or self-generated propositions create new mental links in memory (e.g., Whitfield & Jordan, 2009). As we outlined above, an important determinant of such top-down effects is whether the involved inferences involve an affirmation or negation of the relevant information. Whereas the affirmation of a given proposition should create an evaluative representation that is in line with the meaning of that proposition, negating a given proposition is assumed to have ironic effects.

An illustrative example of the interplay of associative and propositional processes during learning is the difficulty of creating advertisements for products that counteract something negative (e.g., pharmaceutical products, insurance policies). A major challenge in these advertisements is that the product can become directly associated with the negative phenomenon it is supposed to eliminate, thereby eliciting a negative affective response to the product even when propositional inferences suggest a positive evaluation. A recent study by Moran and Bar-Anan (2013) illustrates this problem. In their study, participants were presented with neutral stimuli that started or stopped either pleasant or unpleasant sounds. Consistent with the notion of propositional learning, participants showed more favorable explicit evaluations of stimuli that started pleasant sounds compared with stimuli that started unpleasant sounds. Conversely, participants showed more favorable explicit evaluations of stimuli that stopped unpleasant sounds compared with stimuli that stopped pleasant sounds. In contrast, implicit evaluations showed a pattern consistent with the notion of associative learning, reflecting the mere co-occurrence of the target stimuli with pleasant and unpleasant sounds. That is, participants showed more favorable responses to stimuli that co-occurred with pleasant sounds compared with stimuli that co-occurred with negative sounds regardless of whether they started or stopped the sounds.

Changes in Implicit and Explicit Evaluations

Moran and Bar-Anan's (2013) study illustrates how the same learning experience can have different effects on implicit and explicit evaluations. By specifying the mutual interplay between associative and propositional processes, the APE model implies precise predictions about the conditions under which a given factor should lead to (a) changes in implicit but not explicit evaluations, (b) changes in explicit but not implicit evaluations, and (c) corresponding changes in implicit and explicit evaluations (Gawronski & Bodenhausen, 2006). According to the APE model, the emergence of these patterns depends on (a) which of the two kinds of processes – associative or propositional – is *directly* influenced in the first place, and (b) whether changes in one process lead to *indirect* changes in the other process.

As for direct influences on associative processes, we have already outlined how (a) contextual cues constrain the activation of associated concepts in response to a stimulus and (b) repeated co-occurrences in the environment may create new associative links between concepts in memory. Either of these two factors (i.e., presence of contextual cues, repeated co-occurrences of stimuli) can directly influence the mental concepts that are activated in response to a given stimulus. In addition, the APE model implies that an external factor may directly influence propositional processes when this factor leads to (a) a change in the preferred strategy to achieve consistency or (b) a change in the considered set of propositions. The latter case can be further divided into cases that involve the acquisition of new propositional information and cases that involve the retrieval of previously stored information from memory.

As for indirect influences, it is essential to consider the hypothesized conditions of "bottom-up" effects of associative on propositional processes and "top-down" effects of propositional on associative processes. Bottom-up effects are assumed to occur when

the propositional evaluation implied by a spontaneous affective response is consistent with the set of momentarily considered propositions. However, bottom-up effects are assumed to be disrupted when inconsistency within the set of considered propositions leads to a rejection of the spontaneous affective response. Conversely, top-down effects are assumed to occur when propositional inferences imply an affirmation of new information. However, propositional processes are assumed to leave the activation of associations unaffected (or produce ironic effects) when they involve a negation of activated information. On the basis of these principles, it is possible to identify four basic cases of how a given factor may influence implicit and explicit evaluations of an object (Figure 2).

The first case involves a direct influence on the activation of associations in memory, with the evaluation implied by the resulting affective response being accepted by the propositional validation process (Figure 2, panel A). According to the APE model, such cases should lead to corresponding changes in implicit and explicit evaluations, with changes in explicit evaluations being mediated by changes in implicit evaluations (described as Case 1 by Gawronski & Bodenhausen, 2006). An illustrative example of this case is research on evaluative conditioning, showing that repeated CS–US pairings produce corresponding effects on implicit and explicit

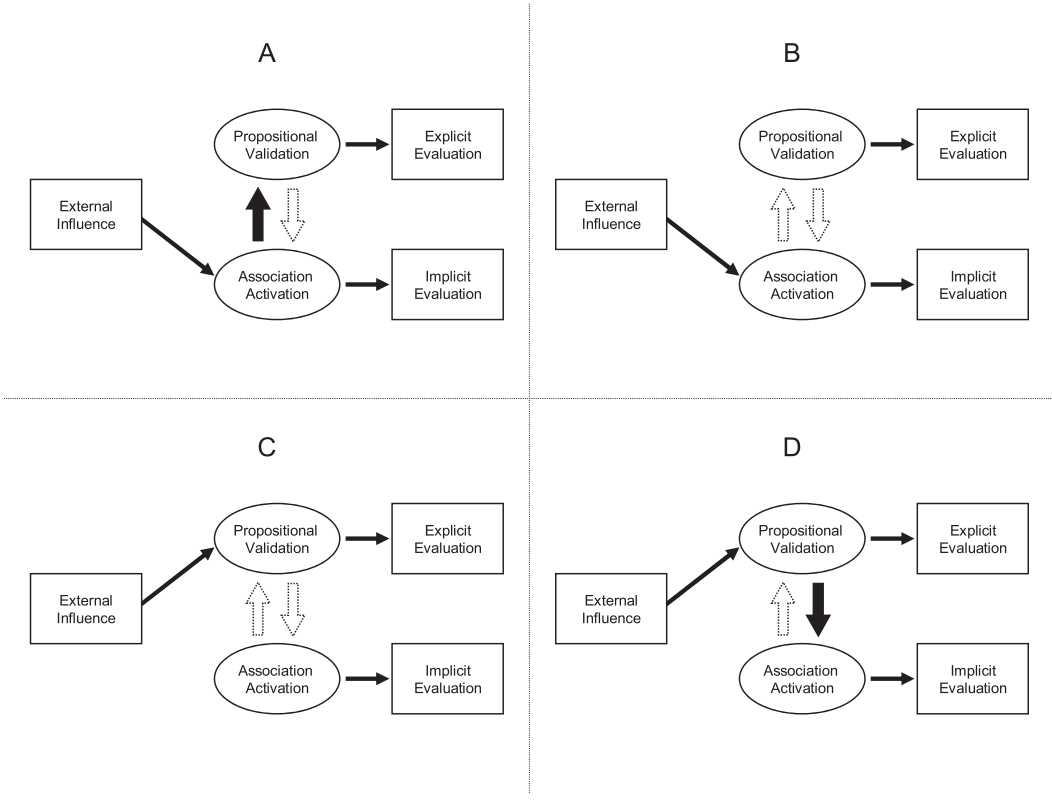

Figure 2 Potential direct and indirect influences of an external factor on associative and propositional processes and their resulting effects on implicit and explicit evaluations. Thin arrows depict direct effects of an external factor on either associative or propositional processes and influences of the two processes on implicit and explicit evaluations, and fat arrows depict mutual influences between associative and propositional processes, with solid arrows depicting the presence of an effect and open arrows the absence of an effect. Figure adapted from Gawronski and Bodenhausen (2011); reprinted with permission.

evaluations when participants focus on their feelings while making an evaluative judgment about the CS (e.g., Gawronski & LeBel, 2008; Whitfield & Jordan, 2009).[3]

The second case involves a direct influence on the activation of associations in memory, with the evaluation implied by the resulting affective response being rejected by the propositional validation process (Figure 2, panel B). According to the APE model, such cases should lead to changes in implicit, but not explicit, evaluations (described as Case 2 by Gawronski & Bodenhausen, 2006). An illustrative example of this case is research showing evaluative conditioning effects on implicit, but not explicit, evaluations when participants reflect on their prior knowledge about the CS (e.g., Gawronski & LeBel, 2008; Grumm, Nestler, & von Collani, 2009).

The third case involves a direct influence on the propositional validation process, with the activation of associations being unaffected by propositional reasoning (Figure 2, panel C). According to the APE model, such cases should lead to changes in explicit, but not implicit, evaluations (described as Case 3 by Gawronski & Bodenhausen, 2006). An illustrative example is research showing that cognitive dissonance resulting from induced compliance (e.g., Gawronski & Strack, 2004) and mere invalidation (or negation) of previously acquired information (e.g., Gregg, Seibt, & Banaji, 2006) produces changes in explicit, but not implicit, evaluations.

Finally, the fourth case involves a direct effect on the propositional validation process, with the activation of associations being influenced in line with the outcome of propositional reasoning (Figure 2, panel D). According to the APE model, such cases should lead to corresponding changes in implicit and explicit evaluations, with changes in implicit evaluations being mediated by changes in explicit evaluations (described as Case 4 by Gawronski & Bodenhausen, 2006). An illustrative example is research showing that newly acquired propositional information (e.g., about positive or negative behaviors performed by a target person) leads to corresponding changes in implicit and explicit evaluations, with changes in implicit evaluations being mediated by changes in explicit evaluations (e.g., Gawronski & Walther, 2008; Whitfield & Jordan, 2009).[4]

Qualifying Some Common Assumptions

By providing an overarching framework that specifies the relation between implicit and explicit evaluations, the APE model also qualifies some common assumptions about the two kinds of evaluations and the nature of their underlying mental entities. These issues involve (a) the presumed unconsciousness of the representations underlying implicit evaluations, (b) the impact of cognitive elaboration on the relation between implicit and explicit evaluations, (c) motivational effects on the two kinds of evaluation, and (d) the relation between the associative–propositional duality and the distinction between automatic and controlled processes.

Consciousness

A common assumption in research using performance-based measures is that they provide access to unconscious mental representations (e.g., Greenwald & Banaji, 1995). This assumption is based on the methodological fact that performance-based measures do not require introspective access for the measurement of mental contents. However, it is important to note that this characteristic does not guarantee that evaluations captured by performance-based measures indeed reflect unconscious representations. Any such claim represents an empirical hypothesis that has to be evaluated on the basis of available data (Gawronski, Hofmann, & Wilbur, 2006). In response to this concern, proponents of the unconsciousness hypothesis often refer to the finding that implicit and explicit evaluations tend to show rather low correlations (for a meta-analysis, see Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005), which is consistent with the claim that performance-based measures capture unconscious representations that

are not accessible for verbal self-reports. Yet, implicit and explicit evaluations can show low correlations for numerous reasons that have nothing to do with lack of introspective access (Gawronski, LeBel, & Peters, 2007). From the perspective of the APE model, dissociations between implicit and explicit evaluations do not reflect unconscious underpinnings of implicit evaluations, but the perceived invalidity of spontaneous affective reactions for verbally endorsed evaluative judgments. As such, dissociations between implicit and explicit evaluations fail to provide an unambiguous criterion for claims about unconscious representations (Hahn & Gawronski, 2014). A more stringent criterion is participants' ability to predict their responses on measures of implicit evaluation. In fact, such predictions tend to show very high levels of accuracy, with mean correlations between predicted and actual scores higher than 0.50 and median correlations in the range of 0.70 (Hahn, Judd, Hirsh, & Blair, in press). These findings are difficult to reconcile with claims that implicit evaluations reflect unconscious representations that are not accessible to introspection.

Cognitive elaboration

Another common assumption is that the correspondence between implicit and explicit evaluations should decrease as a function of cognitive elaboration during the generation of an evaluative judgment (e.g., Fazio, 2007). The APE model agrees with this general claim, albeit with some important qualifications. In our view, the correspondence between implicit and explicit evaluations primarily depends on the (in)consistency of the evaluation implied by one's affective response with other judgment-relevant information. To the extent that the likelihood of inconsistency increases as a function of the amount of information that is considered, higher levels of cognitive elaboration may reduce the correspondence between implicit and explicit evaluations. However, what ultimately reduces the correspondence between implicit and explicit evaluations is not cognitive elaboration per se but the inconsistency of the affective response with other momentarily considered information. Thus, if higher levels of cognitive elaboration lead to a consideration of information that is consistent with the affective response, the correspondence between implicit and explicit evaluations may be unaffected. Moreover, if extensive elaboration helps to identify information that supports the validity of an affective response, enhanced elaboration may increase rather than decrease the correspondence between implicit and explicit evaluations (Galdi et al., 2012). In other words, what moderates the relation between implicit and explicit evaluations is not cognitive elaboration per se but the consistency of additionally recruited information with the evaluation implied by one's spontaneous affective response.

Motivational influences

Similar to the role of cognitive elaboration, it is often assumed that the relation between implicit and explicit evaluations is moderated by motivational factors. For example, it has been argued that implicit evaluations reveal thoughts, opinions, or feelings that people are unwilling to report because of self-presentation or social desirability concerns (for a discussion, see Gawronski et al., 2007). Consistent with this assumption, several studies found that individuals with a low motivation to control prejudiced reactions tend to show positive correlations between implicit and explicit evaluations of minority groups, whereas individuals with a high motivation to control prejudiced reactions show either no or negative correlations (e.g., Dunton & Fazio, 1997; Gawronski, Geschke, & Banse, 2003). According to the APE model, motivational factors can certainly play a role in this regard, but their impact is assumed to be more distal and mediated by cognitive processes of consistency assessment. To illustrate this assumption, consider the

previous example of racial prejudice in which the activation of negative associations was assumed to elicit a negative affective response to African Americans. As outlined above, the propositional evaluation implied by this response may be inconsistent with other propositional beliefs, such as *African Americans are a disadvantaged group* and *Negative evaluations of disadvantaged groups are wrong* (Figure 1, panel A). From the perspective of the APE model, the latter proposition may be interpreted as a propositionally represented goal to the extent that an individual is committed to the action implied by the evaluation (i.e., *I don't want to evaluate disadvantaged groups negatively*). Yet, it may not necessarily lead to a rejection of the negative affective response as a basis for an evaluative judgment, if the inconsistency between the three propositions is resolved by denying the discrimination experienced by African Americans (Figure 1, panel D). In line with this assumption, we repeatedly found that high levels of the motivation to control prejudiced reactions reduced the correlation between implicit and explicit evaluations of minority groups only when these groups were perceived to be a target of discrimination (for a review, see Gawronski, Brochu, Sritharan, & Strack, 2012). When perceptions of discrimination were low, individuals with a high motivation to control prejudiced reactions generally showed a positive correlation between explicit and implicit evaluations. These results suggest that the moderating effect of motivational factors – such as the motivation to control prejudiced reactions – on the relation between implicit and evaluations is more distal, in that their impact depends on the consistency of all propositional beliefs that are considered for an evaluative judgment, including propositionally represented goals, propositional evaluations implied by affective gut responses, and any other evaluative or non-evaluative information that may be regarded as relevant.

Automaticity and control

A final issue concerns the relation between the associative–propositional duality and the distinction between automatic and controlled processes (Gawronski, Sherman, & Trope, 2014). Although the two dichotomies are often assumed to overlap, the APE model treats them as conceptually distinct dimensions (Gawronski & Bodenhausen, 2007, 2009). Our central argument is that the associative–propositional distinction refers to the operating principles that define *what* a particular process is doing (i.e., activation, validation). In contrast, the automatic–controlled distinction refers to operating conditions that specify *when* a given process is operating (i.e., when there is no conscious awareness; when there is no goal to start the process; when cognitive resources are reduced; when there is a goal to alter or stop the process). According to the APE model, there is no one-to-one mapping between operating principles and operating conditions, such that associative processes generally operate automatically, whereas propositional processes generally operate in a controlled fashion. Instead, both associative and propositional processes have automatic and controlled aspects. Moreover, each type of process involves different components, which require separate consideration in a thorough analysis of operating conditions. Because different features of automatic processing do not necessarily co-occur (Bargh, 1994), we also deem it important to distinguish between the unique roles of awareness, intentionality, efficiency, and controllability (Gawronski & Bodenhausen, 2007, 2011). An illustrative example is the assumption that associative processes are generally unintentional, which is qualified by the fact that associations can be activated intentionally through motivated retrieval processes (e.g., Peters & Gawronski, 2011). Similarly, generalized claims about the resource-dependence of propositional processes should be treated with caution, because the amount of resources required by propositional processes depends on the amount of information that is considered and the complexity of the involved inferences (e.g., Richter, Schroeder, & Wöhrmann, 2009). Although a thorough discussion of these issues

goes beyond the scope of this article (for more details, see Gawronski & Bodenhausen, 2014), the APE model explicitly rejects conceptual equations of *associative* with *automatic* and conceptual equations of *propositional* with *controlled*.

Conclusions

Implicit and explicit evaluations have been demonstrated to make unique contributions to the prediction of behavior (Friese, Hofmann, & Schmitt, 2008; Perugini, Richetin, & Zogmaister, 2010). These insights have led to increased interest in the causal antecedents of the two kinds of evaluations as well as the conditions under which they reflect corresponding or diverging responses. The APE model offers answers to both questions by specifying the mental processes involved in the generation of spontaneous affective reactions (*implicit evaluation*) and verbal evaluative judgments (*explicit evaluation*). Over the past years, the APE model has inspired an impressive body of research in a wide range of areas, and we are excited about novel applications in the years to come.

Short Biographies

Bertram Gawronski, PhD, is Professor of Psychology at the University of Texas at Austin. Before he moved to Austin, he held positions as Professor of Psychology and Canada Research Chair in Social Psychology at The University of Western Ontario (Canada). He earned his MA in Philosophy at the Free University Berlin (Germany) and his PhD in Psychology at Humboldt University Berlin (Germany). His research investigates the mental underpinnings and behavioral consequences of spontaneous and deliberate evaluations of objects, individuals, groups, and social issues. Previous and ongoing work includes projects on attitude formation and change, context effects on evaluative responses, evaluative conditioning, cognitive consistency, prejudice and stereotyping, impression formation, political decision-making, and moral judgment. Gawronski is a fellow of the Association for Psychological Science, the Society of Experimental Social Psychology, and the Midwestern Psychological Association.

Galen V. Bodenhausen, PhD, is the Lawyer Taylor Professor of Psychology and Professor of Marketing at Northwestern University. He earned his PhD in Social Psychology at the University of Illinois at Urbana–Champaign. His research addresses the cognitive aspects of social attitudes and stereotypes, particularly their roles in influencing attention, perception, memory, judgment, and behavior. A frequent focus of his recent research is on the relatively automatic and implicit aspects of prejudice and stereotyping. Bodenhausen is a fellow of the Association for Psychological Science, the American Psychological Association, the Society of Experimental Social Psychology, the Society for Personality and Social Psychology, and the Society for the Psychological Study of Social Issues.

Notes

* Correspondence: Department of Psychology, University of Texas at Austin, Austin, TX, USA. Email: gawronski@utexas.edu

[1] The procedural details of performance-based measures are extensively reviewed elsewhere, and we therefore refrain from discussing them in this article. Interested readers are referred to the chapters by Gawronski and De Houwer (2014) and Gawronski, Deutsch, and Banse (2011). A more extensive overview is provided in the chapters of the *Handbook of Implicit Social Cognition* by Gawronski and Payne (2010).

[2] Note that all of these propositions are assumed to be based on associative links. The APE model does not assume a separate storage of propositions in memory (Gawronski & Bodenhausen, 2006).

[3] Note that consciously identified CS–US pairings can trigger inferences about their relation (e.g., inferences that the CS causes the US), which can influence evaluations via propositional processes. Although such inferences can be prevented by eliminating conscious awareness of CS–US contingencies (cf. Gawronski & Walther, 2012; Sweldens, Corneille, &

Yzerbyt, 2014), the conditions under which evaluative conditioning effects are driven by associative or propositional processes (or both) are still the subject of scientific debate (for related evidence, see Gawronski, Balas, & Creighton, 2014; Hütter, Sweldens, Stahl, Unkelbach, & Klauer, 2012).

[4] In addition to the four basic cases depicted in Figure 2, there may be cases in which evaluative responses are influenced by multiple distinct factors. Such cases can be described as combinations of the four basic patterns in Figure 2. These combined patterns have in common that (a) one factor directly influences associative processes, (b) a second factor directly influences propositional processes, and (c) these direct influences may or may not have a corresponding indirect effect on the respective other process. More detailed discussions of such combined cases are provided by Gawronski and Bodenhausen (2006, 2011).

References

Barden, J., Maddux, W. W., Petty, R. E., & Brewer, M. B. (2004). Contextual moderation of racial bias: The impact of social roles on controlled and automatically activated attitudes. *Journal of Personality and Social Psychology*, **87**, 5–22.

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

Brochu, P. M., Gawronski, B., & Esses, V. M. (2011). The integrative prejudice framework and different forms of weight prejudice: An analysis and expansion. *Group Processes and Intergroup Relations*, **14**, 429–444.

Deutsch, R., Gawronski, B., & Strack, F. (2006). At the boundaries of automaticity: Negation as reflective operation. *Journal of Personality and Social Psychology*, **91**, 385–405.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, **56**, 5–18.

Duckworth, K. L., Bargh, J. A., Garcia, M., & Chaiken, S. (2002). The automatic evaluation of novel stimuli. *Psychological Science*, **13**, 513–519.

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin*, **23**, 316–326.

Eastwick, P. W., Eagly, A. H., Finkel, E. J., & Johnson, S. H. (2011). Implicit and explicit preferences for physical attractiveness in a romantic partner: A double dissociation in predictive validity. *Journal of Personality and Social Psychology*, **101**, 993–1011.

Fazio, R. H. (2007). Attitudes as object–evaluation associations of varying strength. *Social Cognition*, **25**, 603–637.

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

Friese, M., Hofmann, W., & Schmitt, M. (2008). When and why do implicit measures predict behaviour? Empirical evidence for the moderating role of opportunity, motivation, and process reliance. *European Review of Social Psychology*, **19**, 285–338.

Galdi, S., Arcuri, L., & Gawronski, B. (2008). Automatic mental associations predict future choices of undecided decision makers. *Science*, **321**, 1100–1102.

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**, 559–569.

Gawronski, B. (2012). Back to the future of dissonance theory: Cognitive consistency as a core motive. *Social Cognition*, **30**, 652–668.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, **132**, 692–731.

Gawronski, B., & Bodenhausen, G. V. (2007). Unraveling the processes underlying evaluation: Attitudes from the perspective of the APE Model. *Social Cognition*, **25**, 687–717.

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

Gawronski, B., & Bodenhausen, G. V. (2011). The associative–propositional evaluation model: Theory, evidence, and open questions. *Advances in Experimental Social Psychology*, **44**, 59–127.

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: Guilford Press.

Gawronski, B., & Cesario, J. (2013). Of mice and men: What animal research can tell us about context effects on automatic responses in humans. *Personality and Social Psychology Review*, **17**, 187–215.

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 edn, pp. 283–310). New York: Cambridge University Press.

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**, 1355–1361.

Gawronski, B., & Payne, B. K. (Eds.). (2010). *Handbook of implicit social cognition: Measurement, theory, and applications*. New York: Guilford Press.

Gawronski, B., & Quinn, K. A. (2013). Guilty by mere similarity: Assimilative effects of facial resemblance on automatic evaluation. *Journal of Experimental Social Psychology*, **49**, 120–125.

Gawronski, B., & Sritharan, R. (2010). Formation, change, and contextualization of mental associations: Determinants and principles of variations in implicit measures. In B. Gawronski & B. K. Payne (Eds.), *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications* (pp. 216–240). New York: 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**, 535–542.

Gawronski, B., & Walther, E. (2008). The TAR effect: When the ones who dislike become the ones who are disliked. *Personality and Social Psychology Bulletin*, **34**, 1276–1289.

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**, 617–623.

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

Gawronski, B., Brochu, P. M., Sritharan, R., & Strack, F. (2012). Cognitive consistency in prejudice-related belief systems: Integrating old-fashioned, modern, aversive and implicit forms of prejudice. In B. Gawronski & F. Strack (Eds.), *Cognitive Consistency: A Fundamental Principle in Social Cognition* (pp. 369–389). New York: Guilford Press.

Gawronski, B., Deutsch, R., & Banse, R. (2011). Response interference tasks as indirect measures of automatic associations. In K. C. Klauer, A. Voss & C. Stahl (Eds.), *Cognitive Methods in Social Psychology* (pp. 78–123). New York: Guilford 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**, 370–377.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, **132**, 692–731.

Gawronski, B., Geschke, D., & Banse, R. (2003). Implicit bias in impression formation: Associations influence the construal of individuating information. *European Journal of Social Psychology*, **33**, 573–589.

Gawronski, B., Hofmann, W., & Wilbur, C. J. (2006). Are "implicit" attitudes unconscious? *Consciousness and Cognition*, **15**, 485–499.

Gawronski, B., LeBel, E. P., & Peters, K. R. (2007). What do implicit measures tell us? Scrutinizing the validity of three common assumptions. *Perspectives on Psychological Science*, **2**, 181–193.

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**, 648–665.

Gawronski, B., Rydell, R. J., Vervliet, B., & De Houwer, J. (2010). Generalization versus contextualization in automatic evaluation. *Journal of Experimental Psychology: General*, **139**, 683–701.

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: Guilford Press.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, **102**, 4–27.

Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: Asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology*, **90**, 1–20.

Grumm, M., Nestler, S., & von Collani, G. (2009). Changing explicit and implicit attitudes: The case of self-esteem. *Journal of Experimental Social Psychology*, **45**, 327–335.

Hahn, A., & Gawronski, B. (2014). Do implicit evaluations reflect unconscious attitudes? *Behavioral and Brain Sciences*, **37**, 28–29.

Hahn, A., Judd, C. M., Hirsh, H. K., & Blair, I. V. (in press). Awareness of implicit attitudes. *Journal of Experimental Psychology: General*.

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

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, **31**, 1369–1385.

Hollands, G. J., Prestwich, A., & Marteau, T. M. (2011). Using aversive images to enhance healthy food choices and implicit attitudes: A experimental test of evaluative conditioning. *Health Psychology*, **30**, 195–203.

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

Jones, E. E., & Gerard, H. B. (1967). *Foundations of Social Psychology*. New York: Wiley.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, **108**, 480–498.

Moran, T., & Bar-Anan, Y. (2013). The effect of object–valence relations on automatic evaluation. *Cognition and Emotion*, **27**, 743–752.

Moss, A. C., & Albery, I. P. (2009). A dual-process model of the alcohol-behavior link for social drinking. *Psychological Bulletin*, **135**, 516–530.

Ouimet, A. J., Gawronski, B., & Dozois, D. J. A. (2009). Cognitive vulnerability to anxiety: A review and an integrative model. *Clinical Psychology Review*, **29**, 459–470.

Perugini, M., Richetin, J., & Zogmaister, C. (2010). Prediction of behavior. In B. Gawronski & B. K. Payne (Eds.), *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications* (pp. 255–277). New York: Guilford Press.

Peters, K. R., & Gawronski, B. (2011). Mutual influences between the implicit and explicit self-concepts: The role of memory activation and motivated reasoning. *Journal of Experimental Social Psychology*, **47**, 436–442.

Redker, C. M., & Gibson, B. (2009). Music as an unconditioned stimulus: Positive and negative effects of country music on implicit attitudes, explicit attitudes, and brand choice. *Journal of Applied Social Psychology*, **39**, 2689–2705.

Richter, T., Schroeder, S., & Wöhrmann, B. (2009). You don't have to believe everything you read: Background knowledge permits fast and efficient validation of information. *Journal of Personality and Social Psychology*, **96**, 538–558.

Rydell, R. J., & Gawronski, B. (2009). I like you, I like you not: Understanding the formation of context-dependent automatic attitudes. *Cognition and Emotion*, **23**, 1118–1152.

Smith, E. R. (1996). What do connectionism and social psychology offer each other? *Journal of Personality and Social Psychology*, **70**, 893–912.

Strick, M., Holland, R. W., Van Baaren, R. B., Van Knippenberg, A., & Dijksterhuis, A. (2013). Humor in advertising: An associative processing model. *European Review of Social Psychology*, **24**, 32–69.

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

Van den Bos, K., & Maas, M. (2009). On the psychology of the belief in a just world: Exploring experiential and rationalistic paths to victim blaming. *Personality and Social Psychology Bulletin*, **35**, 1567–1578.

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

Whitfield, M., & Jordan, C. H. (2009). Mutual influences of explicit and implicit attitudes. *Journal of Experimental Social Psychology*, **45**, 748–759.

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology*, **81**, 815–827.