

RUNNING HEAD: PROPOSITIONAL PERSPECTIVE

**Attitudes Beyond Associations:
On the Role of Propositional Representations in Stimulus Evaluation**

Jan De Houwer, Pieter Van Dessel, & Tal Moran

Ghent University, Belgium

In press. *Advances in Experimental Social Psychology*.

mailing address:

Jan De Houwer
Ghent University
Henri Dunantlaan 2
B-9000 Ghent
Belgium
email: Jan.DeHouwer@UGent.be

phone: 0032 9 264 64 45
fax: 0032 9 264 64 89

Abstract

The idea that attitudes are associations in memory has been highly influential. We highlight an alternative perspective according to which stimulus evaluation is mediated by propositional representations. Unlike simple associations, which are merely links between nodes via which activation can spread, propositional representations are units of information that specify how events are related and that have a truth value. We review research on attitude acquisition and attitude activation that was inspired by this propositional perspective. This research reveals surprising parallels between phenomena that are often assumed to be fundamentally different (e.g., evaluative conditioning vs. persuasion; implicit vs. explicit evaluation) but also interesting differences that provide the impetus for future research.

Keywords: Attitudes, evaluations, associations, propositions, evaluative conditioning, implicit evaluation

It is widely accepted that the concept *attitude* lies at the core of social psychology (e.g., Allport, 1935; Fazio, 2007). From a cognitive perspective, attitudes are mental representations that determine the evaluation of stimuli, that is, whether we respond to stimuli in favorable or unfavorable ways (De Houwer, Gawronski, & Barnes-Holmes, 2013). Based on the idea that behavior in general is instigated by mental mechanisms (Bechtel, 2008), it is indeed logical to assume that mental representations mediate stimulus evaluation because it is the only mechanistic way via which past events can influence current stimulus evaluations (Fazio, 2007). However, merely postulating that there are mental representations that mediate stimulus evaluation contributes little to our understanding of stimulus evaluation. The added value of cognitive theories must come from the answers they provide to four core questions (De Houwer et al., 2013): (a) What is the nature of attitudes? (b) How are they acquired? (c) How are they activated from memory? (d) How do they influence behavior? Within this set of questions, the issue of the nature of attitudes is a central one because it constrains ideas about the other three issues.

Until recently, cognitively inspired attitude research was dominated by the idea that attitudes are associations in memory. This idea has proven to be highly generative, especially with regard to research on attitude acquisition and attitude activation. After sketching the associative perspective and some of its impact on attitude research, we describe an alternative perspective according to which propositional representations mediate stimulus evaluation. The main aim of this chapter is to review research on attitude acquisition and activation that was inspired by this alternative, propositional perspective on attitudes. Although it remains difficult, if not impossible, to prove or refute proposals about the nature of the representations that mediate stimulus evaluation, we hope to show that also the propositional perspective has been highly generative and thus provides a useful tool for attitude researchers.

Background

Attitudes as Associations

Fazio (2007; Fazio, Chen, McDonel, & Sherman, 1982; Fazio, Sanbonmatsu, Powell, & Kardes, 1986) famously proposed that attitudes can be thought of as object-valence associations of varying strength. For instance, a positive attitude toward brownies would correspond to an association between the representation of brownies and the representation of positive valence. When seeing a brownie, this would activate the representation of brownies. This activation can then spread via the association to the representation of positive valence and hence lead to positive evaluative responses (e.g., experiencing positive thoughts or feelings, smiling, approaching the brownie, etc.). In this way, the idea that attitudes are associations provides a bridge between attitude research and the long tradition of associationistic thinking in philosophy and psychology (see Anderson & Bower, 1973, for a historical overview).

This apparently simple idea has been beneficial for attitude research, especially with regard to research on attitude acquisition and attitude activation (see Gawronski & Sritharan, 2010, for an overview of some of the other implications of an associative perspective on attitudes). First, because it is assumed that activation can spread automatically across (strong) associations in memory, theorists argued that attitudes can be activated automatically, that is, even under sub-optimal conditions such as when there is a little time or when people are engaged in order demanding tasks (see Bargh, 1992; Moors, 2016; Moors & De Houwer, 2006, for more details on the concept of automaticity). Importantly, the idea of automatic attitude activation gave rise to research on implicit evaluation, which we define as automatic responding in favorable or unfavorable ways (De Houwer, 2009a; De Houwer et al., 2003; see Ferguson & Zayas, 2009, for a review). For instance, in their seminal study, Fazio et al. (1986) used an evaluative variant of the semantic priming task (e.g., Meyer & Schvaneveldt,

1971) that consisted of a series of trials on which a positive or negative prime word was flashed very briefly before the onset of a positive or negative target word. Participants were asked to indicate the valence of the targets. Results showed that responses were faster when prime valence and target valence matched (e.g., HAPPY – FLOWERS) than when they mismatched (e.g., HAPPY – CANCER). Such an evaluative priming effect indicates that participants evaluate the primes automatically in the sense of quickly (i.e., despite the brief time between prime onset and responding) and efficiently (i.e., despite having the task to respond to the targets as quickly as possible; Bargh, 1992; Moors & De Houwer, 2006). This seminal finding contributed to the development of implicit measures of attitudes that are now being used widely inside and outside of psychology (see Gawronski & Brannon, 2019, for a review).

Second, within philosophy and psychology, it has long been assumed that associations are formed as the result of the spatiotemporal pairing of events (Anderson & Bower, 1973). In its most simple form, this idea entails that an association between representations is gradually strengthened each time that both representations are active simultaneously (i.e., “what fires together, wires together”; Hebb, 1949). Hence, if associations underlie stimulus evaluation, then the spatiotemporal pairing of neutral and valenced events should be an important determinant of whether we respond to stimuli in a favorable or unfavorable manner. This insight led Olson and Fazio (2001) to direct their attention to the phenomenon of evaluative conditioning (EC), that is, changes in liking that are due to the pairing of stimuli (Levey & Martin, 1975; see De Houwer, Thomas, & Baeyens, 2001, for an early review). Many social psychologists followed their lead, resulting in an intensification of research on EC (see Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010, for a more recent review).

Propositions versus Associations

Although the idea that attitudes are associations in memory has been highly successful in generating new research, during the past decade, we have been exploring the merits of an alternative perspective according to which propositional representations provide the basis for stimulus evaluations (see De Houwer, 2014a, 2018, for reviews). This line of research originated from work on classical conditioning and causal learning that the first author of this chapter engaged in (see De Houwer, 2009b; Mitchell, De Houwer, & Lovibond, 2009). Within that literature, a heated debate was taking place between proponents of associative models of learning and proponents of propositional models of learning (see Shanks, 2010, for a review). At the same time, a number of findings in EC research emerged that seemed incompatible with a simple associative perspective, thus providing the impetus for the development of alternative perspectives (De Houwer, Baeyens, & Field, 2005). Finally, attitude researchers started to relate their work to dual-system models that contrast associative with propositional processes (e.g., Strack & Deutsch, 2004), which provided the basis for contrasting the role of associations and propositions in attitude research.

In order to appreciate the unique contribution of propositional theories to attitude research, it is important to have a good understanding of what propositional representations are. They differ from (simple) associations in two important respects (see Hummel, 2010; Lagnado, Waldmann, Hagmayer, & Sloman, 2007, for more details). First, unlike (simple) associations, propositional representations can specify the way in which events are related, that is, they encode relational information. For instance, a propositional representation of the statement “John loves Mary” encodes not only a link between John and Mary but also a specific type of relation (e.g., loves) as well as the role that each element plays within that relation (e.g., the fact that John is the lover and Mary is the beloved; see Hummel, 2010). Second, because the information encoded by propositional representations always entails

assumptions about events in the world (e.g., the belief that John loves Mary), propositional representations can be said to have a truth value in the philosophical sense. In other words, the content of each propositional representation has the potential to be true or untrue, even if its truth cannot be determined in actuality (e.g., as is the case with the statement “angels have wings”; De Houwer, 2018). These two defining characteristics of propositional representations hold for both representations that encode relationally simple or vague information (e.g., “John is somehow related to Mary”) and representations that encode relationally complex or precise information (e.g., “John is the brother-in-law of Mary”). A simple association between concepts, on the other hand, cannot encode specific relations or roles within relations.

Some have argued that more complex associative structures can encode relational information. For instance, one might add a relational label to an association (e.g., adding the label “loves” to the association between John and Mary; see Anderson & Bower, 1973) or design multi-layered associative structures (i.e., connectionist networks) with the aim of encoding relational information (e.g., Gawronski & Bodenhausen, 2006, 2018; Kollias & McClelland, 2013). On the one hand, it remains to be seen whether more complex associative structures are up to the job. For instance, simply adding relational labels to associations does not explain where the labels come from and introduces a non-associative component to the structures. Moreover, the labels do not capture information about the role that elements have within a relation. Multi-layered associative structures do not require relational labels, but the currently available models cannot account for the flexibility with which people construct relations and relational roles (see Hummel, 2010, for an insightful discussion of why associative structures are inherently limited in their capacity to capture relational information). On the other hand, whether associative *structures* can capture complex relational information is not crucial for the aims of the present chapter because our focus is

on the distinction between associative and propositional *representations*. In its essence, the latter distinction is not about the nature of the mental structures that encode information (e.g., links between nodes) but about the nature of the information itself, more particularly whether the represented information specifies relations and roles within relations (e.g., the fact that John loves Mary). Hence, any representational structure that encodes truth-evaluable relational information qualifies as a propositional representation, even when that structure consists of associations via which activation can spread (see also De Houwer, 2014a, 2018). The real question that is raised by the distinction between associative and propositional representations is whether behavior is driven by representations that do not encode information about relations and relational roles (i.e., associative representations) versus representations that do encode information about relations and relational roles (i.e., propositional representations). In the remainder of the chapter, we will therefore use the terms *associations* and *propositions* only in the sense of *associative representations* and *propositional representations*, respectively, that is, as referring to informational content.¹

¹ Gawronski and Bodenhausen (2018, p. 5) recently argued that the distinction between associations and propositions becomes void when viewed from the perspective of a multi-layered associative (i.e., connectionist) memory system because such a system can encode both associations and propositions. We disagree with this position for two reasons. First, as we noted in the main text, it is not at all clear whether a multi-layered associative memory system can encode (complex) propositions. What is clear is that none of the current associative models can simulate the complexities of relational behavior and that there are good logical arguments for why any type of associative model would have difficulties in achieving this (Hummel, 2010). In response to these arguments, Gawronski and Bodenhausen (2018, Footnote 2) wrote that “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”. This argument, however, simply sidesteps the issue by implying that all scientific theories of relational processing are associative in nature. It might well be that the complexities of relational processing are beyond the reach of theories that only allow for associative structures for encoding information but that they are in reach of other scientific theories that allow for other representational structures. Whereas it is impossible to foresee future theoretical developments, we believe that it is at least premature to dismiss the distinction between associations and propositions based on the idea that both *might* be represented within a multi-layered associative structure. Second, and more importantly, Gawronski and Bodenhausen confound the structure that encodes information with the information that is encoded. Even if an associative structure could encode both associative and propositional representations, this would not change the fact that associative representations do not encode the same type of information as propositional representations. Hence, questions can continue to be raised about whether stimulus evaluation is mediated by associative representations or propositional representations (i.e., units without information about relations and relational roles or units with information about relations and relational roles).

As we noted at the start of this chapter, raising questions about the nature of the attitudinal representations that mediate stimulus evaluation is useful because it has important implications for other core issues in attitude research, most prominently attitude acquisition and attitude activation. First, the nature of the representations has implications for the range of events that can shape those representations (and that thus determine the resulting stimulus evaluations), as well as the conditions under which the representations are formed (and thus stimulus evaluations start to emerge). Because associative representations do not specify the type of relation or relational roles, they can be created merely on the basis of spatiotemporal events such as the co-occurrence of stimuli. Propositional representations, on the other hand, require more input than mere co-occurrence simply because the same pattern of spatiotemporal co-occurrence (e.g., John and Mary are often seen together) can result from multiple types of relations (e.g., John loves Mary, John is an adversary of Mary, John is a colleague of Mary, etc.; Lagnado et al., 2007). Hence, the formation of propositions typically requires information that goes beyond spatiotemporal co-occurrence. This extra information can come from a variety of sources such as memory, context, observation of others, and instructions by others. As such, stimulus evaluations that originate from propositional representations are likely to be influenced by a wide variety of sources from which relational information can be derived.

The wide range of events that shape propositional representations also has implications for the conditions under which these representations are formed and thus stimulus evaluations are acquired. Whereas the impact of spatiotemporal events on associations is sometimes assumed to be highly automatic (e.g., Gawronski & Bodenhausen, 2006), the integration of information that is required for the formation of propositions often occurs in a non-automatic manner (i.e., dependent on time, resources, awareness, or goals; Bargh, 1992; Moors & De Houwer, 2006), much like the processes underlying problem

solving (De Houwer, 2018). Hence, whereas stimulus evaluations that depend on the formation of associations should typically be acquired automatically as the result of spatiotemporal events, stimulus evaluations that are driven by propositions are often acquired only under more optimal conditions. In sum, the nature of the mediating representations could have important implications for both the sources that shape stimulus evaluations as well as the conditions under which stimulus evaluations are acquired.

A second reason why studying the representations that mediate stimulus evaluation is important for attitude research relates to the fact that propositional representations, but not associations, have a truth value. This implies that only propositional representations allow for inferential reasoning, that is, the operation of processes via which the truth value of conclusions is derived from the truth value of premises.² Therefore, any behavioral phenomenon that is shaped by inferential reasoning must be mediated by propositional representations. Although drawing inferences requires propositional representations, behavior that is mediated by propositional representations is not necessarily based on (rational) inferences. First, once they have been formed, propositions can be automatically retrieved from memory in a non-inferential manner, for instance, as the result of similarity-based retrieval processes (De Houwer, 2009b, 2018). Second, irrational behavior can result from irrational propositional beliefs or from irrational inferential reasoning (i.e., reasoning that does not follow logical rules) made on the basis of rational beliefs. Another implication of the truth value of propositions is that variables that influence the truth evaluation of a proposition could in principle also influence behavior. Manipulations that target the truth value of a

² Although we do not commit to any assumptions about what these inferential processes might look like, it is unlikely that they can be reduced to spreading of activation. First, activation of a representation does not necessarily imply the truth of the information encoded in that representation (see Jozefowicz, 2018, for an insightful discussion). Second, inferences typically involve multiple premises whose truth value jointly determines the truth value of the conclusion. It is difficult to see how such a complex combination of information could be realized by simple spreading of activation. If, however, we are wrong in rejecting spreading of activation as an inferential process, it would still need to operate on representations that have the potential to be true or false, that is, on propositional representations.

proposition might therefore often influence stimulus evaluations, although this will not always be the case (e.g., when stimulus evaluations reflect invalid propositions that are automatically retrieved from memory).

A Propositional Perspective on EC and Implicit Evaluation

The proposal that at least some instances of stimulus evaluation are mediated by propositions is somewhat trivial. For instance, few attitude researchers would disagree with the claim that persuasion effects require propositional representations, if only because of the relational nature of persuasive messages and the inferences that recipients of persuasive message must make in order to display the predicted changes in their evaluation of stimuli. Nevertheless, the idea that attitudes are associations in memory remains popular in attitude research, especially within research on EC (e.g., Hütter & Sweldens, 2018) and implicit evaluation (e.g., Gawronski & Hahn, 2019),³ probably in part because the initial studies on these topics were inspired by an associative perspective.

The predictive success of a theory does not, however, imply that it is correct. Within the historical context that we sketched at the start of the previous section, we have tried to move beyond associations by proposing that all attitudinal phenomena are mediated by propositional representations, including EC and implicit evaluation. First, De Houwer (2009b, 2018) put forward the idea that the pairing of stimuli can result in changes in

³ Gawronski and Bodenhausen's (2006, 2018) influential Associative-Propositional Evaluation (APE) model focusses on the distinction between associative and propositional processes rather than the distinction between associative and propositional representations. As mentioned in Footnote 1, they do not acknowledge the need for a distinction between associative and propositional representations. However, as we argued in Footnote 1, we believe that their dismissal of this distinction is flawed because of a failure to acknowledge the fact that associative and propositional representations encode different types of information. Propositional processes such as truth validation and inferential reasoning require propositional representations, that is, informational units that can be truth evaluated. Hence, we believe that the APE model also implies a distinction between associative and propositional representations as informational units with a different content. Note, however, that associative (in the sense of similarity-based) processes can operate on propositional representations (i.e., the similarity-based retrieval of propositional representations). Hence, the nature of processes is not fully constrained by the nature of the representations upon which those processes operate. Finally, we agree with Gawronski and Bodenhausen that one should also not conflate different types of processes with different types of operating conditions. Most importantly, propositional processes (e.g., inferential reasoning) can also operate under conditions of automaticity (e.g., Van Dessel, Hughes, & De Houwer, 2019).

(evaluative) behavior only after a proposition has been formed about the relation between the two stimuli. In a typical EC task, a neutral (conditioned) stimulus (CS) is repeatedly paired with a positive or a negative (unconditioned) stimulus (US), and as a result, the former typically acquires a similar valence as the latter (i.e., a standard EC effect). From a propositional perspective, the CS-US pairing can result in the formation of a propositional representation of the spatiotemporal relation between the CS and US (e.g., a representation specifying that the CS co-occurs with a positive US) which can then lead to changes in liking (e.g., based on the inference that stimuli that co-occur often have the same valence; see De Houwer, 2018; Van Dessel, Hughes, & De Houwer, 2019, for more details). Second, De Houwer (2014a) and others (e.g., Mandelbaum, 2016) argued that implicit evaluation depends on the automatic activation of propositions about the evaluated stimulus. For instance, the fact that a brief presentation of the prime word FLOWERS reduces the time to indicate that the target word HAPPY has a positive valence (e.g., Fazio et al., 1986) might result from the activation of the proposition “flowers are good”, which produces a tendency to emit positive responses and therefore facilitates positive responses to the target word HAPPY (see De Houwer, 2014a, for more details).

Note that both proposals make sense only if one clearly separates the to-be-explained phenomenon (i.e., EC as the impact of pairings on liking; implicit evaluation as the automatic impact of stimuli on evaluative responses) from the mental processes and representations that mediate these phenomena (see De Houwer, 2014a; De Houwer et al., 2013). In other words, considering propositional models of EC and implicit evaluation requires definitions of these phenomena that do not refer to associations (e.g., EC as the formation of associations or implicit evaluation as the activation of associations; De Houwer, 2007, 2011; De Houwer et al., 2013). In fact, removing references to associations from the definitions of EC and implicit evaluations is as such an important step forward because it gets rid of (potentially flawed) a

priori assumptions and thereby broadens the scope for theoretical debate (De Houwer, 2007; De Houwer et al., 2013).

At present, propositional models of EC and implicit evaluation incorporate little more than the assumption that these phenomena are mediated by the formation and activation of propositional representations. This is why, in the remainder of the chapter, we will talk about a propositional perspective on, rather than propositional models of, EC and implicit evaluation. We will contrast this propositional perspective with the associative perspective, that is, the idea that EC and implicit evaluation are mediated by the formation and activation of associative representations. Whereas these two perspectives each put forward only one type of attitudinal representation, one could also put forward a dual-representation perspective according to which both propositional and associative representations contribute to EC and implicit evaluation (e.g., Rydell & McConnell, 2006). All three of these perspectives encompass a wide range of potential theories that make specific assumptions about the exact nature, formation, and activation of attitudinal representations. It should be clear from the start that it is impossible to empirically differentiate between these broad perspectives (Miller & Escobar, 2001). For each empirical finding, it is probably possible to find some (variant of a) theory within each class of theories that can account for the finding. However, even broad theoretical perspectives can be generative, that is, have a heuristic and predictive value. In fact, this was already proven by the huge impact of the idea that attitudes are associations (e.g., Fazio, 2007). With a few exceptions (e.g., Hastie, 1988), in attitude research, this idea was never elaborated beyond some general assumptions about the formation of associations (e.g., “what fires together, wires together”) and the activation of associations (e.g., automatic spreading of activation). Nevertheless, these basic assumptions sufficed to generate important new lines of attitude research that provided crucial knowledge

about the moderators of stimulus evaluation and thereby constrained any current and future theory of the mechanisms that mediate stimulus evaluation.

In a similar way, the idea that all attitudinal phenomena (including EC and implicit evaluation) are mediated by propositions can be highly generative if it is combined with what we know about propositions. As noted above, propositional representations (a) can specify relational information, (b) can originate from a variety of sources, (c) have a truth value, and (d) are often formed in non-automatic ways. Hence, merely considering the possibility that attitudinal phenomena like EC and implicit evaluation are grounded in propositional representations encourages researchers to examine whether and how these phenomena (a) depend on relational information, (b) are shaped by different sources of information, (c) depend on truth validation, and (d) might sometimes have feature of non-automaticity.

Aims and Scope

In the remainder of this chapter, we review studies that were inspired by the idea that propositional representations mediate EC and implicit evaluation. More specifically, we review four lines of research on EC and/or implicit evaluation that were for a large part instigated on the basis of what we know about propositions: (a) studies on the effects of relational information in EC, (b) studies on the effects of relational information in implicit evaluation, (c) research testing the impact of instructions about CS-US contingencies on changes in liking, and (d) studies on the effects of truth validation. We also briefly refer to studies on the automaticity features of EC. The latter line of research is not given as much attention as the other four because it originated from an associative perspective on EC, but it still deserves to be mentioned because it was strongly affected by the proposal of propositional theories of EC.

In reviewing the literature, we will ignore whether the study we refer to was designed to confirm or disconfirm propositional theories of EC or implicit evaluation. Our focus will

also not be on whether the results uniquely confirm or disconfirm propositional theories. As noted above, even for findings that confirm the predictions of propositional models, it will often be possible to come up with alternative explanations in terms of associations in memory. We also realize that the structure which we impose on the available literature is, like any other possible structure, somewhat limiting and artificial. However, we hope that this chapter helps to illustrate the generative power of considering the idea that propositional representations mediate attitudinal phenomena like EC and implicit evaluation.

Review of the Evidence

On the Role of Relational Information in EC

Basic idea. Given the dominance of associationistic thinking, EC researchers for a long time ignored the potential impact of relational information. From a propositional perspective, however, spatio-temporal stimulus pairings are often just a first step in forming propositions about the evaluative nature of stimuli. As noted above, CS-US pairings can result in the belief that the CS and US co-occur. Whether this leads to the inference that the CS is good or bad can depend on other information that constrains the relational implications of the CS-US pairings (Van Dessel et al., 2019). For instance, if participants are led to believe that paired stimuli are opposites in some respect (e.g., that a novel CS word is the antonym of a known US word), then the spatio-temporal co-occurrence of a CS and US indicates that the CS has a valence opposite to the US. In these cases, reversed EC effects should emerge (i.e., a CS is liked less after it is paired with a positive US compared to a negative US). If, however, EC depends merely on a contiguity-driven formation of associations in memory, relational information should not moderate the impact of CS-US pairings on CS liking.

EC of Explicit Evaluations. A first set of studies examined whether relational information moderates conditioned changes in explicit evaluation, that is, evaluation under relatively optimal conditions (e.g., when there is ample time and resources and a conscious

goal to evaluate stimuli, as is the case with evaluative ratings on Likert scale). In line with the propositional perspective, many studies showed reversed EC effects when CSs and USs were described as being opposite in some respects (e.g., Förderer & Unkelbach, 2012; Gawronski, Walther, & Blank, 2005; Hu, Gawronski, & Balas, 2017a; Moran, Bar-Anan, & Nosek, 2016; Unkelbach & Fiedler, 2016). For example, Förderer and Unkelbach (2012) paired pictures of unknown individuals (CSs) with generally liked or disliked animals (e.g., a kitten, a pit-bull). On each trial, they also showed a relational qualifier which indicated whether the CS individual *loves* or *loathes* the US animal. For CS-US pairs that were said to be opposite (e.g., CS loathes a kitten), the explicit evaluations of the CS individuals showed a reversed EC effect. For instance, participants liked men that loathed snakes more than men that loathed puppies. In another study (Moran et al., 2016, Experiment 2), CSs (cartoon characters) co-occurred with positive or negative USs (e.g., pictures of puppies or cockroaches) while information was provided about whether the CS *gives* or *takes away* the US. Again, the explicit evaluation of the characters showed a reversed EC effect when the relational qualifier implied opposition. For instance, participants liked characters that took away cockroaches more than characters that took away puppies (see Figure 1). These studies show that conditioned changes in explicit evaluations are not just a function of CS-US pairings but also of information about the relation between the CSs and USs.

It is interesting to note that in all these studies, the reversed EC effects were in absolute terms weaker than standard EC effects (see Bar-Anan & Moran, 2018; Moran et al., 2016, for a review). Although this might indicate a joint influence of associations (as determined by CS-US pairings) and propositions (as determined by the combination of CS-US pairings and relational qualifiers), it is also possible to explain this without reference to associations. For instance, CS liking could be influenced both by propositions about CS-US co-occurrences (as determined by CS-US pairings; e.g., “CS co-occurs with cockroaches”)

and propositions about the evaluative properties of the CS (as determined by CS-US pairings and relational qualifiers; e.g., “CS is good because it takes away cockroaches”).

Alternatively, it is possible that some participants fail to take into account the relational qualifiers (see Moran et al., 2016, for a detailed discussion). Regardless of this issue, it is now clear that relational information can reverse EC effects on explicit evaluations.

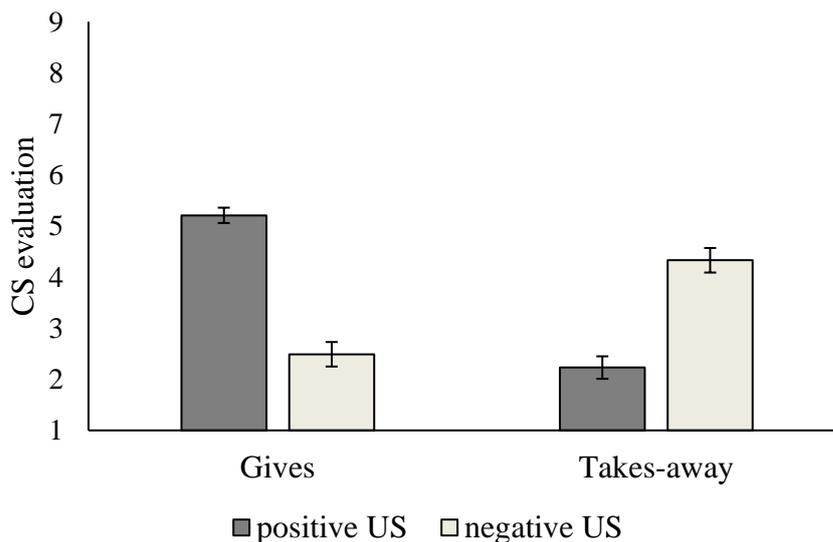


Figure 1. CS evaluation as a function of US valence (positive, negative) and relational information (gives, takes-away). Error-bars reflect +1 and -1 standard error of the mean. Adapted from Moran, T., Bar-Anan, Y., & Nosek, B. A. (2016). *The assimilative effect of co-occurrence on evaluation above and beyond the effect of relational qualifiers*. *Social Cognition*, 34, 343-356, Experiment 2.

Within this line of research, there is one other finding that seems to challenge a propositional perspective. Langer, Walther, Gawronski, and Blank (2009) first presented unknown individuals with positive or negative behavioral descriptions. Pictures of these individuals were used as USs in a second phase during which participants saw pairs of pictures showing a novel individual (CS) and a positive or negative individual (US). On each trial, participants were told whether the US liked or disliked the CS. Afterwards, the liking of half of the USs was changed by presenting them with new behavioral descriptions of the opposite valence. Finally, participants completed a measure of explicit CS evaluations. For

the CSs that were paired with a US whose valence did not change, results were influenced by the relational information. More specifically, a CS that was liked by a positive US was liked more than a CS that was liked by a negative US (i.e., a standard EC effect) whereas a CS that was disliked by a positive US was liked less than a CS that was disliked by a negative US (i.e., a reversed EC effect; see also Gawronski et al., 2005). However, for the CSs that were paired with USs whose valence did change, CS liking was in line with the new valence of the US regardless of whether the US faces liked or disliked the CS faces. Langer et al. interpreted the latter findings as evidence for a direct mental association between the CS and the US, which influenced their CS evaluations regardless of CS-US relations. However, it is possible that participants simply failed to take into account all the different pieces of information (CS-US co-occurrence, CS-US relation, changes in US valence), especially when USs were said to dislike CSs. Future research could test this suggestion by using a simpler learning task (e.g., by using fewer CS-US pairs and more opportunity to learn each piece of information). Moreover, because this study had a small number of participants and has not yet been replicated, some caution is required when interpreting the results.

EC of Implicit Evaluations. The evidence for an effect of relational information on EC of *implicit evaluations* is more ambiguous. Whereas some studies failed to find an effect of relational information (Hu, Gawronski, & Balas, 2017a, Experiments 1-2; Moran & Bar-Anan, 2013), others revealed a strong effect with reversed EC when relational qualifiers implied opposition (Gawronski, Walther, & Blank, 2005; Hu et al., 2017a, Experiment 3), and still other studies found a mere attenuation of standard EC as a result of oppositional relational qualifiers (Hughes, Ye, & De Houwer, 2018; Moran & Bar-Anan, 2019; Zanon, De Houwer, & Gast, 2012). At a first glance, the fact that there is a more consistent effect of relational information on explicit than on implicit evaluation might be seen as a support for the idea that explicit and implicit evaluations depend on the formation of different kinds of

representations (e.g., Rydell & McConnell, 2006). However, it is important to note that differential effects of relational information on implicit and explicit evaluations are theoretically ambiguous because they could be driven not only by learning-related processes (e.g., the formation of associations versus propositions) but also by retrieval-related processes (e.g., the retrieval of different propositions) or a combination of both (De Houwer, 2018).

Interestingly, an inspection of the available evidence reveals a number of potential moderators of the impact of relational information on EC of implicit evaluations. First, the time at which relational information is presented might be important. More specifically, relational information seems to have a bigger impact on implicit evaluation when it is presented simultaneously with the CS-US pairs (e.g., Gawronski et al., 2005; Hu et al., 2017a, Experiment 3; Hughes, Barnes-Holmes, et al., 2018) than when it is presented before or after the CS-US pairings (e.g., Hu et al., 2017a: Experiments 1-2; Moran & Bar-Anan, 2013; Zanon, De Houwer, Gast, & Smith, 2014). This makes sense from a propositional perspective because relational information is bound to have the biggest effect at the time that participants are forming propositions about the CSs, that is, when the CSs are paired with USs. Note, however, that the moderating effect of timing has not yet been tested directly by manipulating timing within a single study.

Second, providing relational information via verbal instructions (i.e., adding verbal relational qualifiers such as “cause” or “prevent”; e.g., Hu et al., 2017a; Gawronski et al., 2005) seems to have more impact than presenting non-verbal contextual cues (e.g., Hughes, Ye, & De Houwer, 2018; Zanon et al., 2012). For example, Hughes, Ye, and De Houwer, (2018) presented not only CS-US trials but also context trials with two words that either were identical (e.g., up-up) or opposite (e.g., up-down). The opposition context intervention was meant to increase the probability that a CS would be contrasted with (i.e., seen as opposite to) the US it was paired with. Although this intervention attenuated EC (relative to the identical

context condition), the effect was small at best. Again, this can be understood from a propositional perspective. In order for relational information to have an impact, participants must realize that information is provided about the nature of the CS-US relation. Whereas this can be made clear via instructions, participants might not realize that context trials, or the way in which they are asked to relate the CS and US, is actually relevant for determining the nature of the CS-US relation.

Third, it seems to be important whether relational information is manipulated within or between participants. The most relevant study in this context was conducted by Hu et al. (2017a) who added the relational qualifiers *cause* and *prevent* to the pairing of novel pharmaceutical products (CSs) with positive and negative health-related conditions (USs). In their first two experiments, Hu et al. manipulated the relational information *between participants*: they informed half of the participants that the products cause the health conditions and informed the other half that the products prevent the health conditions. By contrast, in their third experiment, they manipulated relational information *within participants* (i.e., they informed all participants that some products cause, and other products prevent the health conditions). When relational information was manipulated between participants, only standard EC effects were observed, even when the CS was said to prevent the US. By contrast, when relational information was manipulated *within participants*, there was a strong effect of relational information with a reversed EC effect for CSs that prevented US outcomes (see Fiedler & Unkelbach, 2011, for a related result with explicit evaluation). Note, however, that Experiments 1-2 and Experiment 3 in Hu et al. (2017a) also differed in the timing of the relational information (i.e., relational information presented before vs. during the CS-US pairings). Future research is therefore needed to disentangle the impact of timing and design. From a propositional perspective, it would make sense that the design of the study moderates the impact of relational information. When relational information is

manipulated on a within-basis, each participant is confronted with the two types of relations, which makes this information more salient. In between-designs, on the other hand, participants might be more inclined to ignore the information because it remains constant throughout the task.

A study by Moran, Bar-Anan, and Nosek (2015, see also Experiments 4-5 in Moran et al., 2016, for similar results with explicit evaluations) speaks directly to the issue of the salience of relational information. In line with the studies of Moran and Bar-Anan (2013), a CS (cartoon creature) co-occurred with the appearance or the disappearance of a positive (a melody) or a negative (a scream) US. Participants were instructed before the pairing that they would observe four families of creatures: one that starts an unpleasant noise, one that starts pleasant music, one that ends the unpleasant noise, and one that ends the pleasant music. In one condition, Moran et al. (2015) asked participants to keep track of the nature of the US with which each CS co-occurred (i.e., melody or scream). In the other condition, participants were asked to pay attention to whether a CS performed a good action (i.e., starting the positive melody or stopping the negative scream) or a bad action (i.e., starting the negative scream or stopping the positive melody). Because only the second question requires the processing of relational information, it is safe to conclude that relational information was more salient in the second than in the first condition. This manipulation clearly moderated the impact of relational information on implicit evaluation. When participants focused on the nature of the US, their implicit evaluations were unaffected by the relational information: Participants showed a preference for the CS that co-occurred with the positive USs over the CS that co-occurred with negative USs regardless of the relational information (despite the fact that participants showed perfect memory of the relational information). By contrast, when participants had to focus on the valence of the CS action, an impact of relational information was found on implicit evaluations, with participants showing a tendency for a

reversed EC effect when CSs were said to end USs (see Figure 2).

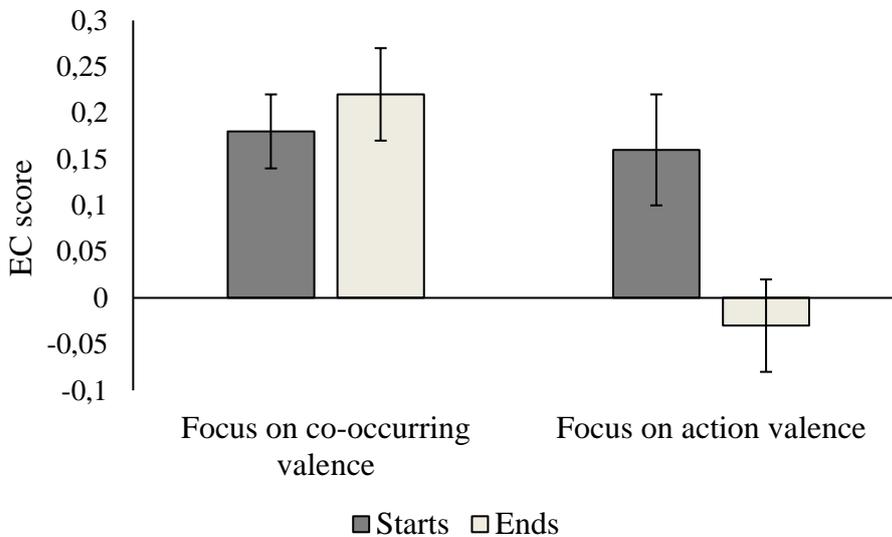


Figure 2. Implicit EC score (i.e., preference for the CS that co-occurred with positive sounds over the CS that co-occurred with negative sounds) as a function of focus condition (focus on co-occurring valence, focus on action valence) and relational information (start, end). Error bars reflect +1 and -1 standard error of the mean. *Adapted from 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, Experiment 2.*

From an associative perspective, one could say that task instructions (i.e., instructions to focus on nature of US or nature of action) change the US (e.g., a scream vs. stopping a scream) and thus the CS-US associations that are formed. A recent study by Bading, Stahl, and Rothermund (2019) argues against this interpretation by showing that the salience of relational information has an effect even when it is manipulated after the CS-US pairings. They used the same paradigm as Moran and Bar-Anan (2013) but manipulated the nature of the IAT that was used to register implicit evaluations. In Moran and Bar-Anan's study, participants completed two IATs, one that compared the CSs that started the sounds, and one that compared the CSs that ended the sounds. For the IAT that compared the CSs that ended the sounds, Moran and Bar-Anan did not find an effect of relational information: Participants showed a preference for the CS that ended the positive USs over the CS that ended the

negative USs. According to Bading et al., the fact that the IATs compared CSs that shared their CS-US relation but differed in the valence of their co-occurring US, led participants to focus on the US valence and ignore the relational information. The results observed by Moran and Bar-Anan could thus be due to the lack of salience of the relational information during test. To test this idea, Bading et al. varied the identity of the two CSs that were compared on the IAT and found that when the two CSs differ in their CS-US relation (e.g., in an IAT that compares the CS that started positive sounds with the CS that ended positive sounds), relational information did have a clear impact on the implicit evaluation (e.g., participants showed a preference for the CS that started positive sounds over the CS that ended positive sounds). It is difficult to explain these results from an associative perspective because the manipulation was implemented after the CS-US pairings and thus after the CS-US associations would have formed. These results can be explained from a propositional perspective if one assumes that the manipulation affected the salience of the relational information and thus the type of propositions that influenced implicit evaluations (i.e., propositions about CS-US co-occurrence or propositions about the valence of the CS action).

Whereas most of the EC studies on relational information contrasted information implying similarity (e.g., *starts, gives, likes*) with information implying opposition (e.g., *ends, takes away, dislikes*), a few studies (Hughes, Ye, Van Dessel, and De Houwer, 2019; Van Dessel, De Houwer, & Smith, 2018) tested the moderating effect of different types of similarity relations. For example, in one study (Hughes et al., 2019), CSs (pictures of men) co-occurred with USs (positive or negative words). Prior to the pairings, participants were either told that the CSs were directly responsible for the USs (i.e., a *cause* relation) or that the CSs merely predicted the USs (i.e., a *predict* relation). Hughes et al. found that EC of both explicit and implicit evaluation was stronger for cause relations than for predict relations. Whereas this finding is difficult to explain from an associative perspective, it makes sense

from a propositional perspective if one assumes that causal relations imply more personal responsibility than predictive relations and therefore lead to more extreme evaluations of the CSs.

Finally, in a recent study, Heycke and Gawronski (2019; see also Kukken, Hütter, & Holland, 2019) used a multinomial modeling approach to separately quantify the effects of CS-US co-occurrences and relational information on EC. Specifically, they used Hu et al.'s (2017a) learning paradigm in which pharmaceutical products (CSs) co-occurred with positive and negative health-related conditions (USs) and added a relational qualifier that indicated a *cause* or *prevent* relation between the paired components (manipulated within participants). Afterwards, participants completed a speeded choice task in which one CS was presented on each trial and participants were asked to indicate whether they would choose this product (yes or no). Although it would require too much space to explain the modeling results in detail, Heycke and Gawronski concluded that participants' responses were influenced by both CS-US co-occurrence and the instructed CS-US relations. In addition, they used different manipulations (e.g., time during encoding, time during judgment) to test the competing predictions of associative and propositional perspectives regarding the moderators of these effects. Overall, the results across five experiments were not entirely consistent with either perspective but the results could, unsurprisingly, be accommodated by both perspectives if auxiliary assumptions are made. Although the results were not conclusive at the theoretical level, the multinomial modeling approach is a promising direction for future research on the (moderators of the) effects of CS-US co-occurrence and relational information on EC.

Summary. An extensive series of experiments has revealed that EC is, at least under certain conditions, moderated by relational information. Overall, the results support the idea that in many cases, the spatio-temporal pairing of a CS and US is not a direct cause of the change in liking of the CS, as would be implied by an associative perspective, but a cue for

forming propositions about the nature of the relation between the CS and the US, as would be implied by a propositional perspective. From the latter perspective, relational information is crucial in shaping EC because it disambiguates the implications of the CS-US pairings for the nature of the CS-US relation (see De Houwer & Hughes, 2016, for an in-depth discussion of the idea that stimulus pairings function as relational cues). Although it is difficult to exclude the possibility that there are conditions under which CS-US pairings do function as a direct cause of changes in liking, one can no longer treat this assumption as a default.

On the Role of Relational Information in Implicit Evaluation

Basic idea. The evidence discussed above already indicates that relational information about novel CS-US pairs can, under certain conditions, have an impact on the implicit evaluation of the CSs involved in those pairs. Because the only systematic source of information about the CSs are the CS-US pairings in combination with the relational information, observing an impact of relational information on the implicit evaluation of CSs strongly suggests that implicit evaluation is mediated by representations that encode relational information, that is, propositions. As such, these EC studies provide clear evidence not only for a propositional perspective on EC but also for a propositional perspective on implicit evaluation. It is difficult to see how such results can be explained from an associative perspective on implicit evaluation. On the other hand, it is impossible to exclude the possibility that implicit evaluations are sometimes mediated by associations in memory. Nevertheless, based on the evidence discussed in the previous section, one should at least allow for a propositional perspective on implicit evaluation and drop the assumption that implicit evaluation is by default mediated by associations in memory.

There is also a second way in which relational information is relevant for research on implicit evaluation. Because of the dominance of associationistic thinking in research on implicit evaluation, tasks that have been designed to capture implicit evaluations typically

provide little or no relational information. For instance, typically nothing is said about the relation between the primes and targets in an evaluative priming task (EPT; Fazio et al., 1986). From a propositional perspective, this creates ambiguity about the nature of the propositions that mediate performance on those tasks. Consider the example discussed by De Houwer (2014a, p. 349) who noted that the concepts *I* and *good* can be related in different ways. For instance, people might vary both in the extent to which they believe that they are good and in the extent to which they desire to be good. In fact, the propositions *I am good* and *I want to be good* capture the essence of the difference between what has been referred to as actual self-esteem (i.e., liking of the current self) and ideal self-esteem (i.e., liking of an optimal future self; e.g., Higgins, 1998; Zentner & Renaud, 2007), respectively. Because simple associations only specify the concepts that are related (e.g., *ME* and *GOOD*) but not the nature of the relation, they cannot differentiate between these two types of self-esteem. Hence, also tasks designed to capture the automatic activation of associations between self and good (e.g., variants of the EPT and the IAT) cannot be used to differentiate between actual and ideal self-esteem. More generally, because all kinds of concepts can be related in all kinds of ways, from a propositional perspective, it is often unclear what knowledge actually drives performance in tasks like evaluative priming and the IAT.

In order to illustrate the practical and theoretical relevance of these concerns, we can again refer to the context of research on implicit self-esteem. Several tasks designed to index the automatic activation of self-good associations indicated positive implicit self-esteem in depressed students and patients, sometimes as positive or even more positive than in non-depressed controls (e.g., De Raedt, Schacht, Franck, & De Houwer, 2006; Franck, De Raedt, Dereu, & Van den Abbeele, 2007; but see Phillips, Hine, & Thorsteinsson, 2010, for meta-analytic evidence showing a negative correlation between depression and implicit self-esteem). These puzzling findings make sense if one considers the possibility that task

performance might reflect different propositions in different populations. More specifically, whereas a positive score might reflect positive actual self-esteem for non-depressed individuals, the same positive score might reflect a strong desire to be good in depressed individuals (i.e., high ideal self-esteem). A second illustration can be found in the context of alcohol addiction. Spruyt et al. (2013) observed that hospitalized alcoholics (a) have an elevated automatic tendency to avoid alcohol stimuli (as compared to a control group of individuals without drinking problems) and (b) are more likely to relapse after treatment when they have a strong automatic tendency to avoid alcohol stimuli. One interpretation of these puzzling findings is that hospitalized alcoholics implicitly endorse the proposition *I should avoid alcohol* rather than the proposition *I want to avoid alcohol*. The more they avoid alcohol on the basis of external social pressure rather than an intrinsic desire to stop drinking, the less they engage in emotional processing and learn new skills in effectively dealing with alcohol related cues in their environment. Again, it is impossible to test this hypothesis with tasks designed to capture simple associations between the concepts of *approach* and *alcohol*.

Creating tasks to capture propositions. The two examples described in the previous paragraph illustrate the need for integrating relational information in tasks that were originally designed to capture the automatic activation of associations in memory. Some variants of the IAT already incorporate such relational information. For instance, in personalized variants of the IAT, labels such as *I like* or *I want* are used instead of the standard labels *positive* and *negative* (e.g., Dewitte & De Houwer, 2008; Olson & Fazio, 2004). This allows one to design IATs that capture relationally specific propositional beliefs (e.g., to capture ideal self-esteem by using the labels *I want to be*, *I do not want to be*, *good*, and *bad*). De Houwer (2014a, p. 350) referred to these tasks as propositionalized IATs: “To the extent that the relational information in the labels actually influences the outcome of the IATs, one can conclude that the IAT effects provide an indirect measure of propositions

rather than associations. It would indeed be difficult to explain the impact of relational information on IAT effects if those effects are mediated by representations that do not specify any relational information.”

Whereas propositionalized IATs specify a relation (e.g., *want to be*) and the role of one element within the relation (e.g., the fact that it is the participant who wants something), the role of the second element is not made explicit (e.g., the fact that *good* or *bad* are the elements that are desired). Inspired by a propositional perspective on implicit evaluation, we created new tasks that specify all elements of a proposition. For instance, in a Relational Responding Task (RRT; De Houwer, Heider, Spruyt, Roets, & Hughes, 2015), participants are asked to act as if they endorse certain beliefs. The logic behind the task is that participants will find it easier to act as if they endorse beliefs when they actually possess those beliefs than when they do not possess those beliefs. In a study on body dissatisfaction, Heider, Spruyt, and De Houwer (2018) either presented *am* statements related to actual body-image (e.g., *I am thin, I weigh too much*) or *want to be* statements related to ideal body-image (e.g., *I want to be thin, I want to weigh too much*). For each set of items separately, participants performed one task in which they were asked to respond *true* to statements relating themselves with thin (e.g., *I am thin, I want to be thin*) and *false* to statements relating themselves with overweight (e.g., *I weigh too much, I want to weigh too much*), as well as a second task in which the response assignments were reversed (i.e., respond *false* to statements relating themselves with thin and *true* to statements relating themselves with overweight). Actual and ideal body satisfaction were assessed by calculating the difference in the speed of performing the two tasks for *am* (i.e., actual) and *want to be* (i.e., ideal) items, respectively. Whereas the index of actual body image significantly differed between participants who scored high versus low on a questionnaire measure of body satisfaction, the index of ideal body image did not differ.

Summary. Evidence showing that relational information moderates EC of implicit evaluations strongly supports the idea that implicit evaluation is at least in some cases mediated by propositions. This implies that tasks that were designed to capture implicit evaluations, but that do not specify relational information, could produce effects that reflect different kinds of propositions. Such ambiguity is reduced in tasks like the RRT and propositionalized variants of the IAT that do specify relational information.

On the Role of Instructions about Spatio-temporal Relations

Basic idea. As noted above, from an associative perspective, the spatio-temporal pairing of stimuli is seen as a direct cause of the formation of associations. From a propositional perspective, however, stimulus pairings are a source of information about when and where stimuli occur, more specifically, about the contingency between the spatio-temporal presence of both. This contingency information could, however, also be provided via other sources such as instructions. Based on this idea, a series of experiments was conducted to examine the effect of contingency instructions on stimulus evaluations. For instance, De Houwer (2006, Experiment 2) informed participants that in an upcoming phase of the study, names of one fictitious social group (e.g., Niffites) would always be followed by positive pictures whereas names of a second fictitious social group (e.g., Luupites) would always be followed by negative pictures. Before the stimulus pairs were actually presented, participants completed an IAT to assess their implicit evaluation of the two social groups. Results showed that responses were faster in IAT blocks that linked the first group with positive stimuli, indicating that participants implicitly preferred this group even though they had not seen any of the pairings.

Stimulus-contingency instructions are one instance of a broader class of regularity instructions, that is, instructions about regularities in the spatio-temporal presence of events. Whereas stimulus-contingency instructions focus on a regularity in the presence of two

stimuli, other instructions specify other types of regularities. For instance, approach-avoid instructions (e.g., *in the next phase, approach names of Niffites and avoid names of Luupites*) refer to stimulus-response contingencies. Likewise, mere exposure instructions (e.g., *in the next phase, names of Niffites will appear often whereas names of Luupites will be presented infrequently*) refer to a regularity in the presence of one stimulus (see De Houwer, Barnes-Holmes, & Moors, 2013). It has now been shown that also approach-avoidance instructions (Van Dessel, De Houwer, Gast, & Smith, 2015) and mere exposure instructions (Van Dessel, Mertens, Smith, & De Houwer, 2017) lead to changes in liking that resemble the changes in liking observed after experiencing the actual events that constitute the regularity (i.e., after actually approaching and avoiding stimuli or actually seeing stimuli presented often or infrequently). Because much of the research on regularity instructions has focused on approach-avoidance instructions and because approach-avoidance instructions are similar to stimulus-contingency instructions in that both refer to the contingency between events, we will review studies on the effects of both types of contingency instructions.

Note that contingency instructions, as well as other regularity instructions, are similar to persuasive messages in that they are verbal in nature. However, they differ with regard to the content of the message. Whereas contingency instructions provide information about the spatio-temporal properties of stimuli (when and where are stimuli presented), persuasive messages provide information about the evaluative properties of stimuli (whether they are good or bad in certain respects). Given that persuasive messages sometimes evoke reactance, contingency instructions could provide a subtler way of changing evaluations (see De Houwer & Hughes, 2016, for a more detailed discussion of the relation between effects of contingency instructions and persuasive messages).

Whereas persuasive messages have been studied extensively by attitude researchers, a systematic study of contingency (and other regularity) instructions had to await the

introduction of a propositional perspective on EC and other forms of evaluative learning. This observation is interesting as such because in principle, effects of contingency instructions on liking could also be explained on the basis of an associative perspective. For instance, it could be argued that giving an instruction about pairings is sufficient to co-activate the representations of the paired stimuli (e.g., the representations of *Niffites* and *good*) and thus to strengthen the relation between the two (e.g., Fazio, 2007; Gawronski & Bodenhausen, 2006). Nevertheless, these ideas were never really pursued, perhaps because they raise important questions about the validity of associative models, such as the question of how an associative mechanism would deal with the grammatical nature of instructions (e.g., whether the instructions that *Niffites* and positive pictures do NOT co-occur also lead to the formation of associations; see De Houwer & Hughes, 2016). In contrast, the idea that propositions mediate (changes in) stimulus evaluations did give rise to many studies on the effects of contingency instructions that we will review in the remainder of this section.

Effects on explicit and implicit evaluations. It has been shown that contingency instructions have an impact on both explicit evaluations (i.e., evaluative ratings; e.g., Van Dessel et al., 2015) and implicit evaluations (i.e., scores on the IAT, personalized IAT, Affect Misattribution Procedure (AMP), and EPT; e.g., De Houwer, 2006; Van Dessel et al., 2015). In fact, in the studies conducted so far, contingency instructions have been shown to influence all the dependent variables that are known to be affected by actual stimulus pairings. This includes automaticity parameters in Multinomial Processing Tree (MPT) analyses (Hütter & De Houwer, 2017) as well as the activation parameter in the Quad model (Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005) of IAT performance (Smith, Calanchini, Hughes, Van Dessel, & De Houwer, 2019), which are both derivatives of task performance designed to tap into automatic processes. MPT and Quad model analyses adopt a multinomial modelling approach to separately quantify automatic versus controlled

processes that are assumed to jointly determine task performance. This approach thus offers a way to sidestep the questionable assumption that overall task performance (e.g., an IAT score) can provide a pure index of automatic processes (e.g., Heycke & Gawronski, 2019; Payne, Burkley, & Stokes, 2008). The fact that contingency instructions influence the automaticity and activation parameters of MPT and Quad models is particularly striking because it was initially assumed that these MPT and Quad parameters by definition reflect associations in memory. Although it is difficult to exclude the possibility that contingency instructions also result in associations and thus only indirectly influence these parameters via their impact on associations, the fact that instructions can have this impact gives credence to the claim that propositions influence automatic aspects of evaluative responding. At the very least, one has to give up the default assumption that those parameters index associations.

Interestingly, Van Dessel, De Houwer, Gast, Smith, and De Schryver (2016; see also Van Dessel, Gawronski, Smith, & De Houwer, 2017, for a replication and extension) observed an impact of contingency instructions on implicit evaluations even when there was no impact on explicit evaluations. In their study, participants were first informed about the evaluative properties of Niffites and Luupites (e.g., that Niffites are peaceful whereas Luupites are violent). Afterwards, they were given instructions about an upcoming task in which they would either approach names of Niffites and avoid names of Luupites or vice versa. Explicit evaluations (i.e., ratings on a Likert scale) and implicit evaluations (i.e., performance on an IAT) of Niffites and Luupites were registered immediately after all instructions had been provided. Instructions about evaluative properties had an impact on both explicit and implicit evaluations such that participants exhibited both an implicit and explicit preference for the social group that had positive traits. In contrast, however, contingency instructions only affected implicit evaluations. Specifically, participants' implicit but not explicit preference for the social group that had positive traits was reduced when

participants were instructed to avoid rather than approach this group (Figure 3).

From a propositional perspective, the latter finding can be explained in the following way: Both the persuasive message about the evaluative properties of the social groups as well as the instructions about the contingency between approach-avoidance responses and the social groups are likely to result in the formation of propositions. However, it is reasonable to assume that participants consider the persuasive message as being more diagnostic than the contingency instructions for determining whether they should like or dislike Niffites and Luupites. By definition, explicit evaluations are easier to control than implicit evaluations. Hence, participants might be able to base their explicit evaluation solely on propositions arising from the persuasive message. In contrast, participants might not, however, be able to fully prevent an impact of propositions arising from contingency instructions on implicit evaluations. The results of Van Dessel et al. (2016, 2017) therefore are one example of the more general point that a propositional perspective can explain dissociations between implicit and explicit evaluations by assuming that both types of evaluations are influenced differentially by different types of propositions. Note that the propositional perspective thus locates dissociations between implicit and explicit evaluation at the level of the retrieval of representations rather than at the level of the formation of representations (De Houwer, 2018).

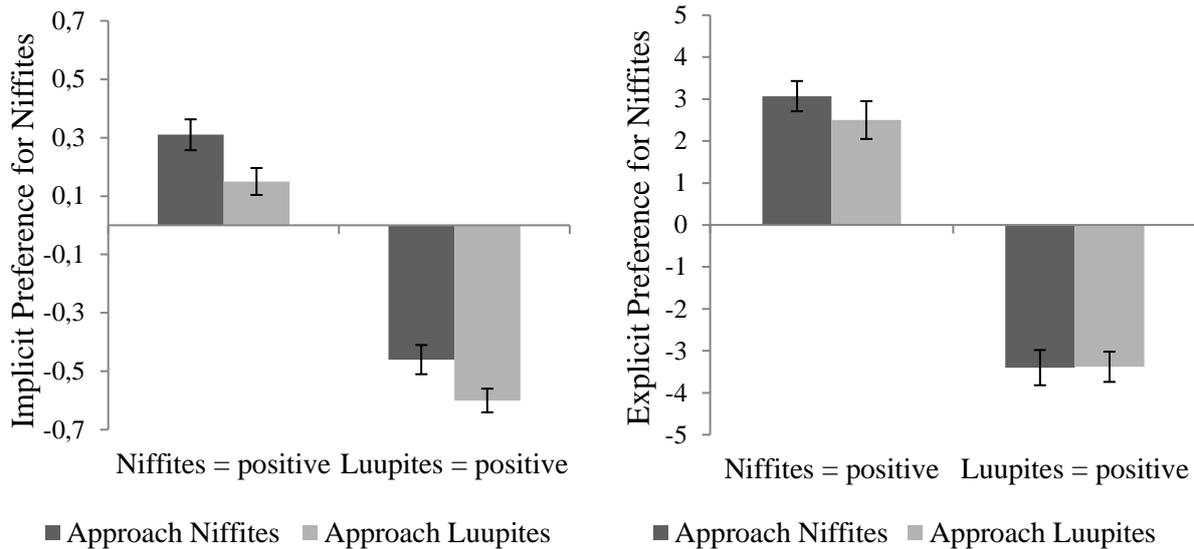


Figure 3. Implicit and explicit preference for Niffites over Luupites (indexed by IAT scores and evaluative ratings, respectively) as a function of trait information (Niffites have positive traits and Luupites have negative traits, Luupites have positive traits and Niffites have negative traits) and contingency instructions (approach Niffites and avoid Luupites, approach Luupites and avoid Niffites). Error-bars reflect +1 and -1 standard error of the mean. *Adapted from Van Dessel, P., De Houwer, J., Gast, A., Smith, C. T., & De Schryver, M. (2016). Instructing implicit processes: When instructions to approach or avoid influence implicit but not explicit evaluation. Journal of Experimental Social Psychology, 63, 1-9.*

Effects on the magnitude of the changes in liking. In an extensive series of studies, Kurdi and Banaji (2017, 2019) found striking similarities and differences between the impact of contingency instructions and actual contingencies on the magnitude of changes in liking.⁴ A first set of studies revealed more extensive changes in liking immediately after participants received instructions about upcoming contingencies between the presentation of specific stimuli (e.g., the social groups Niffians and Laapians) and positive or negative stimuli than immediately after participants actually experienced these contingencies (Kurdi & Banaji, 2017; but see Gast & De Houwer, 2013, who observed equally strong effects). Moreover, actually experiencing stimulus pairings after receiving instructions about those pairings did

⁴ Although the instructions of Kurdi and Banaji (2017; see <https://osf.io/w6qnb/>) specified contingencies (e.g., *Laapians will always be paired with pleasant things. Niffians will always be paired with unpleasant things.*), they could also be understood as specifying evaluative properties (e.g., *Laapians are linked to good things and Niffians are linked to bad things; so please remember well: Laapians = pleasant and Niffians = unpleasant*).

not result in bigger effects than merely receiving the contingency instructions (Figure 4). A second set of studies showed that, when liking was assessed after a delay, actual contingencies produced stronger effects than contingency instructions alone but, surprisingly, also than actual contingencies in combination with contingency instructions (Kurdi & Banaji, 2019).

Kurdi and Banaji (2017) noted that these findings are difficult to reconcile with associative models of attitudes but do make sense from a propositional perspective. When a participant experiences events that constitute a regularity (e.g., the repeated co-occurrence of two stimuli on a computer screen), it is up to the participant to discover the regularity and to generate a proposition about it. A contingency instruction, on the other hand, provides information about the regularity directly to the participant. Hence, if anything, it is more likely that participants will have propositional knowledge about the stimulus pairings after contingency instructions than after actual pairings. This alone can explain why effects of contingency instructions on liking are often stronger than the effects of actual contingencies. Based on the additional assumption that contingency instructions and actual contingencies provide the same information, one can also explain that their combined effect is as strong as the impact of each source separately. Finally, based on the fact that self-generated propositions are remembered better than other propositions, one can explain why effects of actual contingencies are more durable (because it forces participants to generate the crucial propositions themselves) and why this increased durability can be undone by adding contingency instructions to actual contingencies (because participants no longer have to self-generate the crucial proposition on the basis of the actual presentation of the stimuli).

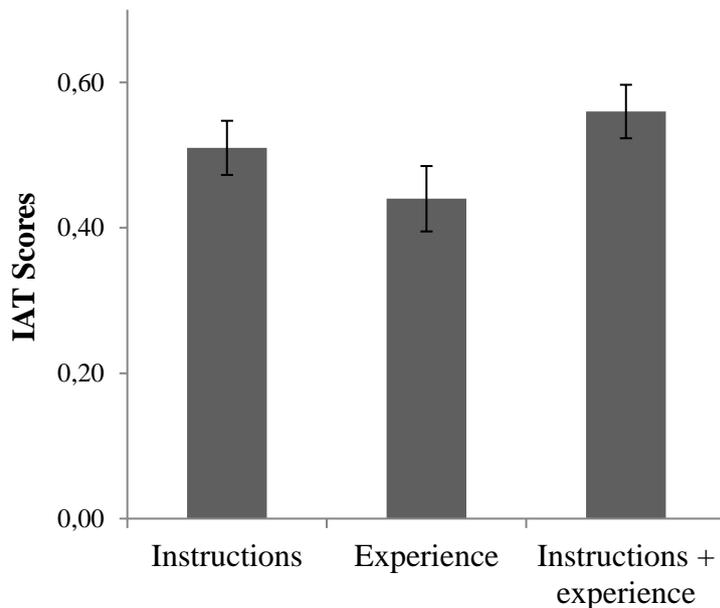


Figure 4. IAT scores indicating a preference for stimuli (instructed to be) paired with positive stimuli as a function of learning procedure (learning via instructions, experience, or both). Error-bars reflect +1 and -1 standard error of the mean. *Adapted from Kurdi, B., & Banaji, M. R. (2017; Study 1). Repeated evaluative pairings and evaluative statements: How effectively do they shift implicit attitudes? Journal of Experimental Psychology: General, 146, 194–213.*

Moderators of the effects of contingency instructions. Whereas Kurdi and Banaji (2017, 2019) focused on the magnitude of effects, other researchers examined parallels between moderators of the effects of contingency instructions and moderators of the effects of actual contingencies. Prior research on EC showed that the impact of actual CS-US pairings on liking is (a) only weakly affected by presentations of the CS on its own after the CS-US pairing (i.e., little or no extinction), (b) strongly affected by pairing the CS with another US of the opposite valence (i.e., strong counterconditioning), (c) not affected by CS-only presentations that are intermixed with the initial CS-US pairings (i.e., no effect of the strength of the contingency), and (d) strongly dependent on participants being able to recall the CS-US contingencies during test (see Hofmann et al., 2010, for a review and meta-analysis of the relevant findings). In parallel, studies have shown that the impact of instructions about a CS-US contingency is (a) only weakly affected by the instruction that the CS will be presented on its own after the CS-US pairings (i.e., little or no instructed

extinction; Gast & De Houwer, 2013), (b) significantly reduced when participants are told that the CS will also be paired with another US of the opposite valence (i.e., instructed counterconditioning; Gast & De Houwer, 2013), (c) not affected by instructions about the number of CS-only trials that will be intermixed with CS-US trials (i.e., no effect of the instructed strength of the contingency; De Houwer, Mattavelli, & Van Dessel, 2019); and (d) strongly dependent on participants correctly remembering the instructed CS-US contingencies (e.g., Gast & De Houwer, 2013). Although these parallels do not prove that effects of contingency instructions and effects of actual contingencies are both mediated by propositional representations, they are clearly in line with this hypothesis.

In addition to revealing parallels between the moderators of the effects of actual and instructed stimulus pairings, these studies also revealed some other intriguing findings. First, counterconditioning *instructions* (i.e., *in the next phase, the CS will occur with a US of the opposite valence of the US it co-occurred with in the first phase*) also counteract the effect of actual CS-US *pairings* (Gast & De Houwer, 2013, Experiment 2b). This finding shows that information from different sources (i.e., actual CS-US pairings and instructions about CS-US pairings) can combine to determine changes in liking, which provides important support for the core assumption of propositional models that propositions can originate from a combination of different sources. Second, whereas De Houwer et al. (2019) did not observe an impact of instructions about the strength of the CS-US relation on changes in liking, they did observe an impact of these instructions on US expectancy ratings (i.e., the extent to which the US was expected to follow the CS). This dissociation of CS-liking and US-expectancy had been observed before but only on the basis of actual CS-US pairings. The original finding was interpreted as evidence for the existence of an evaluative learning system that is separate from an expectancy learning system (e.g., Hermans, Vansteenwegen, Crombez, Baeyens, & Eelen, 2002). De Houwer et al., however, pointed out that the dissociation might also reflect

the operation of a single propositional learning system if one assumes that changes in liking reflect propositions about CS-US co-occurrences whereas changes in US expectancy reflect propositions about the strength of the CS-US contingency. Based on this propositional account, they predicted that the same dissociation should arise as the result of contingency instructions, as was observed. De Houwer et al.'s study not only illustrates the predictive value of the propositional perspective but also supports the general point that dissociations in (evaluative) learning do not necessitate multiple learning systems. Instead, dissociations can also be located at the level of retrieval of information (De Houwer, 2018; Gawronski & Bodenhausen, 2018; Van Dessel, Gawronski, & De Houwer, 2019).

Within this line of research, however, there is one finding that questions the validity of propositional models. More specifically, Hu, Gawronski, and Balas (2017b) found that the effect of actual CS-US pairings on implicit evaluations of the CS can be undone by actual counterconditioning (i.e., presenting the same CS with a US of the opposite valence) but not by counterconditioning instructions (i.e., merely instructing participants that the CS will be paired with a US of the opposite valence) (Figure 5). This observation is especially striking given that the impact of CS-US pairings on explicit evaluation was moderated by counterconditioning instructions. Assuming that counterconditioning instructions and actual counterconditioning provide the same information, from a propositional perspective, one would likely predict similar effects of both interventions.

In this respect, it is again important to note that dissociations between effects on implicit and explicit evaluations could reflect not only differences in learning-related but also retrieval-related processes. As noted by Gawronski and Bodenhausen (2018), counterconditioning requires the integration of conflicting sources of information. It is possible that this integration requires time and resources which is available when providing explicit ratings of liking but not when completing tasks that index implicit evaluations. The

fact that implicit evaluations are impacted by actual counterconditioning but not counterconditioning instructions could be due to the fact that the former provide more opportunity for the integration of conflicting information during the counterconditioning phase. From this perspective, the results of Hu et al. (2017b) could be accounted for from a propositional perspective. In light of the findings of Kurdi and Banaji (2019), it would also be interesting to examine the effect of a combination of actual counterconditioning and instructed counterconditioning on changes in implicit and explicit evaluations. From an associative perspective, adding counterconditioning instructions should not change the effect of actual counterconditioning trials. If it does, it would suggest that the results of Hu et al. can be traced back to the quality of encoding or integration of information. Also note that both studies reported by Hu et al. used evaluative priming effects as an index of implicit evaluations. Because conditioned changes in evaluative priming effects are typically small and unreliable, it would be worthwhile to replicate their studies using other tasks such as the IAT.

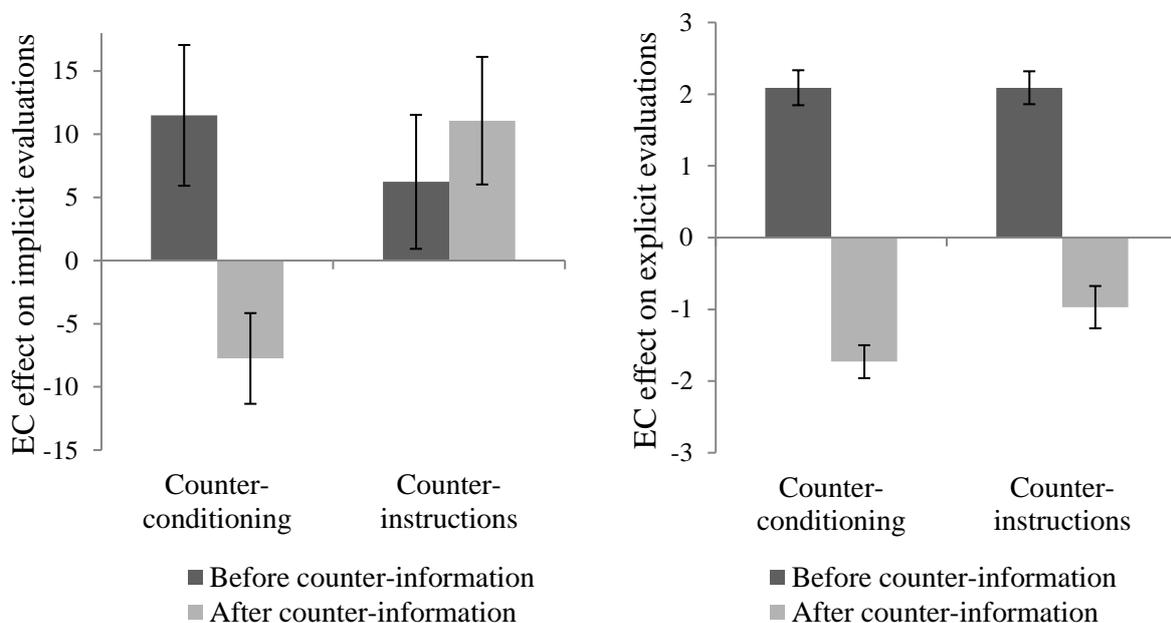


Figure 5. Effect of pairings on implicit and explicit evaluations as a function of Time (before, after counter-information) and counter-information learning procedure (counter-conditioning, counter-instructions). Error-bars reflect +1 and -1 standard error of the mean. Adapted from Hu, X., Gawronski, B., & Balas, R. (2017b; Experiment 1). *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 and Personality Science*, 8, 858-866.

Summary. Although contingency instructions have profound effects on both explicit and implicit evaluations, changes in implicit evaluations can occur even in the absence of changes in explicit evaluations. From a propositional perspective, such dissociations can arise at the level of retrieving information from memory, that is, when different propositions underlie different instances of stimulus evaluation. Whereas immediate effects of contingency instructions on liking are often stronger than the effects of actual contingencies, the reversed is true when liking is measured after a delay. This pattern of results could be related to the fact that contingency instructions directly provide propositions about contingencies whereas participants need to self-generate these propositions when being exposed to events that constitute a contingency (e.g., the pairing of stimuli). The fact that the immediate effects of contingency instructions and actual contingencies are not additive makes sense if both sources provide the same relational information. This also explains why there is a parallel between the moderators of the effects of actual contingencies and the moderators of the effects of contingency instructions.

On the Impact of Truth Information

Basic idea. Because the information encoded in propositional representations specifies assumptions about the nature of events, propositional representations can be said to have a truth value. Therefore, if stimulus evaluations reflect propositional representations, then they might also depend on information that can be used to determine the truth value of propositions. In fact, much of the available research on persuasion is related to examining the impact of truth information on stimulus evaluation. As noted above, persuasive messages provide information about the evaluative properties of stimuli. It has been shown that the impact of those messages on explicit evaluations depends on the extent to which the content

of those messages is thought to be valid, believable, and diagnostic for inferring whether a stimulus is good or bad (see Petty & Cacioppo, 1986, for a relevant review). For instance, the extent to which you like a novel person Bob after being told that Bob helps old ladies cross the street will depend on whether you are told that this information is valid or whether you receive it from a credible source. Likewise, you will dislike Bob more after being told that he is a child molester than after being told that Bob did not help an old lady cross the street, because the first piece of information is more diagnostic of the character of Bob than the second piece of information.

More recently, researchers started to examine whether the impact of persuasive messages on implicit evaluation is also moderated by truth information (see Cone, Mann, & Ferguson, 2017, for a recent review). Based on the assumption that implicit evaluation is mediated by propositions, one would indeed expect not only that persuasive messages can lead to changes in implicit evaluations but also that the impact of persuasive messages on implicit evaluation depends on truth information. Although one could think of ways in which persuasive messages influence implicit evaluations via the formation of associations (but see De Houwer & Hughes, 2016) and perhaps even why this effect would depend on truth information (e.g., by assuming that only propositions that are evaluated as valid and diagnostic lead to changes in associations), it is again striking to see that systematic research on the impact of truth information on implicit evaluations was undertaken only after researchers considered (or tried to falsify) the idea that propositions mediate implicit evaluation.

As noted above, from a propositional perspective, both actual contingencies (e.g., CS-US pairings) as well as contingency instructions provide contingency information that, in combination with other information, can be used to infer whether a stimulus is good or bad. Hence, as is the case for persuasive messages, the impact of actual contingencies and

contingency instructions could also depend on information about whether contingency information reflects the actual contingency (i.e., is valid), as well as on information about the extent to which contingency information can be used to infer whether a stimulus is good or bad (i.e., is diagnostic). In the remainder of this section, we first review research on the effects of persuasive messages, after which we turn our attention to the effects of actual contingencies and contingency instructions.

Validity of persuasive messages. In a seminal study by Rydell, McConnell, Mackie, and Strain (2006), participants learned about a person named Bob by being exposed to two blocks of trials in which they read 50 positive and 50 negative behavioral statements that were each followed by verbal instructions that Bob would or would not engage in the described behavior. In one (positive valid) learning block, instructions indicated that Bob would consistently engage in all of the positive and none of the negative behaviors or vice versa, whereas opposite instructions were given in a second (negative valid) learning block. It was counterbalanced whether participants first completed the positive valid or the negative valid learning block. Importantly, before each behavioral statement, a valenced word (e.g., the word ‘love’) was presented briefly. These words were consistently positive in the negative valid learning block and were consistently negative in the positive valid learning block. Intriguingly, whereas explicit evaluations of Bob reflected the valence of the behavioral information that was learned to be valid for Bob (e.g., Bob was evaluated positively after learning that positive behavioral statements were valid), implicit evaluations, that were assessed with an IAT, did not. Instead, IAT scores reflected the valence of the words that were presented briefly prior to the presentation of the behavioral statements (Figure 6).

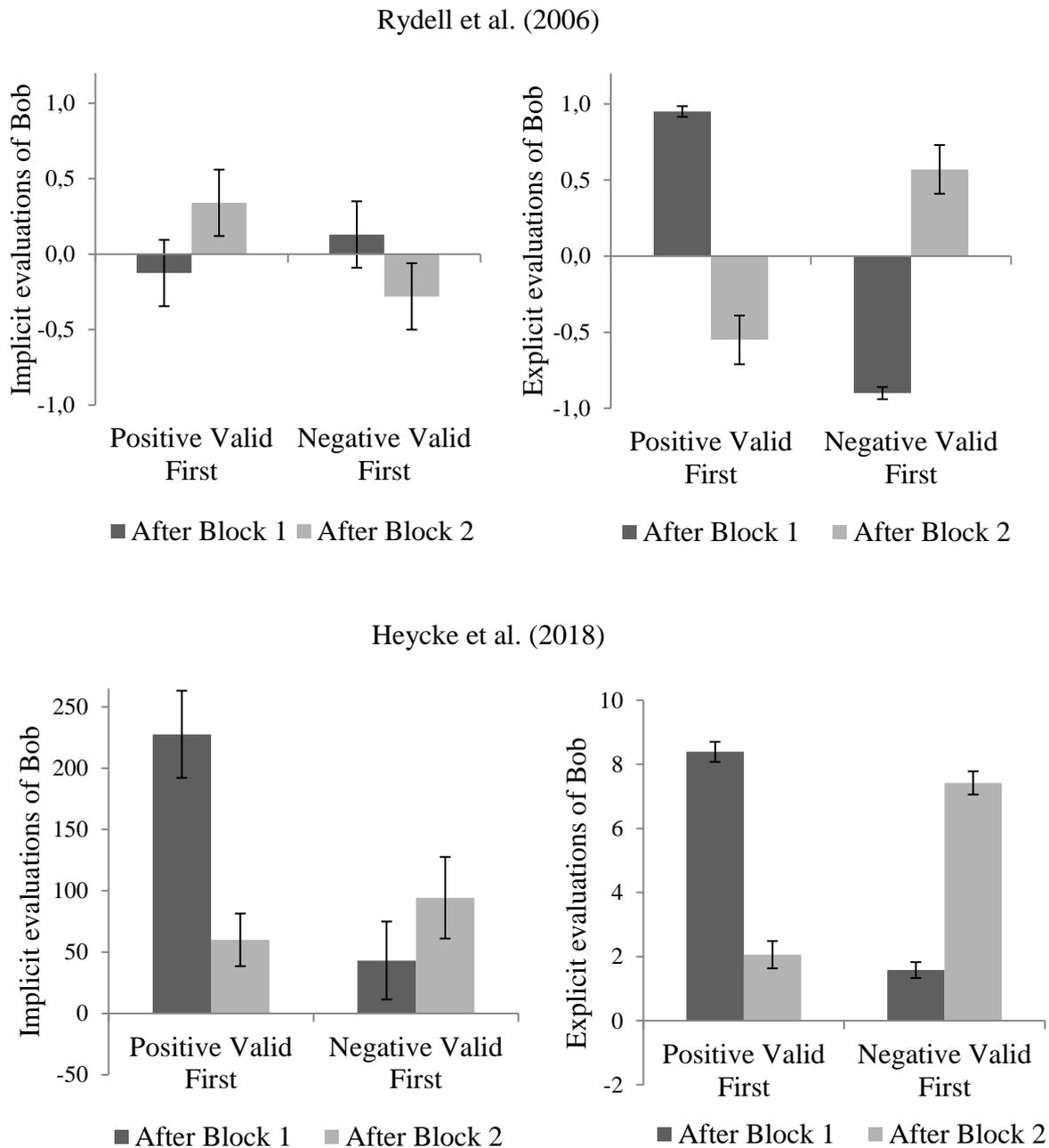


Figure 6. Implicit and explicit evaluations of Bob as a function of Block Order (First Learning Block indicates that positive behaviors or negative behaviors are valid), Time of Evaluation (after first learning block, after second learning block), and study (Rydell et al., Heycke et al.). Error-bars reflect +1 and -1 standard error of the mean. *Adapted from 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, 954–958, and Heycke, T., Gehrman, S., Haaf, J. M., & Stahl, C. (2018). Of two minds or one? A registered replication of Rydell et al. (2006). Cognition and Emotion, 32, 1708-1727.*

This initial study generated a lot of interest in the question of whether validity information can influence implicit evaluations. Importantly, in contrast to the study by Rydell et al., many of these subsequent studies did find evidence for the impact of truth information

on implicit evaluations (Boucher & Rydell, 2012; Mann, & Ferguson, 2015; Peters & Gawronski, 2011). For example, Boucher and Rydell (2012) found a strong effect of validity instructions on implicit evaluations (as captured by scores on the AMP) when they used a similar learning procedure as Rydell et al. that did not include the brief presentations of valenced words and made the validity instructions highly salient. Notably, a recent study by Heycke, Gehrman, Haaf, and Stahl (2018) replicated the procedure of Rydell et al. as closely as possible and found that the validity instructions strongly determined both explicit evaluations and implicit evaluations (as captured by IAT scores) (Figure 6).

Other research established that the impact of validity instructions on implicit evaluations depends on specific moderators. Most importantly, as seems to be the case with the impact of relational information, the impact of validity information on implicit evaluation depends on when the information is given. In a study by Peters and Gawronski (2011), participants first read either positive or negative behavioral descriptions of four target men and were then informed about whether the behavior provided valid information about the character of the target person. Implicit and explicit evaluations reflected the validity information when instructions were provided immediately after participants read the behavioral descriptions. However, when a delay was introduced such that participants only learned about the validity of the behavioral descriptions after they had read all the behavioral descriptions of the four men, validity information had a much stronger impact on explicit than on implicit evaluation. Moran, Bar-Anan, and Nosek (2017) replicated this finding, but showed that the observed moderation depended on the specific task that was used to measure implicit evaluations. Implicit evaluations were more sensitive to delayed validity information when measured with the AMP than with the IAT.

The impact of timing was also demonstrated by Gregg, Seibt, and Banaji (2006). In their studies, participants first learned about two social groups, one that was described as

engaging in positive behavior (e.g., kind) and one that was described as engaging in negative behavior (e.g., aggressive). After completing implicit (IAT) and explicit evaluation measures, they learned that the previous information was the result of a mixed-up and that each group was in fact characterized by the information provided about the other group (i.e., reversal instructions). Whereas participants' explicit preference for one of the two groups completely reversed after participants received the validity information, IAT scores still reflected the initial preferences. However, Wyer (2016) replicated this study and found that the mixed-up manipulation did not have this dissociative effect when participants had enough opportunity (e.g., more time) and motivation (e.g., reduced need for closure) to elaborate on the validity information. These results suggest that dissociations between explicit evaluations and implicit evaluations (as measured with the IAT) can arise on the basis of delayed validity information and that these dissociations are the result of processes operating during retrieval (i.e., at the time of evaluation) rather than during learning. Because implicit evaluation measures typically provide less opportunity and motivation to engage in a comprehensive validation of activated information (Gawronski & Bodenhausen, 2006, 2018), validity information might influence implicit evaluations only when this information has already been extensively elaborated upon (e.g., integrated in a person's belief network).

Believability of persuasive messages. Whereas all studies reviewed until now manipulated the perceived validity of persuasive messages by providing explicit instructions about whether information is true or false, there have also been studies that manipulated validity in a more subtle manner, that is, by varying the believability of the persuasive message. Cone, Flaherty, and Ferguson (in press) replicated the study by Gregg et al. (2006) in which two fictitious social groups were first described as consistently engaging in positive or negative behaviors (e.g., the negative group molested the positive group) and then received new information that the groups had shifted character (e.g., the negative group showed

remorse and the positive group engaged in vengeful behavior). As noted above, Gregg et al. found that explicit but not implicit evaluations shifted on the basis of the counterattitudinal information and interpreted this as evidence that implicit evaluations do not allow for rapid changes. In contrast, Cone et al. observed that the new information was readily incorporated in implicit evaluations, but this revision was observed only when participants believed the events described in the narrative to be plausible. In line with this result, we recently observed rapid revision in the Gregg et al. paradigm by boosting the processing of the counterattitudinal information via hypnotic suggestions (Van Dessel & De Houwer, 2019). These findings support the idea that evaluative inferences underlie implicit evaluation and these inferences are more likely to incorporate new (counterattitudinal) information when this information is considered to be a believable and important source of information.

The role of information believability has also been studied by examining the role of source factors, that is, features of the source that provides the persuasive message. Smith, De Houwer, and Nosek (2013) provided participants with persuasive information about unfamiliar laundry detergents. Results of four experiments showed that information about source trustworthiness and source expertise influenced both explicit and implicit evaluations (measured with the AMP and IAT). Relatedly, Cone et al. (in press) found evidence for the impact of source credibility on implicit evaluation in the context of attitude revision. In their study, participants first learned 100 positive behaviors about a person named Bob and then learned one piece of highly diagnostic negative information (e.g., that Bob had been arrested for domestic abuse). This novel information strongly influenced implicit evaluations when participants were asked to imagine that it came from a reliable source but not when they were asked to imagine that the person who provided this information had a good reason for spreading false rumors about Bob.

Though the effect of believability on implicit evaluations is now well-established,

moderators of this effect have not been systematically examined. In one experiment, Smith et al. (2013) found that a manipulation of source trustworthiness influenced implicit evaluations when participants had been instructed to remember a 2-digit but not a 7-digit number throughout the experiment. This accords with previous findings about the effects of direct instructions about information validity (as we discussed in the previous paragraph), suggesting that the effect of (source credibility) instructions on implicit evaluations requires ample opportunity to elaborate on the instructions.

Diagnosticity of persuasive messages. Several recent studies have examined the role of information diagnosticity in the revision (rather than the formation) of implicit evaluations (see Cone et al., 2017). In contrast to the idea that the modification of implicit evaluations requires changes in associations and is therefore slow and gradual (Gregg et al., 2006; see also Gawronski & Bodenhausen, 2006), from a propositional perspective, persuasive messages can produce immediate changes in implicit evaluations, provided that they offer highly diagnostic information about whether something is good or bad.

A study by Cone and Ferguson (2015) provided strong evidence for this idea, showing that participants who had learned many pieces of positive information about a person named Bob, exhibited a rapid negative shift in implicit evaluations of Bob (measured with an AMP) when they learned new information that they considered more diagnostic of Bob's true character (e.g., that Bob was a convicted child molester). Van Dessel, Cone, Gast, and De Houwer (2019) observed similar effects with a less extreme manipulation of diagnosticity (i.e., a more moderate difference in the rated diagnosticity of initial and new information) and a different measure of implicit evaluation (i.e., EPT rather than AMP). Calanchini and Cone (2019) recently observed an impact of diagnosticity also on the activation parameter of the Quad model (Conrey et al., 2005), which is the parameter that is assumed to reflect automatic processes. Hence, the moderating effect of information about

the diagnosticity of persuasive messages can be observed across a variety of manipulations and indices of implicit evaluation.

We recently examined whether diagnosticity effects depend on the familiarity of the attitude object (Van Dessel, Ye, & De Houwer, 2018). From a propositional perspective, it should be possible for a single piece of highly diagnostic information to not only produce immediate changes in implicit evaluations of unfamiliar attitude objects (e.g., a stranger named Bob; e.g., Cone & Ferguson, 2015) but also in deep-rooted implicit evaluations of a well-known and unambiguously positive attitude object, such as Mahatma Gandhi. Results supported this idea, showing that participants who learned that Gandhi had denied his wife of medical treatment resulting in her death, exhibited a rapid change in implicit evaluations as measured with the AMP and EPT. Notably, this revision was not observed on the IAT. Modeling analyses suggested that this might be due to IAT performance being strongly affected by non-evaluative processes (i.e., cognitive recoding of the IAT categories: see also Rothermund et al., 2009).

Another possible moderator of the effect of diagnosticity is the learning context. From an associative perspective, rapid revision of implicit evaluations can be explained by referring to the context-dependent activation of associations (Gawronski & Cesario, 2013). Drawing on animal research, Gawronski and Cesario suggested that the activation of associations might strongly depend on the match between the context of evaluative learning and testing such that implicit evaluations will strongly reflect associations that are learned in a context that is similar to the evaluation context. From this perspective, diagnostic counterattitudinal information might be stored in a contextualized associative representation and lead to changes in implicit evaluations that are highly context-dependent. To test this, Brannon and Gawronski (2017) presented initial information about a target person against a colored background and then presented highly diagnostic counterattitudinal information

against a background of different color. In line with the idea that diagnosticity effects depend on processes related to validation rather than processes related to the contextual retrieval of (associative) information during evaluation, implicit evaluations reflected the diagnostic information independent of the background that was present (see Gawronski et al., 2018, for a review).

Finally, also retrospective effects of diagnosticity information have been observed. In a study by Mann and Ferguson (2015), participants read a narrative about a man named Francis West who was described as breaking into and causing extensive damage to the homes of two of his neighbors and removing precious things from the bedrooms. This story produced highly negative implicit evaluations of Francis, which were readily revised when participants were informed that that the houses had in fact been on fire, and Francis broke in to rescue the children that were trapped inside. Effects depended on the extent to which the novel information not only negated the new information but also provided information that allowed for reinterpretation of the prior information and were highly resistant across time (Mann & Ferguson, 2017).

Validity of actual contingencies. From a propositional perspective, both explicit and implicit evaluations should depend not on actual events but on propositional beliefs about events. If those beliefs are distorted, then evaluations should reflect the distorted beliefs rather than the actual events. One source of belief distortions are distortions of memory. Many studies revealed that memory plays an important role in EC and other types of evaluative learning (e.g., Gast, 2018; Van Dessel, De Houwer, & Gast, 2016). For instance, the bulk of evidence on EC suggests that typical effects are observed only when participants report accurate memory of the pairings at the time of evaluation (see Gast, 2018). Importantly, participants who report memory of incorrect pairings sometimes exhibit effects that accord with these reports rather than with the actual pairings (e.g., Bar-Anan, De

Houwer, & Nosek, 2010). More direct evidence for the impact of memory distortions was found in recent studies that included a manipulation of pairing memory on the basis of verbal instructions. In an initial study on this topic, Gast and Kattner (2016) instructed participants to remember or forget previously experienced stimulus-stimulus pairings. This moderated EC effects such that typical EC effects were smaller (but not absent) for the stimuli that were part of the to-be-forgotten pairings. Benedict, Richter, and Gast (2019) extended these findings in a study where participants first viewed CS-US pairings and completed a filler task and were then asked questions about the pairings they had seen. These questions either suggested that specific CSs (1) had been paired with the USs they had been paired with (true information), or (2) had been paired with USs of opposite valence than the USs in the actual pairings (misinformation). Interestingly, the misinformation manipulation produced an explicit preference that was in accordance with the manipulation rather than the actual pairings.

A recent study established that verbal suggestions of inaccurate pairings can have a strong impact even on implicit evaluations of well-known stimuli (Van Dessel, De Houwer, Gast, Roets, & Smith, 2019). In this study, participants went through a hypnosis induction procedure followed by the suggestion that participants had approached pictures of Black people and had avoided pictures of White people. Even though they never actually performed this approach-avoidance task, participants who received the hypnotic suggestions exhibited a reduced implicit preference for White people over Black people on a race IAT compared to a control group.

Diagnosticity of actual pairings. As noted above, actual contingencies as well as contingency instructions provide one source of information on the basis of which inferences can be made about the evaluative properties of stimuli. Such inferences also require additional propositions, for instance, the proposition that similar things tend to co-occur in the world (see Van Dessel, Hughes, & De Houwer, 2019, for more details). If participants

evaluate these additional propositions as being valid, then the pairing of a novel (CS) and valenced stimulus (US) provides a basis for inferring the valence of the novel stimulus. This is, of course, a dubious inference because the validity of the additional assumptions is uncertain. For instance, the assumption that similar things tend to co-occur is a generalization that certainly does not always hold (e.g., opponents also occur together in space and time). In other words, in most contexts, actual contingencies provide only low diagnostic information about the evaluative properties of stimuli. This could explain why the effects of actual contingencies on liking can be counteracted by the effects of (valid and diagnostic) persuasive messages (e.g., Van Dessel, De Houwer, Gast, et al., 2016). It also implies, however, that the impact of actual contingencies or contingency instructions on liking can depend on information about the diagnosticity of those contingencies or instructions. This prediction was tested in a number of studies.

In a study by Siegel, Sigall, and Huber (2012), participants were exposed to repeated pairings of names of two fictitious social groups with either positive or negative adjectives. Participants exhibited an explicit preference for the group paired with positive adjectives when instructions indicated that the adjectives accurately described the groups but not when the pairings were described as random. Zanon, De Houwer, Gast, and Smith (2014) extended these findings by showing that the perceived validity of instructions about the meaning of pairings moderate instruction effects on both implicit and explicit evaluation. Participants first received instructions that they would see pairings between specific non-word CSs and valenced USs and that these pairings implied that the CSs either had the same or the opposite semantic meaning than the USs. After the pairings, participants were informed that the semantic meaning information that they had received was valid or invalid and this moderated effects on implicit and explicit evaluations. Interestingly, a group of participants who did not receive any instructions about the meaning of the pairings exhibited EC effects similar in size

to effects of participants who were informed that the words were similar in meaning, suggesting that participants by default interpret pairings as indicating an equivalence relation between the paired components.

In a second line of research, the diagnosticity of pairings was not manipulated via instructions but by varying specific characteristics of the paired components. Fan and Bodenhausen (2017) presented pairings of unfamiliar foods and liquors (CSs) with USs that either were or were not plausible consequences of the USs in real life (e.g., foods paired with obesity versus car accident pictures). EC effects on implicit evaluations as measured with the IAT and AMP were strongly reduced for the implausible compared to the plausible pairings. A recent study corroborated these findings, showing stronger EC effects on implicit and explicit evaluations of target persons when they were paired with trait adjectives that were rated as highly diagnostic for revealing the true nature of a person than when they were paired with nouns that were rated as similar in valence but lower in diagnosticity (Moran, Hughes, Van Dessel, & De Houwer, 2019).

In another recent study, we argued that inferences about the consequences of approach-avoidance responses might strongly moderate approach-avoidance training effects. Based on this idea, we developed a novel task in which such inferences were reinforced (Van Dessel, Hughes, & De Houwer, 2018). More specifically, participants used an avatar to approach or avoid healthy and unhealthy foods and experienced plausible consequences of these actions (e.g., approaching unhealthy foods reduced the health of the avatar). In contrast to participants who completed a typical approach-avoidance task in which they always approached healthy foods (and did not contact any action consequences), participants who completed the consequence-based approach-avoidance training exhibited more positive implicit and explicit evaluations of healthy compared to unhealthy foods and reported eating less unhealthily in the days after the intervention. These preliminary results point to the

potential of (pairing-based) interventions that draw on propositional theories and target specific inferences that might underlie unwanted (evaluative) behavior. The investigation of effects of such interventions represents an important area of future research.

On the (Non-)Automaticity of EC

The issues discussed in the previous three sections were for a large part put on the research agenda because of the introduction of a propositional perspective on EC and implicit evaluation. The question of whether EC occurs under conditions of automaticity, on the other hand, was already a major research theme before propositional models of EC were put forward. Initially, research on this question was inspired predominantly by a simple associative perspective according to which (a) the co-occurrence of stimuli is a necessary and sufficient condition for the formation of associations in memory and (b) associations in memory are necessary and sufficient for EC (e.g., Baeyens, Eelen, Crombez, & Van den Bergh, 1992). This perspective links claims about the operating principles of the processes underlying EC (i.e., the type of representations that are formed) to claims about the operating conditions under which these processes are thought to occur (i.e., the conditions under which those representations are formed; Gawronski & Bodenhausen, 2014). Based on these ideas, researchers predicted that EC should occur even when participants are not aware of the CS-US contingency and when they need to invest effort in other concurrent tasks. Ironically, such a simple associative perspective had already been largely abandoned by learning psychologists in the 1970s and 1980s because of evidence showing that conditioning in humans often if not always requires awareness of the CS-US contingency (e.g., Brewer, 1974; Dawson & Schell, 1985) and the availability of cognitive resources (e.g., Carter, Hofstotter, Tsuchiya, & Koch, 2003). Nevertheless, driven by the idea that EC might be more “basic” than other types of conditioning (e.g., Baeyens et al., 1992), as well as by initial results that seemed to reveal EC in the absence of awareness (e.g., Baeyens, Eelen & Van den

Bergh, 1990) and cognitive resources (e.g., Fulcher & Hammerl, 2001), the simple associative perspective on EC remained quite popular for a long time (e.g., Hütter & Sweldens, 2018; Walther, Nagengast, & Trasselli, 2005).

Because the origins of research on the automaticity of EC lie primarily in associative models (e.g., Baeyens et al., 1992) and work on automaticity in general (e.g., Bargh, 1992) and because research on the automaticity of EC was recently reviewed in detail by Corneille and Stahl (2019), we will not devote too much of the limited space to this topic. Nevertheless, in our opinion, the propositional perspective on EC did have a significant impact on research about the automatic nature of EC. Most importantly, by drawing the parallel between EC and problem solving, it gave researchers a reason to seriously consider the possibility that EC requires contingency awareness and cognitive resources and thus to critically evaluate the available evidence of EC without contingency awareness or cognitive resources.

It is interesting to see that on the basis of the currently available evidence, Corneille and Stahl (2019) concluded that there is little, if any, evidence for EC without contingency awareness or cognitive resources. More generally, they found little support for the idea of an automatic association formation mechanism. Instead, they concluded that out of all the available accounts, the available evidence fits best with the idea that EC and other forms of evaluative learning are mediated exclusively by the formation of propositions. Within the context of the present chapter, however, the main take home message from the evidence reviewed by Corneille and Stahl (2019) is that the propositional perspective had an important impact on research about the automatic nature of EC by providing a clear challenger for a simple associative perspective, as well as for more complex models that acknowledge a role for simple association formation processes in addition to the formation of propositions (e.g., Rydell & McConnell, 2006).⁵

⁵ The associative perspective also instigated research on implicit evaluation, more specifically the nature of the

Summary and Conclusions

We started the chapter by acknowledging the merits of the idea that attitudes are associations in memory. Its success did not result from the introduction of highly formalized associative models that made exact predictions but primarily from providing a link between attitude research and the rich philosophical and psychological literature on associationistic approaches to human behavior and thinking. Most importantly, the associative perspective on attitude research led to an extensive line of research on implicit evaluation and EC.

In this chapter, we argued that there is also merit in a propositional perspective on attitude research. Whereas the importance of propositions is self-evident in certain areas of attitude research such as persuasion, researchers also started to explore the possibility that seemingly association-based phenomena such as implicit evaluation and EC are mediated by propositions. We believe that, in this respect, the propositional perspective has been highly generative. As was the case with the associative perspective, much of this impact was realized despite the absence of precise or formalized models. Instead, it arose from linking attitude research with knowledge about propositional processing from philosophy (e.g., about the nature of propositions) and psychology (e.g., about the formation of propositions via problem solving).

As we noted from the outset, we did not have the aim to empirically differentiate between the associative and propositional perspectives. Because both perspectives encompass a wide range of possible models, it is probably impossible to refute one or the other perspective (Miller & Escobar, 2001). In fact, it might well be that most if not all of the findings that were inspired by the propositional perspective (e.g., the impact of relational information, instructions, and truth information) can somehow be accounted for by some

suboptimal conditions under which stimulus evaluation can take place (see Fazio, 2001, for an early review). A propositional perspective also allows for evaluation to occur under suboptimal conditions (e.g., as the result of automatic inferences or automatic, similarity-based activation of propositions from memory). Hence, the propositional perspective had little impact on research about the automaticity features of implicit evaluation.

version of an associative model of implicit evaluation or EC. It might not even be possible to exclude more specific models within each perspective. At present, there are few models of implicit evaluation or EC that are sufficiently specified to be falsified. Even those models that are sufficiently precise (e.g., Gawronski & Bodenhausen, 2006) could in principle be adjusted in a post-hoc manner to accommodate disconfirmatory results.

The inability to confirm or disconfirm (classes of) models about the mental processes and representations that mediate stimulus evaluation might be worrisome to some. One should realize, however, that this is a general problem in (cognitive) science but fortunately one that does not exclude scientific progress. Although there is certainly merit in trying to formulate precise theoretical models, it is generally recognized that falsification of theoretical models has limitations as a scientific strategy (Lakatos, 1974). First, most scientific models allow for a protective belt of auxiliary assumptions that allow them to accommodate initially unpredicted findings, thus complicating true falsification. Second, even the failure to falsify a model does not imply that it is correct. These problems are exacerbated when studying mental (i.e., information) processing in general (because information is non-physical and can thus be studied only via self-reports or observable behavior) and mental representations specifically (because their impact on self-reports and behavior depends on assumptions about the processes that operate on these representations; see De Houwer et al., 2013, for more details). Hence, we should not be surprised that it is difficult to reach consensus about the nature of the mental representations that mediate stimulus evaluations.

On the other hand, as we already noted at the start of this chapter, entertaining ideas about the nature of attitudinal representations can still be useful for attitude research. Regardless of whether these ideas are best described as a general perspective or a specific model, they can be useful tools because of their heuristic and predictive value, that is, because of their ability to make known findings intelligible (i.e., heuristic value) and to

predict novel findings (i.e., predictive value). As evidenced by the studies reviewed in this chapter, also the propositional perspective has been useful in this respect. Merely considering the possibility that attitudinal phenomena are mediated by propositional representations encourages researchers to (a) look for alternative explanations for phenomena that at first sight seem to fit better with other perspectives (e.g., as was the case with propositional accounts of implicit evaluation and EC) and (b) reexamine those phenomena more critically (e.g., as was the case with evidence for EC in the absence of contingency awareness or cognitive resources).

Readers who now see merit in considering a *possible* role for propositions might still have serious doubts about whether the evidence will end up supporting the idea that *all* attitudinal phenomena are mediated by propositions. On the one hand, from a purely logical point of view, it is necessarily true that any given set of findings can be explained at least as well by allowing for multiple types of representations than by considering only one of those types of representations. On the other hand, allowing for multiple representations also complicates matters, for instance, with regard to the question of when which type of representations influences behavior. Hence, one should not abandon too quickly the idea that propositions underlie all attitudinal phenomena (see De Houwer, 2014b, for a more detailed discussion of why strict propositional models deserve to be defended).

Of course, one should be willing to let go of a specific model or even a broader perspective if its heuristic and predictive value are not satisfactory. Based on an extensive review of the literature on EC and other forms of evaluative learning, Corneille and Stahl (2019) recently concluded that the propositional perspective fits well with the available evidence. Hence, at least in that area of research, the heuristic value of a propositional perspective on EC seems to be high. Also the studies on implicit evaluation that we reviewed in this chapter are largely in line with a propositional perspective. Hence, we see little reason

for abandoning a propositional perspective on EC and implicit evaluation.

Until now, we have argued that there is merit in adopting a propositional perspective on EC and implicit evaluation mainly because it provides us with new information about the moderators of those phenomena (i.e., its predictive value). But what do the studies that were inspired by a propositional perspective actually tell us about the moderators of EC and implicit evaluation? In many respects, the evidence that emerged as a result of these studies suggests that EC and implicit evaluation are not as special as we initially thought they were. EC seems to have a lot in common with persuasion (such as the impact of relational information, contingency instructions and persuasive messages, truth information, awareness, and mental resources). Implicit evaluation seems to have many things in common with explicit evaluation (such as the impact of relational information, persuasive messages, and truth information). In sum, research that was inspired by a propositional perspective has certainly complicated the picture in research on attitude acquisition and attitude activation.

One response to this complexity at the empirical level is to argue that many studies failed to capture “true” instances of EC and implicit evaluation that do have fundamentally different properties than persuasion and explicit evaluation, respectively. One could indeed argue that some results were “contaminated” by propositional processes such as demand effects which trumped the associative processes. Such arguments are difficult to refute. We can only point out that the impact of factors such as relational information, instructions, and truth information on EC and implicit evaluation has been observed across a wide range of situations using a wide range of indices of liking. Even parameters of complex modeling techniques that were specifically designed to capture “truly” associative processes have proven sensitive to these factors. Researchers who still wish to maintain the position that there are instances of EC and implicit evaluation that are not sensitive to factors such as relational information, instructions, and truth information need to clearly specify the

conditions under which those instances can be observed and substantiate their claims empirically.

Another response to complexity is to embrace it. In our opinion, the fact that EC and implicit evaluation can be moderated by factors such as relational information, instructions, and truth information, highlights new opportunities for putting these phenomena to practical use. Let us start with EC. Even though EC does not seem to be the “primitive means of changing attitudes” that it is often thought to be (Briñol, Petty, & McCaslin, 2009, p. 287), it continues to provide a unique avenue for shaping stimulus evaluation. The evidence that we reviewed in this chapter supports the idea that EC procedures and persuasive messages are both sources of information for deciding whether a stimulus is good or bad. However, the nature of the information provided is fundamentally different. Whereas persuasive messages provide ready-made information about the evaluative properties of stimuli, EC procedures merely provide spatio-temporal information. From this spatio-temporal information, propositions about the relation between stimuli can be derived which in turn can lead to inferences about the evaluative properties of stimuli. On the one hand, the information provided by EC procedures is much more indirect (i.e., requires more additional processing steps before a conclusion can be drawn) than the information provided by persuasive messages. This could explain why persuasive messages are much more impactful than EC procedures in determining stimulus evaluations (e.g., Van Dessel, De Houwer, Gast, et al., 2016). On the other hand, whereas persuasive messages are typically generated by others, inferences about evaluative stimulus properties that are derived from spatio-temporal information are self-generated. As we noted in the context of the recent studies of Kurdi and Banaji (2019), self-generated propositions are easier to remember and might therefore have more long-lasting effects than merely instructed propositions. Moreover, persuasive messages often lead to reactance, which can be sidestepped by using EC procedures as a source of

information. Finally, knowledge about the different inferential steps that are involved in EC effects can lead to new ways of strengthening these inferences and thus EC effects (see Van Dessel, Hughes, & De Houwer, 2018, for an example of this approach). In sum, also from a propositional perspective, EC research can continue to provide unique insights in how to change stimulus evaluations (see De Houwer & Hughes, 2016, for an in-depth discussion).

What about implicit evaluations? If implicit evaluations are sensitive to many of the same factors as explicit evaluations, what is the use of acknowledging the existence of and developing tasks for capturing implicit evaluations? We see implicit evaluations as evaluative responses that are evoked by stimuli under suboptimal conditions, for instance, when there is little time to process the stimulus or when there are other tasks that need to be performed (see De Houwer, Gawronski, & Barnes-Holmes, 2013, for more details). Explicit evaluations, on the other hand, are evaluative responses to stimuli that are evoked under optimal conditions, such as when there is ample time to process the stimulus and no other tasks to be fulfilled. Although we realize that it is difficult to draw an exact line between conditions that qualify as optimal versus suboptimal, the term “implicit” remains useful as an umbrella term that captures a variety of conditions that are in some way suboptimal for cognitive processing (Moors, 2016; Moors & De Houwer, 2006).

Although, from a propositional perspective, all instances of stimulus evaluation are mediated by propositions, dissociations between implicit and explicit evaluation can result from differences in the retrieval of propositions from memory. Which propositions determine the evaluation of stimuli is likely to depend on the conditions under which stimulus evaluation takes place, including whether those conditions are in some way suboptimal. As noted by Gawronski and Bodenhausen (2006, 2018), many of the dissociations that have been observed between implicit and explicit evaluations might reflect the fact that tasks designed to capture implicit evaluations provide little opportunity for the integration of different

sources of information. Likewise, as suggested by the data of Van Dessel et al. (2016, 2017), participants might be less able to control which propositions influence their evaluative responses in tasks designed to capture implicit evaluations than in tasks designed to reflect explicit evaluations. Because a propositional perspective does not deny the possibility of dissociations between implicit and explicit evaluations, it continues to underline the importance of research on implicit evaluations.

Finally, we believe that the propositional perspective can also inspire research on phenomena other than EC and implicit evaluation. To give one example, consider research on implicit stereotyping, that is, automatic stereotypical thoughts and behavior. Research on this topic is currently dominated by an associative perspective (e.g., Devine, 1989; Greenwald et al., 2002; Kawakami, Amodio, & Hugenberg, 2017; Kunda & Spencer 2003) even though stereotypical thoughts typically involve relational beliefs about the characteristics of members of certain groups (e.g., Hilton & Von Hippel, 1996). The relational nature of these stereotypical beliefs can be illustrated by the distinction between descriptive stereotypes (e.g., *women are emotional*) and prescriptive stereotypes (e.g., *woman should be emotional*; e.g., Burgess & Borgida, 1999). A propositional perspective would highlight the possibility that implicit stereotypes rely on the automatic activation of propositions and thereby instigate new research on the role of relational information, instructions, and truth validation. Such potential moderators are currently ignored because of the dominant associative perspective on implicit stereotyping.

Regardless of these possible future developments, we believe that a propositional perspective on EC and implicit evaluation has already been highly generative. It is, of course, always difficult to know what the world would have looked like if past events had unfolded in a different way. However, we are convinced that the propositional perspective on EC and implicit evaluation changed attitude research for the better. By highlighting this contribution

in the present chapter, we hope that more researchers will start exploring the merits of the propositional perspective.

References

- Allport, G. (1935). Attitudes. In Murchison, C. (Ed.). *A Handbook of Social Psychology* (pp. 789-843). Worcester, MA: Clark University Press.
- Anderson, J. R., & Bower, G. H. (1973). *Human associative memory*. Washington DC: Winston.
- Bading, K., Stahl, C., & Rothermund, K. (2019). Why a standard IAT effect cannot provide evidence for association formation: the role of similarity construction. *Cognition and Emotion*. Advance online publication. doi:10.1080/02699931.2019.1604322
- Baeyens, F., Eelen, P., Crombez, G., & Van den Bergh, O. (1992). Human evaluative conditioning: Acquisition trials, presentation schedule, evaluative style and contingency awareness. *Behaviour Research and Therapy*, *30*, 133–142
- Baeyens, F., Eelen, P., & Van den Bergh, O. (1990). Contingency awareness in evaluative conditioning: A case for unaware affective–evaluative learning. *Cognition & Emotion*, *4*, 3–18.
- Bargh, J. A. (1992). The ecology of automaticity: Toward establishing the conditions needed to produce automatic processing effects. *The American Journal of Psychology*, *105*, 181-199.
- Bar-Anan, Y., De Houwer, J., & Nosek, B. (2010). Evaluative conditioning and conscious knowledge of contingencies: A correlational investigation with large samples. *Quarterly Journal of Experimental Psychology*, *63*, 2313-2335.
- Bar-Anan, Y., & Moran, T. (2018). Simple first: A skeleton for an evaluative learning model. *Social Psychological Bulletin*, *13*, e28761.
- Bar-Anan, Y., & Nosek, B. A. (2012). Reporting intentional rating of the primes predicts priming effects in the affective misattribution procedure. *Personality and Social*

Psychology Bulletin, 38, 1194-1208.

- Bechtel, W. (2008). *Mental mechanisms: Philosophical perspectives on cognitive neuroscience*. London: Routledge.
- Benedict, T., Richter, J., & Gast, A. (2019). The influence of misinformation manipulations on evaluative conditioning. *Acta Psychologica*, 194, 28-36.
- Boucher, K. L., & Rydell, R. J. (2012). Impact of negation salience and cognitive resources on negation during attitude formation. *Personality and Social Psychology Bulletin*, 38, 1329-1342.
- Brannon, S. M., & Gawronski, B. (2017). A second chance for first impressions? Exploring the context (in)dependent updating of implicit evaluations. *Social Psychological and Personality Science*, 3, 275-283.
- Brewer, W. F. (1974). There is no convincing evidence for operant or classical conditioning in adult humans (pp. 1-42). In W. B. Weimer & D. S. Palermo (Eds.), *Cognition and the symbolic processes*. Oxford, England: Lawrence Erlbaum.
- Briñol, P., Petty, R. E., & McCaslin, M. J. (2009). Changing attitudes on implicit versus explicit measures: What is the difference? In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 285–326). New York: Psychology Press.
- Burgess, D., & Borgida, E. (1999). Who women are, who women should be: Descriptive and prescriptive gender stereotyping in sex discrimination. *Psychology, Public Policy, and Law*, 5, 665-692.
- Calanchini, J., & Cone, J. (2019). *A Quad modeling analysis of effects of information diagnosticity in the revision of implicit evaluations*. Manuscript in preparation.
- Carter, R. M., Hofstotter, C., Tsuchiya, N., & Koch, C. (2003). Working memory and fear conditioning. *Proceedings of the National Academy of Sciences of the United States of*

America, 100, 1399–1404.

- Cone, J., & Ferguson, M. J. (2015). He Did what? The Role of Diagnosticity in Revising Implicit evaluations. *Journal of Personality and Social Psychology, 108*, 37-57.
- Cone, J., Flaharty, K., & Ferguson, M. J. (in press). Believability of evidence matters for correcting social impressions. *Proceedings of the National Academy of Sciences of the United States of America*.
- Cone, J., Mann, T. C., & Ferguson, M. J. (2017). Changing our implicit minds: How, when, and why implicit evaluations can be rapidly revised. *Advances in Experimental Social Psychology, 56*, 131-199.
- Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. (2005). Separating multiple processes in implicit social cognition: The Quad-Model of implicit task performance. *Journal of Personality and Social Psychology, 89*, 469-487.
- Corneille, O., & Stahl, C. (2019). Associative attitude learning: A closer look at evidence and how it relates to attitude models. *Personality and Social Psychology Review, 23*, 161-189
- Dawson, M.E., & Schell, A.M. (1985). Information processing and human autonomic classical conditioning. *Advances in Psychophysiology, 1*, 89-165
- Dewitte, M., & De Houwer, J. (2008). Proximity and distance motives in adult attachment. *European Journal of Personality, 22*, 675–694.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*, 5-18.
- De Houwer, J. (2006). Using the Implicit Association Test does not rule out an impact of conscious propositional knowledge on evaluative conditioning. *Learning and Motivation, 37*, 176–187.

- De Houwer, J. (2007). A conceptual and theoretical analysis of evaluative conditioning. *The Spanish Journal of Psychology, 10*, 230-241.
- De Houwer, J. (2009a). How do people evaluate objects? A brief review. *Social and Personality Psychology Compass, 3*, 36-48.
- De Houwer, J. (2009b). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior, 37*, 1-20.
- De Houwer, J. (2011). Why the cognitive approach in psychology would profit from a functional approach and vice versa. *Perspectives on Psychological Science, 6*, 202-209.
- De Houwer, J. (2014a). A Propositional Model of Implicit Evaluation. *Social and Personality Psychology Compass, 8*, 342-353.
- De Houwer, J. (2014b). Why a propositional single-process model of associative learning deserves to be defended. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual processes theories of the social mind* (pp. 530-541). New York: Guilford.
- De Houwer, J. (2018). Propositional models of evaluative conditioning. *Social Psychological Bulletin, 13*(3), Article e28046. doi:10.5964/spb.v13i3.28046
- De Houwer, J., Baeyens, F., & Field, A. (2005). Associative learning of likes and dislikes: Some current controversies and possible ways forward. *Cognition & Emotion, 19*, 161-174.
- De Houwer, J., Barnes-Holmes, D., & Moors, A. (2013). What is learning? On the nature and merits of a functional definition of learning. *Psychonomic Bulletin & Review, 20*, 631-642.
- De Houwer, J., Gawronski, B., & Barnes-Holmes, D. (2013). A functional-cognitive framework for attitude research. *European Review of Social Psychology, 24*, 252-287.
- De Houwer, J., Heider, N., Spruyt, A., Roets, A., & Hughes, S. (2015). The Relational

- Responding Task: Toward a new implicit measure of beliefs. *Frontiers in Psychology*. 6:319. DOI: 10.3389/fpsyg.2015.00319
- De Houwer, J., & Hughes, S. (2016). Evaluative conditioning as a symbolic phenomenon: On the relation between evaluative conditioning, evaluative conditioning via instructions, and persuasion. *Social Cognition*, 34, 480-494.
- De Houwer, J., Mattavelli, S., & Van Dessel, P. (2019). Dissociations between learning phenomena do not necessitate multiple learning processes: Mere instructions about upcoming stimulus presentations differentially influence liking and expectancy. *Journal of Cognition*, 2, 7.
- De Houwer, J., Thomas, S., & Baeyens, F. (2001). Association learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin*, 127, 853.-869.
- De Raedt, R., Schacht, R., Franck, E., & De Houwer, J. (2006). Self-esteem and depression revisited: Implicit positive self-esteem in depressed patients? *Behaviour Research and Therapy*, 44, 1017–1028.
- Fan, X., & Bodenhausen, G. (2017). Changing Brand Attitudes Using Relevant Affective Cues: an Investigation of Relational Processing in Evaluative Conditioning. In Gneezy, A. et al. (Eds.), *Advances in Consumer Research*: Vol. 45 (pp. 595-595). Duluth, MN: Association for Consumer Research.
- Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion*, 15, 115—141.
- Fazio, R. H. (2007). Attitudes as object-evaluation associations of varying strength. *Social Cognition*, 25, 603-637.
- Fazio, R. H. , Chen, J. , McDonel, E. C. , & Sherman, S. J. (1982). Attitude accessibility, attitude-behavior consistency, and the strength of the object-evaluation association.

Journal of Experimental Social Psychology, 18, 339-357.

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, 1013-1027.

Fazio, R.H., Sanbonmatsu, D.M., Powell, M.C., & Kardes, F.R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, 50, 229-238.

Ferguson, M., & Zayas, V. (2009). Automatic evaluation. *Current Directions in Psychological Science*, 18, 362–366.

Fiedler, K., & Unkelbach, C. (2011). Evaluative conditioning depends on higher order encoding processes. *Cognition and Emotion*, 25, 639-656.

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

Franck, E., De Raedt, R., Dereu, M., & Van den Abbeele, D. (2007). Implicit and explicit self-esteem in currently depressed individuals with and without suicidal ideation. *Journal of Behavior Therapy and Experimental Psychiatry*, 38, 75–85.

Fulcher, E. P., & Hammerl, M. (2001). When all is revealed: A dissociation between evaluative learning and contingency awareness. *Consciousness and Cognition*, 10, 524–549.

Gast, A. (2018). A declarative memory model of evaluative conditioning. *Social Psychological Bulletin*, 13(3), e28590.

Gast, A., & De Houwer, J. (2013). The influence of extinction and counterconditioning instructions on evaluative conditioning effects. *Learning and motivation*, 44, 312-325.

Gast, A., & Kattner, F. (2016). Single-trial evaluative conditioning can be moderated by instructed forgetting. *Learning & Behavior*, 44, 260–269.

- 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. (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: Guilford Press.
- Gawronski, B., & Bodenhausen, G. V. (2018). Evaluative conditioning from the perspective of the associative-propositional evaluation model. *Social Psychological Bulletin, 13*(3), e28024.
- Gawronski, B., & Brannon, S. M. (2019). Attitudes and the implicit-explicit dualism. In D. Albarracín & B. T. Johnson (Eds.), *The handbook of Attitudes. Volume 1: Basic Principles* (2nd edition, pp. 158-196). New York, NY: Routledge.
- 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., & Hahn, A. (2019). Implicit measures: Procedures, use, and interpretation. In H. Blanton, J. M. LaCroix, & G. D. Webster (Eds.), *Measurement in social psychology* (pp. 29-55). New York, NY: Taylor & Francis.
- Gawronski, B., Rydell, R. J., De Houwer, J., Brannon, S. M., Ye, Y., Vervliet, B., & Hu, X. (2018). Contextualized Attitude Change. *Advances in Experimental Social Psychology, 57*, 1-52.
- 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, NY: Guilford Press.

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*, 618-626.

Gentner, D. (2016). Language as cognitive toolkit: How language supports relational thought. *American Psychologist, 71*, 650-657.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*, 3-25.

Greenwald, A. G., & De Houwer, J. (2017). Unconscious conditioning: Demonstration of existence and difference from conscious conditioning. *Journal of Experimental Psychology: General, 146*, 1705-1721.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*, 1464-1480.

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.

Hastie, R. (1988). A computer simulation model of person memory. *Journal of Experimental Social Psychology, 24*, 423-447.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (Eds.). (2001). *Relational Frame Theory: A Post-Skinnerian account of human language and cognition*. New York: Plenum Press.

Hebb, D.O. (1949). *Organization of behavior: A neuropsychological theory*. New York: John Wiley & Sons.

- Heider, N., Spruyt, A., & De Houwer, J. (2018). Body dissatisfaction revisited: On the importance of implicit beliefs about actual and ideal body image. *Psychologica Belgica*, 58, 158-173. DOI:<https://doi.org/10.5334/pb.362>
- Hermans, D., Vansteenwegen, D., Crombez, G., Baeyens, F., & Eelen, P. (2002). Expectancy- learning and evaluative learning in human classical conditioning: Affective priming as an indirect and unobtrusive measure of conditioned stimulus valence. *Behaviour Research and Therapy*, 40, 217-234.
- Heycke, T., & Gawronski, B. (2019). Co-occurrence and relational information in evaluative learning: A multinomial modeling approach. *Journal of Experimental Psychology: General*. Advance online publication. doi:10.1037/xge0000620
- Heycke, T., Gehrman, S., Haaf, J. M., & Stahl, C. (2018). Of two minds or one? A registered replication of Rydell et al. (2006). *Cognition and Emotion*, 32, 1708-1727.
- Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. *Advances in Experimental Social Psychology*, 30, 1-46.
- Hilton, J. L., & Von Hippel, W. (1996). Stereotypes. *Annual review of psychology*, 47, 237-271.
- Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: a meta-analysis. *Psychological Bulletin*, 136, 390–421.
- Hogarth, L. (2018). A critical review of habit theory of drug dependence. In B. Verplanken (Ed.), *The Psychology of Habit* (pp. 325-341). Springer, Cham.
- 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, 17-32.
- 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 and Personality Science*, 8, 858-866.

Hughes, S., & Barnes-Holmes, D. (2014). Associative concept learning, stimulus equivalence, and relational frame theory: Working out the similarities and differences between human and non-human behavior. *Journal of the Experimental Analysis of Behavior*, 101, 156-160.

Hughes, S., & Barnes-Holmes, D. (2016). Relational Frame Theory: The basic account. In S. Hayes, D. Barnes-Holmes, R. Zettle, and T. Biglan (Eds.), *Handbook of contextual behavioral science* (pp.129-178). New York: New Harbinger.

Hughes, S., Barnes-Holmes, D., & Vahey, N. (2012). Holding on to our functional roots when exploring new intellectual islands: A voyage through implicit cognition. *Journal of Contextual Behavioral Science*, 1, 17-38.

Hughes, S., Barnes-Holmes, D., Van Dessel, P., de Almeida, J. H., Stewart, I., & De Houwer, J. (2018). On the symbolic generalization of likes and dislikes. *Journal of Experimental Social Psychology*, 79, 365-377.

Hughes, S., De Houwer, J., & Barnes-Holmes, D. (2016). The moderating impact of distal regularities on the effect of stimulus pairings: A novel perspective on evaluative conditioning. *Experimental Psychology*, 63, 20-44.

Hughes, S., De Houwer, J., & Perugini, M. (2016). The functional-cognitive framework for psychological research: Controversies and resolutions. *International Journal of Psychology*, 51, 4-14.

Hughes, S., Ye, Y., & De Houwer, J. (2018). Evaluative conditioning effects are modulated by the nature of contextual pairings. *Cognition and Emotion*, 1-14.

Hughes, S., Ye, Y., Van Dessel, P., & De Houwer, J. (2019). When people co-occur with

good or bad events: Graded effects of relational qualifiers on evaluative conditioning. *Personality and Social Psychology Bulletin*, 45, 196-208.

Hummel, J. E. (2010). Symbolic vs. associative learning. *Cognitive Science*, 34, 958-965.

Hütter, M., & De Houwer, J. (2017). Examining the contributions of memory-dependent and memory-independent components to evaluative conditioning via instructions. *Journal of Experimental Social Psychology*, 71, 49-58.

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

Jozefowicz, J. (2018). Associative versus predictive processes in Pavlovian conditioning. *Behavioral Processes*, 154, 21–26.

Kawakami, K., Amodio, D. M., & Hugenberg, K. (2017). Intergroup perception and cognition: An integrative framework for understanding the causes and consequences of social categorization. *Advances in Experimental Social Psychology*, 55, 1-80.

Kawakami, K., Phills, C. E., Steele, J. R., & Dovidio, J. F. (2007). (Close) distance makes the heart grow fonder: Improving implicit racial evaluations and interracial interactions through approach behaviors. *Journal of Personality and Social Psychology*, 92, 957-971.

Kollias, P., & McClelland, J. L. (2013) Context, cortex, and associations: A connectionist developmental approach to verbal analogies. *Frontiers in Psychology*, 4:857.

Kukken, N., Hütter, M., & Holland, R. W. (2019). Are there two independent evaluative conditioning effects in relational paradigms? Dissociating the effects of CS-US pairings and their meaning. *Cognition & Emotion*. Advance online publication. doi: 10.1080/02699931.2019.1617112

Kunda, Z., & Spencer, S. J. (2003). When do stereotypes come to mind and when do they

- color judgment? A goal-based theoretical framework for stereotype activation and application. *Psychological Bulletin*, *129*, 522-544.
- 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*, 194–213.
- Kurdi, B., & Banaji, M. R. (2019). Attitude change via repeated evaluative pairings versus evaluative statements: Shared and unique features. *Journal of Personality and Social Psychology*, *116*, 681–703
- Lagnado, D. A., Waldmann, M. R., Hagmayer, Y., & Sloman, S. A. (2007). Beyond covariation: Cues to causal structure. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 154-172). New York, NY: Oxford University Press.
- Lakatos, I. (1974). The Role of Crucial Experiments in Science. *Studies in the History and Philosophy of Science*, *4*, 309-325.
- 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*, 1232-1237.
- Levey, A. B., & Martin, I. (1975). Classical conditioning of human 'evaluative' responses. *Behaviour Research and Therapy*, *4*, 205-207.
- Mandelbaum, E. (2016). Attitude, inference, association: On the propositional structure of implicit bias. *Noûs*, *50*, 629-658.
- Mann, T. C., & Ferguson, M. J. (2015). Can we undo our first impressions? The role of reinterpretation in reversing implicit evaluations. *Journal of Personality and Social Psychology*, *108*, 823-849.

- Mann, T. C., & Ferguson, M. J. (2017). Reversing implicit first impressions through reinterpretation after a two-day delay. *Journal of Experimental Social Psychology, 68*, 122–127.
- Melnikoff, D.E., & Bargh, J.A. (2018). The Mythical Number Two. *Trends in Cognitive Sciences, 22*, 280-293.
- Mertens, G., Van Dessel, P., & De Houwer, J. (2018). The contextual malleability of approach-avoidance training effects: Approaching or avoiding fear conditioned stimuli modulates effects of approach-avoidance training. *Cognition & Emotion, 32*, 341-349.
- Meyer, D.E., & Schvaneveldt, R.W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227-234.
- Miller, R.R., & Escobar, M. (2001). Contrasting acquisition-focused and performance-focused models of acquired behavior. *Current Directions in Psychological Science, 10*, 141–145.
- Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences, 32*, 183-198.
- Moors, A. (2016). Automaticity: Componential, causal, and mechanistic explanations. *Annual Review of Psychology, 67*, 263-287.
- Moors, A., & De Houwer, J. (2006). Automaticity: A conceptual and theoretical analysis. *Psychological Bulletin, 132*, 297-326.
- Moran, T., & Bar-Anan, Y. (2013). The effect of object–valence relations on automatic evaluation. *Cognition & Emotion, 27*, 743-752.
- Moran, T., & Bar-Anan, Y. (2019). The effect of co-occurrence and relational information on speeded evaluation. *Cognition and Emotion*. Advance online publication.

doi:10.1080/02699931.2019.1604321

- 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.
- Moran, T., Bar-Anan, Y., & Nosek, B. A. (2016). The assimilative effect of co-occurrence on evaluation above and beyond the effect of relational qualifiers. *Social Cognition, 34*, 343-356.
- Moran, T., Bar-Anan, Y., & Nosek, B. A. (2017). The effect of the validity of co-occurrence on automatic and deliberate evaluation. *European Journal of Social Psychology, 47*, 708-723.
- Moran, T., Hughes, S., Van Dessel, P., & De Houwer, J. (2019). *The role of diagnosticity in evaluative conditioning*. Manuscript in preparation. OSF project available at <https://osf.io/6gp2h/>
- Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science, 12*, 413-417.
- Olson, M. A., & Fazio, R. H. (2004). Reducing the influence of extra-personal associations on the Implicit Association Test: Personalizing the IAT. *Journal of Personality and Social Psychology, 86*, 653-667.
- Payne, B. K., Burkley, M. A., & Stokes, M. B. (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology, 94*, 16-31.
- Peters, K. R., & Gawronski, B. (2011). Are we puppets on a string? Comparing the impact of contingency and validity on implicit and explicit evaluations. *Personality and Social Psychology Bulletin, 37*, 557-569.

- Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology, 19*, 123-205.
- Petty, R. E., Tormala, Z. L., Briñol, P., & Jarvis, W. B. G. (2006). Implicit ambivalence from attitude change: An exploration of the PAST model. *Journal of Personality and Social Psychology, 90*, 21–41.
- Phillips, W. J., Hine, D. W., & Thorsteinsson, E. B. (2010). Implicit cognition and depression: A meta-analysis. *Clinical Psychology Review, 30*, 691-709.
- Rothermund, K., Teige-Mocigemba, S., Gast, A., & Wentura, D. (2009). Minimizing the influence of recoding in the Implicit Association Test: The Recoding-Free Implicit Association Test (IAT-RF). *The Quarterly Journal of Experimental Psychology, 62*, 84–98.
- 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*, 995–1008. doi: 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*, 954–958.
- Schmidt, J. R., De Houwer, J., & Rothermund, K. (2016). The Parallel Episodic Processing (PEP) model 2.0: A single computational model of stimulus-response binding, contingency learning, power curves, and mixing costs. *Cognitive Psychology, 91*, 82-108.
- Shanks, D. R. (2010). Learning: From association to cognition. *Annual Review of Psychology, 61*, 273–301.
- Siegel, E., Sigall, H., & Huber, D. E. (2012). The IAT is sensitive to the perceived accuracy of newly learned associations. *European Journal of Social Psychology, 42*, 189-199.

- Smith, C. T., Calanchini, J., Hughes, S., Van Dessel, P., & De Houwer, J. (2019). The impact of instruction and experience-based evaluative learning on IAT performance: A quad model perspective. *Cognition & Emotion*. Advance online publication. doi:10.1080/02699931.2019.1592118
- Smith, C. T., De Houwer, J., & Nosek, B. (2013). Consider the source: persuasion of implicit evaluations is moderated by source credibility. *Personality and Social Psychology Bulletin*, 39, 193-205.
- Spruyt, A., De Houwer, J., Tibboel, H., Verschuere, B., Crombez, G., Verbanck, P., Hanak, C., Brevers, D., & Noël, X. (2013). On the predictive validity of automatically activated approach/avoidance tendencies in abstaining alcohol-dependent patients. *Drug and Alcohol Dependence*, 127, 81-86.
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, 8, 220-247.
- Unkelbach, C., & Fiedler, K. (2016). Contrastive CS-US relations reverse evaluative conditioning effects. *Social Cognition*, 34, 413-434.
- Van Dessel, P., Cone, J., Gast, A., & De Houwer, J. (2019). The impact of valenced verbal information on implicit and explicit evaluation: The role of information diagnosticity, primacy, and memory cueing. *Cognition & Emotion*. Advance online publication. doi:10.1080/02699931.2019.1594703
- Van Dessel, P., & De Houwer, J. (2019). *Hypnotic suggestions can induce rapid change in implicit attitudes*. Manuscript under review. OSF project available at <https://osf.io/wkjp6/>
- Van Dessel, P., De Houwer, J., Gast, A., & Smith, C.T. (2015). Instruction-based approach-avoidance effects: Changing stimulus evaluation via the mere instruction to approach or avoid stimuli. *Experimental Psychology*, 62, 161-169.

- Van Dessel, P., De Houwer, J., & Gast, A. (2016). Approach-Avoidance Training effects are Moderated by Awareness of Stimulus-Action Contingencies. *Personality and Social Psychology Bulletin*, *42*, 81-93.
- Van Dessel, P., De Houwer, J., Gast, A., Roets, A., & Smith, C. T. (2019). *The effectiveness of approach-avoidance instructions and training for the reduction of prejudice*. Manuscript in preparation. OSF project available at <https://osf.io/wkjp6/>
- Van Dessel, P., De Houwer, J., Gast, A., Smith, C. T., & De Schryver, M. (2016). Instructing implicit processes: When instructions to approach or avoid influence implicit but not explicit evaluation. *Journal of Experimental Social Psychology*, *63*, 1-9.
- Van Dessel, P., De Houwer, J., & Smith, C. T. (2018). Relational information moderates approach-avoidance instruction effects on implicit evaluation. *Acta Psychologica*, *184*, 137-143.
- Van Dessel, P., Gawronski, B., & De Houwer, J. (2019). Does explaining social behavior require multiple memory systems? *Trends in Cognitive Science*, *23*, 368-369.
- Van Dessel, P., Gawronski, B., Smith, C. T., & De Houwer, J. (2017). Mechanisms underlying approach-avoidance instruction effects on implicit evaluation: Results of a preregistered adversarial collaboration. *Journal of Experimental Social Psychology*, *69*, 23-32.
- Van Dessel, P., Hughes, S., & De Houwer, J. (2018). Consequence-Based Approach-Avoidance Training: A New and Improved Method for Changing Unwanted Behavior. *Psychological Science*, *29*, 1899-1910.
- Van Dessel, P., Hughes, S., & De Houwer, J. (2019). How do actions influence attitudes? An inferential account of the impact of action performance on stimulus evaluation. *Personality and Social Psychology Review*. Advance online publication. doi:10.1177/1088868318795730

- Van Dessel, P., Hughes, S., Hussey, I., & De Houwer, J. (2019). *Inferential influences in evaluative learning effects*. Manuscript in preparation. OSF project available at <https://osf.io/pw7xa/>
- Van Dessel, P., Mertens, G., Smith, C. T., & De Houwer, J. (2017). The Mere Exposure Instruction Effect: Mere Exposure Instructions Influence Liking. *Experimental Psychology*, *64*, 299-314.
- Van Dessel, P., Ye, Y., & De Houwer, J. (2019). Changing deep-rooted implicit evaluation in the blink of an eye: Negative verbal information shifts automatic liking of Gandhi. *Social Psychological and Personality Science*, *10*, 266-273.
- Walther, E., Nagengast, B., & Trasselli, C. (2005). Evaluative conditioning in social psychology: Facts and speculations. *Cognition & Emotion*, *19*, 175-196. Wyer, N. A. (2016). Easier done than undone... by some of the people, some of the time: The role of elaboration in explicit and implicit group preferences. *Journal of Experimental Social Psychology*, *63*, 77-85.
- Zanon, R., De Houwer, J., & Gast, A. (2012). Context effects in evaluative conditioning of implicit evaluations. *Learning and Motivation*, *43*, 155-165.
- Zanon, R., De Houwer, J., Gast, A., & Smith, C. T. (2014). When does relational information influence evaluative conditioning? *Quarterly Journal of Experimental Psychology*, *67*, 2105-2122.
- Zentner, M., & Renaud, O. (2007). Origins of adolescents' ideal self: An intergenerational perspective. *Journal of Personality and Social Psychology*, *92*, 557-574.

Author Note

Jan De Houwer, Pieter Van Dessel, Tal Moran; Ghent University, Ghent, Belgium.

The preparation of this chapter was made possible by Methusalem Grant

BOF16/MET_V/002 of Ghent University. We thank Sean Hughes for comments on an earlier

draft of the chapter. Correspondence should be addressed to Jan De Houwer, Ghent

University, Henri Dunantlaan 2, B-9000 Ghent, Belgium. Electronic mail can be sent to

Jan.DeHouwer@UGent.be.