

A functional-cognitive framework for attitude research

Jan De Houwer¹, Bertram Gawronski²,
and Dermot Barnes-Holmes³

¹Department of Experimental Clinical and Health Psychology, Ghent University, Ghent, Belgium

²Department of Psychology, The University of Western Ontario, London, Canada

³Department of Psychology, National University of Ireland, Maynooth, Ireland

In attitude research, behaviours are often used as proxies for attitudes and attitudinal processes. This practice is problematic because it conflates the behaviours that need to be explained (*explanandum*) with the mental constructs that are used to explain these behaviours (*explanans*). In the current chapter we propose a meta-theoretical framework that resolves this problem by distinguishing between two levels of analysis. According to the proposed framework, attitude research can be conceptualised as the scientific study of evaluation. Evaluation is defined not in terms of mental constructs but in terms of elements in the environment, more specifically, as the effect of stimuli on evaluative responses. From this perspective, attitude research provides answers to two questions: (1) Which elements in the environment moderate evaluation? (2) What mental processes and representations mediate evaluation? Research on the first question provides explanations of evaluative responses in terms of elements in the environment (*functional level of analysis*); research on the second question offers explanations of evaluation in terms of mental processes and representations (*cognitive level of analysis*). These two levels of analysis are mutually supportive, in that better explanations at one level lead to better explanations at the other level. However, their mutually supportive relation requires a clear distinction between the concepts of their explanans and explanandum, which are conflated if behaviours are treated as proxies for mental constructs. The value of this functional-cognitive framework is illustrated by applying it to four central questions of attitude research.

Keywords: Attitudes; Cognitivism; Evaluation; Functional analysis; Levels of explanation.

Correspondence should be addressed to Jan De Houwer, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium. E-mail: Jan.DeHouwer@UGent.be

We thank Galen Bodenhausen, Christoph Klauer, and the members of the Learning and Implicit Processes lab at Ghent University for valuable comments on earlier versions of this chapter.

The preparation of this paper was made possible by Ghent University grants [grant number BOF/GOA2006/001] and [grant number BOF09/01M00209] to Jan De Houwer.

Attitudes are typically conceived of as mental entities that have a profound impact on behaviour. In social psychology the construct of *attitude* is commonly considered to be the “most distinctive and indispensable concept” (Allport, 1935, p. 798), because “understanding attitudes is the first step to understanding human behavior” (Conrey & Smith, 2007, p. 718). Attitude researchers therefore aim to develop and test theories about the mechanisms by which attitudes are formed and activated, the manner in which they are represented, and the processes by which they influence behaviour. Towards this end, researchers often treat behavioural effects as proxies for attitudes and attitudinal processes. For example, responses on an evaluative rating scale are frequently equated with the mental attitude that is assumed to underlie these responses (Krosnick, Judd, & Wittenbrink, 2005). Similarly, evaluative priming effects are often equated with the automatic activation of attitudes through processes of spreading of activation (e.g., Fazio, Sanbonmatsu, Powell, & Kardes, 1986).

Although the use of behavioural effects as proxies for attitudes and attitudinal processes is rather common, it is well known that this practice can be problematic for theory construction and the interpretation of empirical data (see Fazio, 2007; Krosnick et al., 2005). In the first part of this chapter we review these problems and argue that they can be traced back to confounds at the conceptual level. Specifically, we argue that using a behavioural effect as a proxy for a mental construct presupposes the validity of a priori assumptions about how the mental construct is related to the behavioural effect. If these assumptions turn out to be false, problems arise for theorising that is based on these assumptions. Yet, despite the widespread awareness of these issues, a closer look at the attitude literature reveals that it is still common practice to equate behavioural effects and mental constructs. We believe that one reason for this discrepancy is the absence of a general, phenomenon-independent framework that avoids such conceptual conflations.

The main goal of the current chapter is to resolve the problem of proxies in attitude research by proposing an overarching framework that (1) conceptually separates attitudinal effects and explanatory mental constructs and (2) encompasses many, if not all, effects that are studied in attitude research. In the second part of this chapter we outline the conceptual basis of our meta-theoretical framework: the definition of evaluation as the effect of stimuli on evaluative responses. Based on this definition, different attitudinal phenomena can be conceptualised as instances of evaluation involving different kinds of moderators. In the third part we explain how this overarching conceptualisation leads to a comprehensive alternative for the use of proxies in attitude research by revealing two mutually supportive levels of analysis: (1) a functional level that focuses on the environmental moderators of evaluation and (2) a cognitive level that investigates the mental processes and representations

that mediate evaluation.¹ In the fourth part we illustrate the implications of our functional-cognitive framework for theory construction and the interpretation of empirical data by applying it to four central questions of attitude research. Finally, the last part addresses some questions about our functional-cognitive framework to avoid potential misunderstandings about its central arguments and implications. One such potential misunderstanding concerns the primacy of the two levels of analysis. To be clear from the outset, we do not claim that the functional level is primary to, can replace, or is in any way superior to the cognitive level of analysis in attitude research. Instead, we believe that the two levels of analysis mutually support each other, in that progress at the functional level can strengthen the cognitive level and vice versa. However, this mutually supportive relation can be realised only to the extent that attitude researchers consistently avoid using behavioural proxies for mental constructs by separating to-be-explained effects from explanatory mental constructs. Defining attitudinal effects in terms of evaluation allows attitude researchers to do so.

ON THE PITFALLS OF TREATING BEHAVIOURS AS PROXIES FOR MENTAL CONSTRUCTS

What is the problem that we are trying to solve?

Although prominent scholars have repeatedly outlined the pitfalls of equating behavioural effects with mental constructs (e.g., Eagly & Chaiken, 2007; Fazio, 2007; Krosnick et al., 2005), the use of behaviours as proxies for attitudes and attitudinal processes is still very common. For example, in research on attitude-behaviour relations, a person's behavioural response on an evaluative rating scale is often treated as a proxy of this person's attitude towards an object and his or her object-related actions in real-life situations are treated as behavioural instances that may or may not be influenced by the attitude (Ajzen & Fishbein, 2005). This practice serves as the foundation for research guided by the theory of planned behaviour (Ajzen, 1991), which represents one of the most influential research programmes in social psychology. Yet, despite the widespread use of this practice, it is important to realise that a given behaviour can be treated as a proxy for a mental construct only by virtue of a priori assumptions about how the mental construct is related to that behaviour (De Houwer, 2011; Poldrack, 2006). Specifically, the practice of using behaviour as a proxy for mental

¹Note that we do not use the term *functional* in the sense of adaptive or something that serves a particular goal (e.g., Maio & Olson, 2000). Instead, the term is used to refer to relations between environment and behaviour; that is, to functions that map the environment onto behaviour. Moreover, the term *cognitive* is used in a broad sense that is synonymous with the term *mental*. It is supposed to refer to the complete set, rather than a subset, of all mental processes and representations, including affective and motivational constructs.

constructs presupposes that the presence and properties of the relevant behaviour (e.g., response on an evaluative rating scale) directly reflect the presence and properties of the mental construct that is thought to underlie the behaviour (e.g., attitude). In technical terms, treating behaviour as an index of a particular mental construct requires a bi-conditional relation between the two, such that variations in one unambiguously reflect variations in the other (i.e., “if p then q ” and, at the same time, “if q then p ”). However, most attitude researchers agree that such claims of bi-conditional relations are untenable, because variations in attitudes are neither necessary nor sufficient to produce variations in attitude-relevant behaviour (Krosnick et al., 2005). On the one hand, variations in attitudes are not necessary because attitude-relevant behaviour can be influenced by various other factors. On the other hand, variations in attitudes are not sufficient because attitudes may not always be expressed in attitude-relevant behaviours. For example, doubts have been raised about the usefulness of evaluative judgements as direct indices of attitudes, because evaluative judgements can be influenced by various other factors (e.g., transient mood states; see Schwarz, 1990) and the impact of attitudes on evaluative judgements can sometimes be disrupted (e.g., when people are motivated to conceal their attitudes; see Fazio, 2007). Similarly, evaluative priming effects can provide an unambiguous index of automatically activated attitudes only if variations in automatically activated attitudes are both necessary and sufficient to produce variations in evaluative priming. Yet priming effects are influenced by various factors over and above automatically activated attitudes (e.g., processes involved in response interference; see Gawronski, Deutsch, LeBel, & Peters, 2008) and the impact of automatically activated attitudes on evaluative priming can be reduced under certain conditions (e.g., through strategic counteraction; see Teige-Mocigemba & Klauer, 2013).

These considerations are relevant not only when probing mental representations, such as attitudes, but apply also when behavioural effects are used as proxies for the operation of particular attitudinal processes. As an example, consider the spreading-of-alternatives effect, which describes the phenomenon that choosing between two equally attractive alternatives leads to more favourable evaluations of chosen as compared to rejected alternatives (Brehm, 1956). Drawing on Festinger’s (1957) theory of cognitive dissonance, this effect is often interpreted as an index of post-decisional dissonance. Specifically, it is assumed that people experience an aversive feeling of post-decisional dissonance when they recognise either (1) that the rejected alternative has positive features that the chosen alternative does not have, or (2) that the chosen alternative has negative features that are not present in the rejected alternative. To reduce this uncomfortable feeling of cognitive dissonance, people are assumed to emphasise or search for positive characteristics of the chosen alternative and negative characteristics of the rejected alternative, which in turn leads to more favourable evaluations of the chosen compared with the rejected alternative.

Although the contribution of post-decisional dissonance to the spreading-of-alternatives effect has been confirmed in a substantial number of studies, treating this effect as a proxy of cognitive dissonance presupposes that the presence versus absence of the effect is diagnostic for the presence versus absence of cognitive dissonance (i.e., “if p then q ” and, at the same time, “if q then p ”). Yet, assuming such a bi-conditional relation seems problematic, because post-decisional dissonance is neither necessary nor sufficient for the spreading-of-alternatives effect to occur. It is not necessary, because the spreading-of-alternatives effect can emerge in the absence of cognitive dissonance as a result of mere ownership (Gawronski, Bodenhausen, & Becker, 2007) or simple methodological factors (Chen & Risen, 2010). Moreover, it is not sufficient, because cognitive dissonance can be resolved in many ways other than attitude change (Festinger, 1957). Nevertheless, the mere emergence of a spreading-of-alternatives effect is often interpreted as an index of cognitive dissonance, including prominent claims of post-decisional dissonance in children and monkeys (Egan, Santos, & Bloom, 2007) and amnesic patients who do not even remember their choice (Lieberman, Ochsner, Gilbert, & Schacter, 2001).

More generally, uncertainty about the mental causes of a particular behaviour in a given situation creates uncertainty about whether the presence of that behaviour provides a valid indicator of the presence of a particular mental construct. Furthermore, uncertainty about the mental causes of the behaviour can be reduced only if there are other behavioural indices that are known to be perfect indicators of the presence of particular mental constructs (for a discussion of these issues, see Bechtel, 2005; Chiesa, 1994; De Houwer, 2011; Poldrack, 2006). This dilemma creates a “catch-22” situation that seems virtually impossible to resolve.

In essence, the use of behavioural proxies for mental constructs is problematic because it conflates the behaviours that need to be explained (*explanandum*) with the mental constructs that are used to explain these behaviours (*explanans*). In a strict sense, scientifically sound explanations are supposed to keep the explanans conceptually independent of the explanandum, such that the explanans should not refer to the concepts of the explanandum and vice versa (Hempel, 1970). Treating attitude-relevant behaviours as indices of mental attitudes or attitudinal processes ignores the distinction between the behaviours that need to be explained and the mental constructs that are supposed to explain these behaviours. Common results of this practice are logical fallacies in theory construction and the interpretation of empirical data, such as backward inferences and the fallacy of affirming the consequent (i.e., inferring the presence of “ p ” from the presence of “ q ” on the basis of the conditional “if p then q ”; see Gawronski & Bodenhausen, in press).

What can we do about it?

One option to deal with these problems is to treat certain behaviours as “tentative” proxies of mental attitudes and attitudinal processes. Unfortunately such a strategy can hamper theory construction (see De Houwer, 2011). For example, if a certain behaviour is treated as a tentative proxy for a particular mental construct, theorising about this construct can be distorted when variations in that behaviour are due to mental constructs other than those it is thought to capture (e.g., when the spreading-of-alternatives effect is due to mechanisms other than cognitive dissonance). Moreover, using behaviours as tentative proxies can undermine advances in theorising by directing researchers’ attention away from alternative mechanisms that may produce the same behavioural effects (e.g., when mere ownership is ignored as a potential cause of the spreading-of-alternatives effect).

A better solution, we would argue, may be found at the conceptual level. As we noted above, the use of behavioural proxies for mental constructs is problematic because it conflates the behaviours that need to be explained (*explanandum*) with the mental constructs that are used to explain these behaviours (*explanans*). Such a conflation can be eliminated by defining the to-be-explained effects independently from explanatory mental constructs. Although prominent scholars have raised related arguments before (e.g., Eagly & Chaiken, 2007; Fazio, 2007), violations of this principle are still very common in the attitude literature. We believe that one reason for this discrepancy is the absence of an overarching meta-theoretical framework that avoids conceptual confluences of behavioural effects and mental constructs in a way that is applicable to a wide range of attitudinal phenomena. Until now, those researchers who defined attitudinal effects without referring to mental constructs often did so by simply describing the superficial properties of the independent and dependent variables that are involved in the effect at hand. For instance, instead of equating evaluative priming with the automatic activation of attitudes through processes of spreading of activation (e.g., Fazio et al., 1986), one can define it as faster responding to targets when they are preceded by primes with the same valence compared to primes with a different valence. Although such a definition does not refer to explanatory mental constructs and thereby avoids a conflation of effect and mental construct, it is a relatively concrete definition that is couched in terms that refer only to the procedure typically used in evaluative priming studies. In the following section we go beyond such idiosyncratic non-mental definitions by proposing a comprehensive, unifying way to define *all* attitudinal effects in non-mental terms. More specifically, we define attitudinal effects functionally as instances of evaluation. After explaining the core concepts of our approach and how they relate to attitudinal effects and mental constructs, we discuss the advantages of our approach and illustrate the implications of our framework for attitude research.

ATTITUDE RESEARCH AS THE STUDY OF EVALUATION

The basis of our meta-theoretical framework is the concept *evaluation*, which we define in functional, non-mental terms as the impact of stimuli on evaluative responses. In this section we discuss the meaning of the concept *evaluative response*, explain our definition of *evaluation*, and illustrate how attitudinal phenomena can be conceived of as instances of evaluation.

What are evaluative responses?

We propose that attitude research can be conceptualised as being concerned with a particular category of behaviours that we refer to as *evaluative responses*. This conceptualisation implies a distinction between behavioural responses that are evaluative and those that are not. From a methodological perspective, there are at least two ways in which such a distinction can be made.

First, the decision can be made on the basis of scientific convention. For some responses, such as self-reported evaluative responses (e.g., the number circled on a rating scale ranging from *very unpleasant* to *very pleasant*) or particular kinds of facial muscle actions (e.g., smiling versus frowning), consensus among researchers might be relatively easy to reach. Yet, for other kinds of responses such as response latencies in performance-based paradigms (e.g., facilitation scores obtained in an evaluative priming task), consensus about whether these responses should be regarded as evaluative responses may be more difficult to achieve (see Arkes & Tetlock, 2004, for an example).

Second, the decision can be based on the results of empirical validation research, which serves to support (or not) the relevant arguments in cases lacking consensus in the scientific community. For example, responses that are correlated with other responses that are thought to be evaluative in nature (e.g., self-reported evaluative responses) and that are influenced by stimuli that are known to evoke evaluative responses (e.g., normatively pleasant or unpleasant stimuli) could be considered as evaluative. Ultimately, however, the second approach depends on the first one (i.e., decision based on convention), because its validation criteria involve a reference to existing conventions. After all, relating a particular kind of response to another response that is thought to be evaluative requires consensus about the evaluative nature of the latter. Similarly, testing whether a particular kind of response is influenced by stimuli that are known to evoke evaluative responses presupposes that the evaluative quality of these stimuli has been established by means of another response that is consensually considered as evaluative.

Interpreted in this manner, definitions of the concept *evaluative response* are arbitrary insofar as they depend on scientific conventions about the criteria that can be used to distinguish evaluative from non-evaluative responses. However, whatever criteria are used to make this distinction, our meta-theoretical

framework presupposes that these criteria do not refer to mental constructs such as attitudes (e.g., defining evaluative responses as responses that are caused by attitudes). Instead, evaluative responses should be identified on the basis of physical or functional properties (e.g., responses that serve to increase or decrease the physical distance to the relevant object). Evaluative responses could either have those physical or functional properties themselves or be correlated with responses that have those properties (as established in validation research). Our meta-theoretical framework is consistent with any such conceptualisation to the extent that it does not refer to mental constructs. As we pointed out earlier, a definition of *evaluative response* that refers to mental constructs such as “attitudes” would violate the conceptual independence of explanans and explanandum, thereby reintroducing the problems that we aim to resolve. Note also that defining evaluative responses in terms of attitudes does not allow one to escape the question of how the physical or functional features of evaluative responses (i.e., the subset of behaviours that attitude research focuses on) differ from those of other responses. Such a mental definition leaves open the question of what it is about evaluative responses that makes them susceptible to attitudes whereas other responses are not. To answer this question in a non-circular manner (i.e., without referring to attitudes), reference has to be made to differences between the physical or functional properties of evaluative and non-evaluative responses.

What is evaluation?

Expanding on the notion of *evaluative response*, we define *evaluation* as the causal effect of stimuli on evaluative responses. Defined in this manner, propositions about the occurrence of evaluation involve three components: (1) a statement about the presence of an evaluative response, (2) a statement about the presence of a stimulus, and (3) a statement about a causal link between the two. Whereas the presence of a stimulus and an evaluative response can most often be observed, a causal relation between the two has to be inferred. In many cases it can be difficult to determine whether an evaluative response has been caused by a particular stimulus, because there are many stimuli that could be responsible for the evaluative response. This situation is not much different for individuals themselves, in that it can be rather difficult to discern the true environmental causes of one’s own evaluative responses (Wilson & Dunn, 2004).

Although philosophers and psychologists debated for centuries about how causality can be inferred (e.g., Cheng, 1997; Hume, 1739/1987; Kant, 1781/1965; Sosa & Tooley, 1993), there is general agreement that the experimental method provides a useful tool to establish the presence of a causal relation. For example, attitude researchers can compare evaluative responses between an experimental and a control condition that differ only with regard to the presence of a particular stimulus (property). If the evaluative response is different in the

experimental condition than in the control condition, one can conclude that the stimulus (property) had a causal effect on the evaluative response, and thus that evaluation of the stimulus (property) has occurred.

The statement that a stimulus is the cause of a given behaviour merely implies that the behaviour is a function of that stimulus (*functional causation*); it does not commit itself to any ideas about the processes by which the stimulus influences behaviour (e.g., as the result of a specific force or mechanism; see Chiesa, 1992, for a detailed discussion). The presence of the stimulus can be conceived of as an independent variable that determines the dependent variable that is behaviour. Of course, researchers often make theoretical assumptions about the mental mechanisms that mediate functional relations (*mechanistic causation*), but it is important to realise that functional causes can be identified even when the mediating mechanism is unknown. For instance, provided that adequate control conditions are implemented, experimental studies on evaluative priming allow one to determine whether the speed of responding to target stimuli is determined by the relation between the prime and the target, or by the relation between the prime and the required response (see De Houwer, 2003). Although such experiments might have important implications for cognitive theories of evaluative priming, it is possible to formulate conclusions about the functional relations without knowing the mental processes by which prime–target or prime–response relations influence the speed of responding to the targets.

Attitudinal phenomena as instances of evaluation

The functional definition of evaluation allows one to systematically define in a functional, non-mental manner a wide range of the behavioural phenomena that are studied in attitude research. As such, the concept *evaluation* forms the conceptual basis of the comprehensive meta-theoretical framework that we were looking for. More specifically, we propose that different attitudinal phenomena involve different instances of evaluation or different moderators of evaluation. For example, evaluative ratings can be defined as the effect of stimuli on one particular type of evaluative response, namely the selection of a valenced stimulus (e.g., a value on an evaluative rating scale). Evaluative priming can be conceived of as an instance of evaluation that involves a different type of evaluative response, being the time needed to respond to a valenced stimulus (e.g., a positive or negative target word). The spreading-of-alternatives effect (i.e., choosing between two equally attractive alternatives leads to a more favourable evaluation of the chosen as compared to rejected alternatives) could be seen as a demonstration of the moderating impact of a contextual factor (i.e., a choice between certain alternatives) on evaluation. Importantly, because evaluation implies only functional causation, it can be moderated not only by elements of the environment that are present when the evaluative response is observed but also by the history of the organism. For example, the fact that a stimulus evokes

positive rather than negative evaluative responses in the present may be attributed to pairings of that stimulus with other (positive) stimuli in the past (e.g., De Houwer, 2007). In a similar way, many other behavioural phenomena that are studied in attitude research can be said to involve evaluation, that is, some kind of stimulus that evokes some kind of evaluative response in a particular organism and a particular context. Although they all involve evaluation, different phenomena can be distinguished on the basis of the kind of stimulus, evaluative response, organism, and context that are involved. We will provide more detailed examples of these functional definitions of behavioural phenomena later in this chapter.

The main advantage of adopting our functional conceptualisation of attitudinal phenomena is that it provides a systematic way to avoid the conflation of to-be-explained behavioural effects and explanatory mental constructs. It allows us to define a wide variety of attitudinal effects without reference to mental constructs, thereby going far beyond the idiosyncratic non-mental definitions that have been proposed in the past (e.g., evaluative priming as better performance when target and prime have the same valence). Thus a major contribution of our conceptualisation is that it provides a unifying perspective entailing that all attitude research is concerned with the study of evaluation. This overarching conceptualisation sheds light not only on what different attitudinal phenomena have in common, but also on the differences between these phenomena. There is, however, an important additional benefit of our conceptualisation. Because we have defined evaluation in functional, non-mental terms, and because we did this in a manner that is applicable to a wide variety of attitudinal phenomena, our conceptualisation helps to identify two distinct, yet mutually supportive, levels of analysis in attitude research. These two levels of analysis, as well as their interaction, are described in the following section.

TOWARDS A FUNCTIONAL-COGNITIVE FRAMEWORK FOR ATTITUDE RESEARCH

Defining attitudinal phenomena as instances of evaluation allows researchers to clearly delineate and relate two levels of analysis in attitude research: (1) a functional level of analysis that provides explanations of evaluative responses in terms of stimuli in the environment, and (2) a cognitive level of analysis that offers explanations of evaluation in terms of mental processes and representations.

Functional level of analysis

Our definition of evaluation as the causal impact of a stimulus on evaluative responses implies that statements about evaluation go beyond mere observations. Rather, statements about evaluation are theoretical in the sense that they imply a

causal hypothesis about which element in the environment is responsible for an observed evaluative response. To illustrate this issue, imagine that you attend a talk during a conference. During the talk you see a member of the audience smiling. If you as an observer conclude that the smile reflects a positive evaluation of the content of the talk, then you put forward the hypothesis that the content of the talk is the functional cause of the smile. However, it is also possible that the smile was elicited by other elements in the (current or past) environment (e.g., the way in which the speaker dresses). In this case the smile would no longer qualify as an evaluation of the content of the talk. Thus statements about evaluations are not “mere” descriptions of observations, but instead represent hypothetical functional explanations of evaluative responses.

Attitude researchers usually test their hypotheses about a causally effective stimulus (property) by studying naturally occurring or experimentally produced covariations between the stimulus (property) and evaluative responses (e.g., by comparing the effects of Black vs. White face primes in an evaluative priming task). Such investigations fit within the functional approach in psychology that aims to explain behaviour by identifying its environmental causes (see Chiesa, 1992, 1994). Learning more about the conditions under which attitudinal effects occur helps us to understand how evaluative responses are determined by the environment. In functional explanations, variations in evaluative responses represent the events that need to be explained (*explanandum*) and elements in the environment serve as the entities that explain these events (*explanans*). Thus functional explanations respect the general scientific principle that the *explanandum* needs to be separated conceptually from the *explanans*.

There is also another way in which functional explanations go beyond “mere” description. Functional hypotheses about the cause of an evaluative response can differ in their level of sophistication. At a low level of sophistication, one can describe causal effects in terms that are linked directly to a specific procedure (e.g., evaluative priming as the impact of words on the speed of responding to other words). More sophisticated explanations use abstract terms that can be applied to a well-delineated subset of different situations or elements of situations (i.e., terms that have a broad scope and, at the same time, a high level of precision; see Hayes, Barnes-Holmes, & Wilson, 2012). Functional researchers strive towards sophisticated functional explanations because they can reveal similarities and differences between different effects that would otherwise remain hidden. Our definition of *evaluation* fits well with this aim because it has a scope and level of precision that allows for more sophisticated functional explanations of attitudinal phenomena than idiosyncratic explanations of individual effects that are unrelated to the epistemic goals of attitude researchers (e.g., evaluative priming as the impact of stimuli on the speed of responding to other stimuli).

Another important goal of functionally oriented attitude research is to identify environmental conditions that moderate the effect of a given stimulus feature on evaluative responses. These moderators may include prior experiences with the

stimulus (stimulus history), the presence of other stimuli (stimulus context), and the nature of the evaluative response. For example, researchers interested in racial prejudice may investigate whether the effect of racial outgroup faces on evaluative responses is moderated by prior contact with outgroup members (e.g., Pettigrew & Tropp, 2006), features of the context in which these faces are encountered (e.g., Barden, Maddux, Petty, & Brewer, 2004), and the type of task that is used measure evaluative responses (e.g., Fazio, Jackson, Dunton, & Williams, 1995). These moderators do not have to be studied in isolation, but can also be investigated for interactive effects. For example, researchers may be interested in whether the moderating effects of intergroup contact differs as a function of the type of evaluative response that is recorded and features of the context in which those evaluative responses are observed.

The overarching goal of functionally oriented research is to generate increasingly abstract explanations of evaluative responses that stipulate in an accurate (i.e., in line with the observed data), unambiguous (i.e., using concepts that leave little room for misunderstandings), and economical way (i.e., using as few concepts and assumptions as possible) the conditions under which elements of the environment cause and moderate evaluative responses. This abstract knowledge goes beyond a simple list of individual empirical facts (e.g., findings in single experiments) in that it is assumed to apply to a wide range of situations within and outside the laboratory. Thus the ultimate challenge for functional research is to derive from a large set of empirical observations those common elements that are crucial in moderating evaluation. At the same time, purely functional research is limited in its focus, in that it is primarily concerned with the question of *when* evaluation occurs (i.e., moderators of evaluation); it does not direct itself towards understanding *how* evaluation occurs (i.e., mediators of evaluation). The latter question is central to the cognitive level of analysis.

Cognitive level of analysis

When attitude research is conceived of as the study of evaluation, cognitive attitude theories can be conceptualised as hypotheses about the mental processes and representations that mediate evaluation. All cognitive attitude theories have in common the assumption that a stimulus can cause an evaluative response only by virtue of mental representations that are formed and activated by stimuli in the environment and that influence behaviour by virtue of certain mental processes. Yet this minimal assumption does not go beyond the basic claim that effects of stimuli on behaviour can be explained by mediating mental processes and representations, which is the bedrock of cognitive science. Various concepts have been used to refer to these mediating mental processes and representations. Although all definitions are to some extent arbitrary and open to debate, our meta-theoretical framework implies that a clear distinction should be made between concepts at the mental level and concepts at the functional level. In

other words, concepts should be defined either at the functional level (e.g., evaluation) or the mental level (e.g., attitude).

Using mental constructs, cognitive attitude theories aim to describe mental mechanisms that explain (1) why a given stimulus has the capacity to cause a particular evaluative response, and (2) why the effect of the stimulus on evaluative responses is moderated by particular environmental conditions. Thus, whereas the functional level of analysis provides explanations of evaluative responses (explanandum) in terms of stimuli in the environment (explanans), the cognitive level of analysis offers explanations of evaluation (explanandum) in terms of mental processes and representations (explanans).

The merit of cognitive attitude theories can be evaluated on the basis of whether they (1) account for existing knowledge about the environmental conditions that moderate evaluation (i.e., heuristic value) and (2) lead to the discovery of new knowledge about the environmental conditions that moderate evaluation (i.e., predictive value). For example, the hypothesis that evaluation is based on the activation of nodes in associative networks (e.g., Fazio, 2007) has merit because it correctly predicts that stimuli can cause evaluative responses automatically (because activation in the network can spread automatically through associations). Because the structural characteristics that maximise the heuristic versus predictive value of a given theory can sometimes be in conflict (e.g., when a theory accounts for everything, yet predicts nothing; see Quine & Ullian, 1978), an empirically confirmed hypothesis is typically assigned more value when it is made before (i.e., predictive value) rather than after (i.e., heuristic value) the hypothesised effect has been observed.

Importantly, this way of testing cognitive attitude theories provides an alternative to testing those theories by using proxies. In other words, our conceptualisation of attitude effects as instances of evaluation not only provides a comprehensive way to avoid treating those effects as proxies for mental constructs; it also removes the need for proxies of mental constructs in testing cognitive attitude theories. To illustrate the difference between the proxy approach and our alternative approach, again take the example of associative network theories. The presence of an evaluative priming effect is often interpreted as a proxy of the automatic activation of attitudes via spreading of activation in an associative network. Hence the emergence of an evaluative priming effect is typically viewed as a *demonstration* of this process. Within our conceptual framework, however, observing an evaluative priming effect is merely an instance of evaluation (i.e., the impact of the prime on the speed of emitting an evaluative response) that occurs under certain conditions (e.g., when the prime is presented only briefly and immediately before the target). The fact that such an effect occurs is in line with the idea that attitudes are stored as evaluative associations in an associative network through which activation can spread automatically. However, the consistency between theory and data does not guarantee the absence of another cognitive theory that allows for evaluation

under those conditions. Hence the observation that an evaluative priming effect occurs under certain conditions can never be treated as proof of the hypothesis that a specific mental process (e.g., spreading of activation) occurs under those conditions (cf. Gawronski & Bodenhausen, in press). Doing so would imply that the effect is a proxy of a specific process. Nevertheless, knowledge about when evaluative priming occurs *constrains* theories of the mental processes that mediate evaluative priming effects. A theory of evaluative priming should be able to account for the fact that evaluative priming occurs under certain conditions. This example illustrates that the more we know about the conditions under which evaluation occurs, the more we can constrain theoretical ideas about the mental processes and representations that mediate evaluation (see Jacoby & Sassenberg, 2011). Moreover, the higher the constraints that are imposed by an individual finding, the stronger will be the conclusion that can be reached at the cognitive level. In most cases, however, it is more likely that multiple findings are needed to achieve such progress. In any case, our framework implies that the validation of cognitive attitude theories can be achieved by testing the conditions under which evaluation occurs rather than by simply searching for behavioural effects that provide proxies for mental processes.

As we noted earlier, behavioural effects and mental constructs can also be separated by adopting idiosyncratic definitions of attitudinal effects. Again consider an idiosyncratic definition of evaluative priming as faster responses when prime and target have the same rather than opposite valence. Such a definition would allow one to investigate moderators of the evaluative priming effect, and to use this information to inform the development of cognitive theories of evaluative priming, without making any a priori assumptions about whether and how mental processes such as spreading of activation are related to evaluative priming effects. This idiosyncratic approach would be a major improvement to the common practice of equating evaluative priming effects with mental processes such as spreading of activation. However, because our approach is much more comprehensive, it has several advantages compared to an idiosyncratic approach. When defining effects such as evaluative priming in idiosyncratic terms, one runs the risk that both research on the moderators of the effect and mental process theories about the mediators of the effect also become very idiosyncratic; that is, effect-specific. Functional research on the moderators of the effect will be limited to examining when this particular effect occurs. Likewise, mental process theories about the effect will focus on accounting for the functional knowledge about when that specific effect occurs. As a result, both functional and cognitive attitude research becomes fractionated. Within our approach it is still possible to examine separately when and how a given effect arises, but it is also possible to relate different attitudinal phenomena in functional terms; that is, as instances of evaluation. As such, our conceptualisation provides a comprehensive and overarching framework for research on evaluation in general (rather than for research on individual attitudinal effects). In addition

to providing such an overarching framework, our conceptualisation encourages researchers to determine in a systematic manner for each attitudinal phenomenon how it can be defined functionally as an instance of evaluation. In that way the framework can help to detect instances in which attitudinal effects are not defined functionally and hence to reveal hidden theoretical assumptions in the definition of those effects.

The mutual relation between functional and cognitive explanations

Our discussion of the way in which cognitive theories are evaluated on the basis of their heuristic and predictive value indicates that functional explanations and cognitive explanations are not in competition, but mutually support each other (Bechtel, 2005; De Houwer, 2011). Progress in functional explanations of evaluative responses (i.e., knowledge about the elements in the environment that determine evaluative responses) provides the empirical basis that is necessary to develop and test cognitive explanations of evaluation (i.e., knowledge about the mental mechanisms by which stimuli produce evaluative responses). Vice versa, progress in cognitive explanations of evaluation can lead to progress in functional explanations of evaluative responses by organising existing knowledge about the elements in the environment that determine evaluative responses (i.e., heuristic value) and through the generation of new predictions about elements in the environment that might influence evaluative responses (i.e., predictive value). However, this mutually supportive relation between functional and cognitive explanations can succeed only if the two levels of explanation are separated conceptually, because they differ in terms of their explanandum (i.e., variations in evaluative responses vs. variations in evaluation) and in terms of their explanans (i.e., elements in the environment that moderate the causal effect of stimuli on evaluative responses vs. mental processes and representations that mediate the causal effect of stimuli on evaluative responses). We refer to this meta-theoretical framework of conceptually distinct, but mutually supportive, functional and cognitive levels of analysis as the functional-cognitive framework for attitude research.

We want to emphasise that there is no fixed hierarchy between the two levels of analysis in terms of their scientific importance. Which level is given primacy depends on philosophical considerations (for an overview of the philosophical underpinnings of the functional and cognitive approach in psychology, see Chiesa, 1992, 1994; Gardner, 1987; Hayes, 1995). Yet, regardless of which level of explanation is preferred by an individual researcher, it is important to note that progress at one level of explanation is tightly connected to progress at the other level of explanation. Therefore both functionally and cognitively oriented researchers can benefit from embracing both levels of explanation. Moreover, an integrative focus combining both functional and cognitive

explanations can avoid potential problems resulting from an exclusive focus on one type of explanation, as we will explain in the final section of this chapter.

PUTTING THE FUNCTIONAL-COGNITIVE FRAMEWORK TO WORK

In the previous section we briefly explained in general terms how functional research on evaluation can lead to progress in cognitive attitude theories, and vice versa. In the present section we provide more concrete examples of how a functional-cognitive approach can promote progress in attitude research. Towards this end, we will focus on four core questions in traditional attitude research: (1) How are attitudes acquired? (2) How are attitudes activated? (3) How do attitudes influence behaviour? (4) How are attitudes represented? We argue that each of these questions relates to functional knowledge about particular moderators of evaluation; that is, to specific sets of conditions under which evaluation does or does not occur. In each case we illustrate how our conceptualisation can solve problems that have resulted from the use of behavioural proxies for mental constructs.

The acquisition of attitudes: Varying the nature of the stimulus history

Theories about the acquisition of attitudes are constrained primarily by findings about changes in evaluation. More specifically, one can argue that a new attitude has been acquired if two conditions are met. First, it needs to be verified that a stimulus initially evokes a particular evaluative response, but a different evaluative response at a later point in time.² Second, this change in evaluation needs to be the result of experiences that are somehow related to this stimulus. The second condition distinguishes the acquisition of attitudes from the spontaneous development or emergence of attitudes (e.g., because of genetic factors; see Olson, Vernon, Harris, & Jang, 2001). This conceptualisation implies that research on the acquisition of attitudes focuses on the conditions under which prior experiences with a stimulus determine the nature of the evaluative response that is

²The term *attitude acquisition* is meant to subsume both attitude change and attitude formation. Attitude change can be defined as a change in the particular quality of the evaluative response that is evoked by a stimulus as the result of new experiences related to that stimulus. It differs from attitude formation, in which the stimulus initially does not evoke an evaluative response and at a later time does evoke an evaluative response. Yet both research on attitude change and research on attitude formation deal with the effect of new experiences, which is supposed to be captured by the term *attitude acquisition*. Like the study of attitude acquisition, research on the malleability of attitudes also deals with variations in the evaluative response that is evoked by a given stimulus. In this case, however, the focus is not on the moderating impact of new experiences with a stimulus, but on the way in which changes in context moderate evaluation.

evoked by that stimulus. Insights into the acquisition of attitudes can be achieved by comparing at different moments in time evaluations of one stimulus or by comparing at one moment in time evaluations of stimuli that differ only with regard to certain prior experiences. For example, stimuli can differ with regard to whether they were experienced before (as is the case in studies on mere exposure; see Bornstein, 1989, for a review), whether they co-occurred with other valenced stimuli—as is the case in studies on evaluative conditioning (EC), (see De Houwer, Thomas, & Baeyens, 2001, for a review)—or whether they were described in a positive or negative manner in verbal statements (as in the studies on persuasive communication; see Crano & Prislin, 2006, for a review). Importantly, the observed change in evaluation is not treated as a proxy for a specific mental process (e.g., formation of associations, see below) but merely as a behavioural effect that constrains cognitive theories on attitude acquisition.

Conceptualised in this manner, the goal of research on the acquisition of attitudes is twofold. First, it aims at examining the environmental conditions under which prior experiences moderate evaluation. For example, one can examine whether the effect of stimulus pairings on evaluative responses (i.e., EC effect; see De Houwer, 2007) depends on aspects of the stimulus pairings (e.g., whether they appear simultaneously or sequentially; see Sweldens, Van Osselaer, & Janiszewski, 2010), the stimuli as such (e.g., how salient they are; see Jones, Fazio, & Olson, 2009), the type of evaluative response that is registered (e.g., evaluative ratings or evaluative priming; see Gawronski & LeBel, 2008), or the broader context during the presentation of the stimulus pairings (e.g., whether participants are engaged in other tasks; see Pleyers, Corneille, Yzerbyt, & Luminet, 2009). The second aim is to develop cognitive theories that describe the mental mechanisms by which prior experiences moderate evaluation (e.g., Gawronski & Bodenhausen, 2006; Petty & Cacioppo, 1986). Thus research on the acquisition of attitudes represents a particular segment of our functional-cognitive framework, namely the segment that deals with the moderating effect of the history of stimuli on evaluative responses.

Conceptualising attitude acquisition in this manner offers an alternative to the use of behavioural proxies for mental constructs. As argued earlier, behavioural effects have often been used as proxies for mental processes that were assumed to underlie the acquisition of attitudes. Consider, for example, research on EC. EC is often conceived of as a mental process by which the pairing of stimuli automatically and gradually results in the formation of associations in memory (e.g., Petty & Briñol, 2010). As discussed in more detail by De Houwer (2007), defining EC in this manner restricts theoretical ideas about EC to those that refer to automatic association formation. However, recent evidence raises doubts about the widespread assumption that EC can occur unconsciously and thus about the idea that automatic association formation mediates EC (for a meta-analysis, see Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010). When the definition of EC as a process of automatic association formation is taken seriously,

such evidence would lead to the conclusion that EC does not exist and therefore should not be studied.

The framework that we propose in this chapter makes a clear distinction between EC as a behavioural effect and the mental processes that might underlie this effect. Hence evidence against a particular mental process theory of EC cannot be mistaken as evidence against EC as an effect. Moreover, it allows one to take seriously the idea that EC effects may be due to several mental processes, including processes that do not involve the automatic formation of associative links (De Houwer, 2007; Gawronski & Bodenhausen, 2011). For example, it has been proposed that EC effects might often result from the non-automatic acquisition and truth evaluation of propositional knowledge about stimulus relations (De Houwer, 2009a; Mitchell, De Houwer, & Lovibond, 2009), which is supported by the results of a recent meta-analysis (Hofmann et al., 2010). The idea that EC effects can be produced by different types of mental processes also sheds light on many conflicting results that have been observed in the literature (De Houwer, 2007; Sweldens et al., 2010).

Although the proposal to define EC as an effect was put forward some years ago (see De Houwer, 2007), our functional-cognitive framework encompasses many other effects. As such, it reveals that EC is concerned with one type of moderator of evaluation, being the stimulus history. Within the class of effects that deals with stimulus history, EC is unique in that it concerns one specific type of experience (i.e., stimulus pairings; see De Houwer, Barnes-Holmes, & Moors, 2013). There are many other effects that also deal with changes in evaluation but that concern other types of experiences (e.g., mere exposure effects as a change in evaluation that is due the repeated presentation of a stimulus).

The functional-cognitive approach can be applied not only to EC but also to these other effects. Consider the recently introduced self-referencing effect (e.g., Prestwich, Perugini, Hurling, & Richetin, 2010; see also Ebert, Steffens, von Stülpnagel, & Jelenec, 2009, for a related proposal). In studies on the self-referencing effect, participants are asked to respond to stimuli from four categories by pressing one of two keys. For instance, Prestwich et al., (2010) asked participants to press a left key when they saw pictures of a novel Drink A or words related to the self (e.g., me, I) and to press a right key when pictures of a novel Drink B or words related to others (e.g., they, he) were presented. As a result, Drink A was liked more than Drink B. Originally, the self-referencing effect was introduced as a type of EC because it was said to be “based on associating one’s self with a particular drink and others with a contrast drink” (Prestwich et al., 2010, p. 62). However, from a strict functional point of view, the two effects differ because the change in liking is due to other types of experiences. Whereas EC refers to changes in evaluation that are due to the pairing of stimuli in space and time, the self-referencing effect involves changes in evaluation that result from assigning the same discriminative function to stimuli (i.e., the function of signalling that a particular response should be

selected; see Perugini, Richetin, & Zogmaister, 2012). Based on this functional analysis, we are now exploring a whole new class of effects in which evaluation changes as the result of overlap in different kinds of functions. For instance, in one as yet unpublished experiment participants experienced that a first response was followed by a positive stimulus on some trials and by a neutral stimulus A on other trials. The second response was sometimes followed by a negative stimulus and sometimes by a neutral stimulus B. As a result of this procedure, neutral stimulus A was subsequently liked more than neutral stimulus B. Hence an overlap in outcome function (e.g., the positive stimulus and stimulus A both function as an outcome of the first response) can also lead to a change in liking. The discovery of these new effects was a direct consequence of functional analyses of EC and the self-referencing effect.

The activation of attitudes: Varying the nature of the stimulus context

From a cognitive perspective, one could argue that evaluation as an effect can occur only if some kind of attitudinal representation has been formed, activated, and influences behaviour.³ Hence cognitive theories about the activation of attitudes are constrained by knowledge about the environmental conditions that moderate evaluation. Traditionally, however, certain classes of moderators are excluded from this research. Most importantly, prior experiences should be kept constant between conditions, because differences in prior experiences could influence the attitude itself rather than the activation of the attitude. This ambiguity is the reason why most studies on the activation of attitudes focus on how properties of the momentary context moderate evaluation.⁴ For example, one can test whether a stimulus elicits different evaluative responses depending on the presence versus absence of other stimuli in the environment (e.g., Gawronski, Rydell, Vervliet, & De Houwer, 2010). Other properties of the context that can be manipulated are whether participants are asked to provide their response under time pressure (e.g., Ranganath, Smith, & Nosek, 2008) or the extent to which participants can discriminate above chance-level between the

³Some might argue that an attitude is involved only in instances of evaluation that involve a dispositional component that is stable over time (e.g., Eagly & Chaiken, 2007). Such a surplus meaning can be captured in functional terms by specifying the conditions under which attitudes are assumed to underlie evaluation (e.g., instances of evaluation that remain stable over time).

⁴The context may also contain other stimuli that influence evaluative responses directly and independent of the target object. For example, evaluative judgements about an object can be influenced by whether the judgements are registered on a sunny day or on a rainy day (e.g., Schwarz & Clore, 1983). Within the functional-cognitive framework, effects of this kind represent behavioural effects that can be explained by referring to mental constructs such as mood. The critical difference is whether the context *directly* influences evaluative responses or instead *moderates* the causal effect of the target stimulus.

presence versus absence of the response-evoking stimulus (e.g., Ottati, Coats, Mae, DeCoster, & Smith, 2010). Ideally the effects of these variables should generalise across different types of evaluative responses to make sure that the contexts influenced the activation of the attitude, rather than the impact of the attitude on behaviour.⁵

Again we want to point out that observing an effect of context on evaluation does not provide a proxy for any single attitude activation process. It only provides information about the moderators of evaluation, which in turn contributes to the cognitive level by constraining the mental process theories that are possible at that level. Hence, from the perspective of the functional-cognitive framework, research on the activation of attitudes has two aims: (1) to describe whether and when certain aspects of the stimulus context moderate evaluation, and (2) to develop cognitive theories that explain why these contextual factors moderate evaluation in that manner.

As is the case for research on the acquisition of attitudes, the functional-cognitive approach provides an alternative for the use of proxies for mental constructs in research on attitude activation. As an example, consider the literature on implicit and explicit attitudes. A common assumption in this literature is that implicit attitudes are activated automatically upon the encounter of a stimulus, whereas explicit attitudes require controlled processes to be retrieved from memory (e.g., Wilson, Lindsey, & Schooler, 2000). To the extent that indirect, performance-based measures (e.g., evaluative priming tasks) reduce participants' opportunity to engage in controlled processing, responses on indirect measures are often treated as an index of implicit attitudes, whereas responses on direct self-report measures are equated with explicit attitudes (e.g., Greenwald & Banaji, 1995).

Such a use of proxies can, however, hamper attitude research in several ways. Mental states such as attitudes influence behaviour on a given task only by virtue of task-specific mechanisms that translate the mental state into an overt behavioural response (see De Houwer, 2009b; Gawronski et al., 2008). Therefore the absence of an evaluative response on an indirect measure might be due either to the fact that the implicit attitude was not activated or to the fact that it did not influence the observed response (e.g., Gawronski, Cunningham, LeBel, & Deutsch, 2010). Likewise, variables that influence evaluative responses on indirect measures might do so not because they influence the activation of implicit

⁵ Even though the activation of attitudes is typically studied through manipulations of the stimulus context, effects of the stimulus context are not diagnostic about differences in attitude activation when these effects are limited to particular types of evaluative responses. For example, the presence versus absence of other people may influence evaluative responses that are relatively easy to control (e.g., evaluative ratings), but not those that are relatively difficult to control (e.g., responses in an evaluative priming task). Such discrepancies between different types of evaluative responses represent the basis for studies on the impact of attitudes on behaviour, which are discussed in more detail in the following section.

attitudes, but because they affect the mechanisms through which implicit attitudes influence behaviour (e.g., Deutsch & Gawronski, 2009).

We believe that research on implicit attitudes would benefit from conceptualising it as the study of automatic evaluation. Whether, and the way in which, evaluation is automatic needs to be specified in terms of the impact of elements of the context on evaluation. For example, evaluation can be said to be automatic if it occurs even when the context has properties that could potentially reduce the likelihood of evaluation. More specifically, evaluation can be studied in contexts in which participants (1) are not instructed to respond in an evaluative manner to the evoking stimuli, (2) are instructed to prevent, stop, or alter the evaluative response that a stimulus evokes, (3) are instructed to engage in other difficult tasks, (4) are not able to discriminate between the presence or absence of the evoking stimuli, between the presence or absence of the causal impact of the stimuli on the evaluative response, or between the presence or absence of the evaluative response that is evoked by the stimulus, or (5) have little time to respond. If evaluation occurs under one or more of these conditions, it can be described as automatic. Another approach would be to avoid using the term *automatic* as an umbrella concept and refer directly to the conditions themselves. If, however, one chooses to continue to use the concept *automatic*, one should recognise that it only makes sense to do so only if one always specifies the conditions to which the concept *automatic* is meant to refer in that specific instance (De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009).⁶

Once it has been established that a particular instance of evaluation is automatic in a certain sense, it can be further examined (1) what environmental conditions moderate automatic evaluation and (2) which mental processes mediate automatic evaluation. The functional-cognitive approach thus allows one to clearly separate automatic evaluation as a behavioural effect and implicit attitudes as a mental construct that can be used to explain automatic evaluation. One could even entertain theories that dispense with the concept *implicit attitudes* as a separate mental entity to *explicit attitudes* (e.g., Fazio, 2007). There is also no need to limit oneself to theories that invoke associations as an explanatory mental construct. Historically, research on implicit attitudes has almost always been conceptualised in terms of the cognitive process of spreading of activation via associations in memory, presumably because it provides a possible mechanism by which stimuli could automatically evoke evaluative responses (see Hughes, Barnes-Holmes, & De Houwer, 2011, for a discussion). From a functional-

⁶Note that, in contrast to cognitive analyses of the concept *automaticity* (e.g., Bargh, 1992; Moors & De Houwer, 2006), we do not refer to mental constructs such as goals or mental resources in our definition of automaticity features (for more details, see De Houwer & Moors, 2012). The reasoning behind this approach is that it remains agnostic with regard to specific theories about the mental processes that underlie automatic effects (e.g., particular assumptions about the nature of mental resources; see Baddeley, 2010).

cognitive perspective, however, it is entirely possible that non-associative processes mediate automatic evaluation. In line with this possibility, De Houwer (in press; Hughes et al., 2011) recently proposed that automatic evaluations might be driven by the automatic retrieval of propositional knowledge. Whereas associations involve simple unqualified links between mental representations, propositions include information about *how* events are related. De Houwer (2013) reviewed evidence showing that the automatic activation of propositional knowledge is a possible mechanism underlying automatic evaluation, as well as evidence in line with a propositional account of automatic evaluation. For instance, context information about how stimuli are related has been shown to moderate the impact of stimulus pairings on the automatic evaluation of those stimuli (e.g., Zanon, De Houwer, & Gast, 2012). These novel ideas and findings make perfect sense within (and were actually inspired by) a functional-cognitive framework in which automatic evaluation is seen as an effect that can be due to different types of mental processes.

The effect of attitudes on behaviour: Varying the nature of the evaluative response

In addition to the question of how attitudes are acquired and activated, there is the important question of how attitudes influence behaviour (e.g., Ajzen, 1991; Fazio, 1990). As noted above, from a cognitive perspective, one could argue that evaluation as an effect can occur only if some kind of attitudinal representation is acquired, activated, and influences behaviour. Given this assumption, cognitive theories about how attitudes influence behaviour can be constrained by learning more about the effect of stimuli on different types of evaluative responses while keeping constant those factors that influence the acquisition (i.e., stimulus history) and activation (i.e., stimulus context) of attitudes. Hence research on the impact of attitudes on behaviour can be conceptualised in terms of how the nature of the evaluative response moderates evaluation. Consistent with our general approach, this conclusion does not imply that (the impact of the nature of the evaluative response on) evaluation is seen as a proxy for specific attitudes or specific processes by which attitudes influence behaviour, because such an approach would require additional assumptions (e.g., that evaluation is determined *only* by those specific attitudes or processes).

We believe that a clear separation between functional and cognitive levels of analysis could advance research on the relation between attitudes and behaviour in significant ways. In the past, attitude researchers have invested much effort in devising behavioural proxies of attitudes because they believed that these proxies, as attitude measures, would allow them to predict a wide variety of attitudinal behaviours (Krosnick et al., 2005). For example, a person's evaluative response on a rating scale is often treated as a proxy of this person's attitude towards a given object and his or her object-related actions in real-life situations

as are treated as behavioural instances that may or may not be influenced by the attitude (Ajzen & Fishbein, 2005). However, there is no a priori reason that justifies the treatment of one evaluative response as a superior measure of attitudes (see Kaiser, Byrka, & Hartig, 2010). After all, it is also possible that object-related actions in real-life situations represent more reliable reflections of attitudes than responses on evaluative rating scales. In fact, there is no empirical way to support either of the two claims, because the ambiguity about which of the two behaviours is the more reliable indicator of attitudes can be resolved only if there are other behavioural indicators that are known to be perfect indicators of attitudes. As we outlined in our introduction, this problem creates a “catch-22” situation for the use of behavioural proxies that seems impossible to resolve.

From the perspective of our functional-cognitive framework, the question of attitude-behaviour consistency is not concerned with the consistency between a mental construct and behaviour, but the consistency between the different types of behaviour, more specifically, different types of evaluative responses. Interpreted in this manner, research on attitude-behaviour consistency is focused on whether evaluation is moderated by the type of evaluative response that is observed. At the functional level, the primary aim is to investigate whether the same stimulus can elicit different responses for different kinds of evaluative responses (e.g., saying that one likes the stimulus but avoiding the presence of that stimulus). At the cognitive level, the main goal is to develop mental process theories that explain why the same stimulus can elicit functionally different types of evaluative responses.

From a functional perspective, there are several reasons why effects of stimuli on one type of evaluative response might fail to covary with the effect of the same stimuli on another type of evaluative response. First, the two kinds of evaluative responses might be influenced by different stimuli or different features of a stimulus (see Ajzen & Fishbein, 1977). Consider, for example, the observation that the verbal endorsement of ecologically friendly products often fails to predict the purchase of those products (e.g., Vantomme, Geuens, De Houwer, & De Pelsmacker, 2005). This inconsistency might be due to the fact that verbal endorsement is determined primarily by whether the products are ecologically friendly, whereas the actual purchase of those products is determined primarily by the visual attractiveness of those products (e.g., ecologically friendly but unfashionable clothes). Similarly, evaluative ratings might fail to predict other kinds of evaluative responses because the context is different during their registration. Returning to the example of ecologically friendly products, verbal endorsements might fail to predict purchase behaviour if the context includes cues related to ecologically friendly behaviour while giving the verbal endorsement, but such cues might be absent when people make a purchasing decision (e.g., Schuldt, Konrath, & Schwarz, 2011). Hence progress in predicting particular instances of evaluation (e.g., purchasing behaviour) on the basis of other instances of evaluation (e.g., evaluative ratings) can be achieved by carefully

analysing the elements in the environment that influence the two kinds of evaluative responses (see Kaiser et al., 2010). We contend, therefore, that important advances may be made by clearly distinguishing the functional and cognitive levels and by engaging in sophisticated functional analyses of the different instances of evaluation (i.e., analyses of what it is in the environment that influences the evaluative responses).

Expanding on observed covariations between different instances of evaluation, the central task for cognitive attitude theories can be conceived of as the development of mental explanations of why certain types of evaluative responses to a given object covary while others do not. Ideally these theories should also generate novel predictions that lead to the discovery of previously undetected covariations (or a previously undetected lack of covariation) between instances of evaluation that involve different evaluative responses (e.g., Dunton & Fazio, 1997). Within our functional-cognitive approach, the nature of the relation between attitudes and behaviour (or lack thereof) is not something that needs to be explained (i.e., explanandum), but constitutes a cognitive hypothesis that is supposed to explain observed covariations between different instances of evaluation (i.e., explanans). For example, discrepancies between evaluative ratings of environmentally friendly products and the actual purchase of these products represent a behavioural event that needs to be explained, and this event may be explained by the hypothesis that evaluative ratings are mostly influenced by attitudes, whereas the impact of attitudes on purchasing decisions depends on other factors that are unrelated to attitudes. However, in hypothesising a lacking influence of attitudes on behaviour as the explanans for discrepancies between different instances of evaluation, it is critical that the proposed explanation provides novel predictions about potential moderators of the observed discrepancy. Otherwise, such explanations involve the risk of explanatory circularity, in that discrepancies between two behaviours are interpreted as evidence for the ineffectiveness of attitudes in influencing one of these behaviours, but the only evidence for this assumption is the lacking covariation that needs to be explained in the first place. Thus a clear distinction between the functional and the cognitive levels of analysis, including their explanans and explanandum, can resolve the “catch-22” problem that is created by the use of behavioural proxies for mental attitudes.

The representation of attitudes: All possible moderators of evaluation

The last three sections have focused on the mental processes that attitude theories aim to describe: those involved in the acquisition of attitudes, the activation of attitudes, and the impact of attitudes on behaviour. However, a mental explanation is complete only if it also specifies the nature of the representations on which these processes operate. In fact, assumptions about the nature of mental

representation impose important constraints on the nature of the mental processes that can operate on these representations (Smith & Queller, 2001). Because of these constraints, attitude researchers have rightfully devoted a lot of attention to the question of how attitudes are represented (for an overview, see Gawronski, 2007).

Unfortunately it is very difficult, if not impossible, to arrive at definite conclusions regarding the exact nature of mental representations (Wyer, 2007). Within the functional-cognitive framework, we can assume that there must be some kind of attitude towards a given stimulus if the presence of that stimulus causes an evaluative response. However, the mere occurrence of evaluation does not allow one to draw strong conclusions about *how* the attitude is represented (Greenwald & Nosek, 2009). The same is true for moderating effects on evaluation. Because mental representations always influence evaluation by virtue of mental processes, one can never be sure that a given moderator has an effect because attitudes are represented in a certain manner or because the mental processes that operate on attitudes have certain properties (Wyer, 2007). To illustrate this point, consider the longstanding debate in the attitude literature between proponents of the idea that attitudes are abstract summary representations that simply need to be retrieved from memory (e.g., Fazio, 2007), and researchers who argue that attitudes are constructed on the spot from concrete memories each time that an attitude object is encountered (e.g., Schwarz, 2007). The fact that evaluation often depends on the characteristics of the context seems to argue for the latter position, but is also compatible with the former position if it is assumed that contexts modulate which abstract representation is activated in response to a given stimulus (e.g., Fazio, 2007). This controversy resembles the debate between abstractive and exemplar-based representations in the cognitive literature, which some have argued is impossible to resolve (Barsalou, 1990).

We argue that these epistemological problems should not stop researchers from speculating about the nature of the mental representations that underlie evaluation. After all, hypotheses about mental representation are a central component of cognitive attitude theories. However, the framework that we have put forward in this chapter entails a somewhat different perspective on debates about mental representations. First, according to the functional-cognitive approach, the primary input for these debates is functional knowledge about the moderators of evaluation. In order for this knowledge to have a maximal impact on theories about attitude representation, moderating relations should be formulated as much as possible in terms that do not refer to mental processes or representations. A major contribution of our framework is to propose that these formulations can be achieved in a systematic and coherent manner by defining behavioural phenomena in terms of the moderators of evaluation. Second, although well-formulated hypotheses about the nature of mental representations have heuristic and predictive value, researchers should realise that it will be difficult to arrive at definite conclusions about the exact nature of the mental representations that mediate

evaluation (e.g., Greenwald & Nosek, 2009). As outlined by prominent theorists in this domain, empirical findings that have been inspired by one theory of mental representation can almost always be re-described in terms of alternative theories (e.g., Wyer, 2007), which suggests that competing theories of mental representation rarely differ in terms of their utility in organising existing knowledge (i.e., their heuristic value). Thus, in our view, the more important criterion in the evaluation of competing theories of mental representation is their utility in inspiring research that leads to new empirical discoveries (i.e., their predictive value).

Third, and directly related to the last point, many theories of mental representation are rather abstract with a relatively low level of precision, which makes it difficult to determine which behavioural observations are consistent or inconsistent with these theories (for a discussion, see Gawronski & Bodenhausen, in press). The result is that these theories can explain almost every empirical finding in a post-hoc fashion (i.e., high heuristic value). Yet, they seem less suitable for making a priori predictions that could challenge these theories in the case of disconfirmation (i.e., low predictive value). As noted above, theories of mental representation should not only organise functional knowledge about evaluation, but also generate new functional knowledge by virtue of novel predictions about functional relations between stimuli and evaluative responses. An important implication of our functional-cognitive framework therefore is the call for conceptual precision at both levels of explanation, which will be helpful in maximising the predictive value of theories of attitude representation.

POSSIBLE QUESTIONS ABOUT THE FUNCTIONAL-COGNITIVE FRAMEWORK

Our functional-cognitive framework provides a meta-theoretical conceptualisation of attitude research that resolves several problems resulting from the use of behavioural proxies for mental constructs. Although many of its basic concepts have been directly adopted from traditional attitude research, it also involves several concepts that may be less familiar to some researchers (e.g., the notion of functional explanation; see Chiesa, 1992; De Houwer, 2011; De Houwer et al., 2013). Because these concepts require a different way of interpreting empirical data (e.g., interpreting attitudinal phenomena as instances of evaluation and their environmental moderators rather than direct reflections of mental attitudes), we anticipate that these differences will raise a number of questions regarding the premises and implications of our meta-theoretical analysis. In the remainder of this chapter we address several issues to avoid potential misunderstandings about the central arguments and implications of our functional-cognitive framework.

Does the functional-cognitive framework require a change in the way attitude research is done?

The functional-cognitive framework is a meta-theoretical framework that has been developed to avoid theoretical pitfalls in the interpretation of empirical data and the construction and evaluation of attitude theories. Thus adopting the functional-cognitive framework does not require any methodological changes in the way attitude research is conducted. In fact, we consider research on attitudes as one of the most impressive areas in social psychology in terms of its creativity, research design, methodological rigor, and data analysis. Our framework does, however, require a major change in the interpretation of empirical data. The cornerstone of our proposal is to conceptualise attitudinal phenomena in a coherent non-mental, functional way as instances of evaluation.

What is new about the functional-cognitive framework of attitude research?

The framework presented in this chapter builds on our previous work in which we provided arguments for the need to separate behavioural effects and mental processes (e.g., De Houwer, 2007, 2011; Deutsch & Gawronski, 2009; Gawronski et al., 2008). Some of this work already provided non-mental definitions of specific attitudinal phenomena (e.g., EC; see De Houwer, 2007) while other work pointed to the mutually reinforcing nature of functional and cognitive levels of analysis in psychology in general (De Houwer, 2011; Hughes et al., 2011). The framework proposed in the current chapter integrates and extends this earlier work by providing a coherent and comprehensive way of defining attitudinal phenomena in functional terms. More specifically, we introduced the functional concept of evaluation and proposed the idea that a wide variety of attitudinal phenomena can be conceived of as instances of evaluation. Based on this conceptual innovation, we also put forward a functional-cognitive framework for attitude research. This framework provides a novel meta-theoretical perspective on attitude research that helps to avoid the well-known problems of using behavioural proxies for attitudes and attitudinal processes.

Hasn't attitude research been functional-cognitive all along?

One might argue that attitude research has always been functional-cognitive in nature, in that it always focused on mental explanations of behavioural effects. We do indeed believe that any cognitively oriented research necessarily involves a functional level of analysis. However, attitude research has been lacking a coherent and encompassing framework that (1) makes explicit the nature of and relation between the functional and cognitive levels of analysis and (2) does so in a conceptually sophisticated way that encompasses a wide variety of attitudinal

phenomena. In fact, we believe that the absence of such an overarching framework is the primary reason why many attitude researchers continue to treat behavioural effects as proxies for mental constructs. By providing a functional-cognitive framework for attitude research, we offer the conceptual tools to optimise progress at both the functional and cognitive levels of analysis.

Does the functional-cognitive framework endorse radical behaviourism?

The central goal of functionally oriented attitude research is to provide explanations of evaluative responses in terms of stimuli in the environment. This endeavour resembles the notion of radical behaviourism (e.g., Skinner, 1938), in that both focus on stimulus–response relations. Yet radical behaviourism has been rejected by psychologists for several reasons (for a review, see Chiesa, 1994), which raises the question of whether the criticisms of radical behaviourism also apply to the functional-cognitive framework. The short answer to this question is: “No, it does not.” Whereas radical behaviourism rejects cognitive explanations, the functional-cognitive framework not only embraces cognitive explanations, but also illustrates the interdependent relation between functional and cognitive explanations (for more details, see De Houwer, 2011; De Houwer et al., 2013). Functional explanations of evaluative responses provide the empirical input that is necessary to develop and test cognitive explanations of evaluation. Vice versa, progress in cognitive explanations of evaluation can advance functional explanations of evaluative responses by organising existing knowledge about the elements in the environment that determine evaluative responses (i.e., heuristic value) and through the generation of new predictions about elements in the environment that might influence evaluative responses (i.e., predictive value). In fact, an exclusive focus on cognitive explanations that does not consider its empirical roots in functional stimulus–response relations will be susceptible to the problems of behavioural proxies outlined in the beginning of this chapter. Conversely, an exclusive focus on functional explanations that does not consider the mediators of stimulus–response relations lacks any cognitive explanation for how those relations emerge. By integrating both functional and cognitive research in a single framework with two interdependent levels of analysis, the functional-cognitive framework avoids both problems.

Is functional research meant to achieve “objectivity”?

Our suggestion to adopt non-mental, functional definitions of attitudinal phenomena may also raise the question of whether these definitions are meant to achieve “objectivity”. Again, the short answer to this question is: “No, they are not.” Functional definitions are not more or less objective than any other type of definition (Barnes & Roche, 1997; Chiesa, 1992). They are always theory-

laden in the sense that they are shaped by the theories and concepts in the mind of the researcher (e.g., a researcher's classification of different stimuli as belonging to the same category). Nevertheless, it is crucial that functional definitions of attitudinal phenomena do not refer to those mental constructs that are proposed to be in the mind of the participant (e.g., a participant's attitude towards particular stimuli). Undoubtedly, concepts in the mind of the researcher have a significant impact on his or her classification of observations and his or her functional explanations of these observations. For the functional-cognitive framework, it suffices that the behavioural effects shown by a person are not defined in terms of the mental constructs that are proposed to explain these effects. Adopting functional definitions also does not result in "objectivist" attitude research in the sense that it would consider only the "objective" environment rather than a person's subjective construal of the environment. As prominently explained by Ross and Nisbett (1991), the functional notion of *situationism* is perfectly consistent with the cognitive notion of *subjective construal* (see also Reis, 2008). Because each individual has a unique learning history, the same stimulus can have dramatically different effects for different individuals. Similarly, the same stimulus may have dramatically different effects for the same individual if the stimulus is encountered in different contexts. In fact, one of the main goals of attitude research, as conceptualised in functional-cognitive terms, is precisely to describe and explain the way in which individual learning histories and stimulus contexts moderate evaluation.

Does the functional-cognitive framework simply re-label existing constructs?

Some readers may wonder if the functional-cognitive framework merely involves a re-labelling of existing constructs. Again the answer is: "No, it does not." In our framework it is essential that each construct is defined either at the functional level or at the cognitive level. For instance, the term *evaluation* refers to a strictly functional concept whereas the term *attitude* refers to a strictly cognitive concept. Ideas about attitudes can be used to explain why evaluation occurs under certain conditions and knowledge about the moderators of evaluation constrains ideas about attitudes, but the two should never overlap in an a priori manner. By insisting on this conceptual distinction, our framework helps to resolve the problems that arise from the use of behavioural proxies for mental constructs.

Can we investigate cognitive mediation with statistical mediation analysis?

If cognitive attitude theories are conceived of as propositions about the mediators of evaluation, one might wonder whether these propositions could also be

tested by means of statistical mediation analysis, such as Baron and Kenny's (1986) multiple regression approach or Preacher and Hayes' (2004) bootstrapping method. The common idea underlying these procedures is that causal chains can be established by statistically controlling for a potential mediator (e.g., mental process) when testing the effect of a given distal factor (e.g., environmental stimuli). Although this approach has become one of the most important tools to test mediation in social psychology, it requires that the proposed mediator can be measured directly. Because we reject the use of behaviour as a proxy for mental constructs, statistical mediation analysis does not provide a viable means to assess the merit of cognitive attitude theories. Instead, their merit has to be evaluated on the basis of their heuristic value (i.e., the extent to which they account for existing knowledge about the moderators of evaluation) and their predictive value (i.e., the extent to which they correctly predict previously unknown effects of moderators). In other words, theories about the cognitive *mediators* of evaluation have to be evaluated on the basis of their implications regarding the environmental *moderators* of evaluation (see Jacoby & Sassenberg, 2011). These conclusions resonate with earlier concerns by Spencer, Zanna, and Fong (2005) who suggested that experimental manipulations are often superior in establishing a causal chain compared with traditional regression-based approaches. In fact, the only exceptions for which they deemed traditional regression-based approaches as superior are cases in which the measurement of the proposed mediator is easy and manipulation of that mediator is difficult. Because the measurement of mental constructs requires a treatment of behaviour as a proxy for these constructs, Spencer et al. (2005) conclusion implies that experimental manipulations are generally superior in testing hypotheses about the mediating role of mental constructs. Nevertheless, it is worth noting that regression-based approaches are well-suited for cases in which a behaviour itself (e.g., relative frequency of drinking red versus white wine) is assumed to serve as the mediator for another behaviour (e.g., self-reported preference for red over white wine), as is the case in self-perception (Bem, 1967).

Can neuroscience solve the problem of behavioural proxies?

A major development in research on attitudes is the growing interest in their neurological underpinnings (Ito & Cacioppo, 2007). From the perspective of the functional-cognitive framework, there are two ways in which neuroscientific research on attitudes can be seen as part of the study of evaluation. First, neurological responses (e.g., changes in electrical activity as captured by EEG; changes in bloodflow as captured by fMRI) can be conceptualised as a particular type of evaluative response next to other types such as verbal or motor responses (see De Houwer & Moors, 2010). From this perspective, one can examine whether and when a given stimulus leads to certain kinds of

neurological evaluative responses. Second, environmental influences on the state of a participant's brain are an important aspect of the broader stimulus context, because the brain is assumed to regulate most behaviour. It is therefore interesting to examine whether evaluation is moderated by environmental influences on the neurological state of the participant (e.g., lesions in particular brain areas; transcranial magnetic stimulation of particular brain areas). Thus, from a functional-cognitive perspective, neuroscience can provide both dependent and independent variables for studies on evaluation. It can provide dependent variables in the form of novel kinds of evaluative responses. In addition, it can provide independent variables in the form of environmental factors that may moderate evaluation. In this manner neuroscience can contribute to the development of mental theories of evaluation because it generates new functional knowledge that can be used to constrain mental explanations of evaluation. Nevertheless, it is important to note that neuroscientific data do not provide direct access to the mental constructs proposed by cognitive attitude theories (see also Poldrack, 2006). The meaning of neuroscientific data generally depends on their relation to behaviour, which places them at the level of functional relations between stimuli and evaluative responses.

CONCLUDING COMMENTS

We started this chapter by reviewing the problems that arise from using behaviour as a proxy for attitudes and attitudinal processes. In search of an alternative for this practice, we explored the idea that attitude research can be conceptualised as the study of the moderators and mediators of evaluation. Drawing on a functional definition of evaluation as the effect of stimuli on evaluative responses, we argued that attitude research provides answers to two questions: (1) Which elements in the environment moderate evaluation? (2) What mental processes and representations mediate evaluation? These functional and cognitive levels of analysis do not compete but mutually support each other. Although many of the reviewed problems inherent to the use of behavioural proxies are well-known to attitude researchers, the treatment of behavioural effects as direct indices of attitudes and attitudinal processes is still rather common. A potential reason for this paradox is the lack of an overarching, phenomenon-independent framework that provides an alternative to the use of behaviours as proxies for mental constructs. The current chapter aimed to fill this gap. We hope that our analysis revealed the benefits of the functional-cognitive framework for attitude research, and in this way will contribute to progress in this important domain of research.

REFERENCES

- Ajzen, I. (1991). The theory of planned behavior. *Organisational Behavior and Human Decision Processes*, *50*, 179–211. doi:10.1016/0749-5978(91)90020-T
- Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological bulletin*, *84*, 888–918. doi:10.1037/0033-2909.84.5.888
- Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 173–221). Mahwah, NJ: Erlbaum.
- Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.), *Handbook of social psychology* (pp. 798–844). Worcester, MA: Clark University Press.
- Arkes, H. R., & Tetlock, P. E. (2004). Target article: Attributions of implicit prejudice, or “would Jesse Jackson ‘Fail’ the implicit association test?”. *Psychological inquiry*, *15*, 257–278. doi:10.1207/s15327965pli1504_01
- Baddeley, A. (2010). Working memory. *Current biology*, *20*, R136–R140. doi:10.1016/j.cub.2009.12.014
- Barden, J., Maddux, W. W., Petty, R. E., & Brewer, M. B. (2004). Contextual moderation of racial bias: The impact of social roles on controlled and automatically activated attitudes. *Journal of Personality and Social Psychology*, *87*, 5–22. doi:10.1037/0022-3514.87.1.5
- 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. doi:10.2307/1423027
- Barnes, D., & Roche, D. (1997). A behavior-analytic approach to behavioral reflexivity. *The psychological record*, *47*, 543–572.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182. doi:10.1037/0022-3514.51.6.1173
- Barsalou, L. W. (1990). On the indistinguishability of exemplar memory and abstraction in category representation. In T. K. Srull & R. S. Wyer (Eds.), *Advances in social cognition* (Vol. 3, pp. 61–88). Hillsdale, NJ: Lawrence Erlbaum.
- Bechtel, W. (2005). The challenge of characterizing operations in the mechanisms underlying behavior. *Journal of the Experimental Analysis of Behavior*, *84*, 313–325. doi:10.1901/jeab.2005.103-04
- Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, *74*, 183–200. doi:10.1037/h0024835
- Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin*, *106*, 265–289. doi:10.1037/0033-2909.106.2.265
- Brehm, J. W. (1956). Postdecision changes in the desirability of alternatives. *The Journal of Abnormal and Social Psychology*, *52*, 384–389. doi:10.1037/h0041006
- Chen, M. K., & Risen, J. L. (2010). How choice affects and reflects preferences: Revisiting the free-choice paradigm. *Journal of Personality and Social Psychology*, *99*, 573–594. doi:10.1037/a0020217
- Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*, 367–405. doi:10.1037/0033-295X.104.2.367
- Chiesa, M. (1992). Radical behaviorism and scientific frameworks: From mechanistic to relational accounts. *American Psychologist*, *47*, 1287–1299. doi:10.1037/0003-066X.47.11.1287
- Chiesa, M. (1994). *Radical behaviorism: The philosophy and the science*. Boston, MA: Authors’ Cooperative.
- Conrey, F. R., & Smith, E. R. (2007). Attitude representation: Attitudes as patterns in a distributed, connectionist representational system. *Social Cognition*, *25*, 718–735. doi:10.1521/soco.2007.25.5.718

- Crano, W. D., & Prislin, R. (2006). Attitudes and persuasion. *Annual Review of Psychology*, *57*, 345–374. doi:10.1146/annurev.psych.57.102904.190034
- De Houwer, J. (2003). A structural analysis of indirect measures of attitudes. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 219–244). Mahwah, NJ: Lawrence Erlbaum.
- De Houwer, J. (2007). A conceptual and theoretical analysis of evaluative conditioning. *The Spanish Journal of Psychology*, *10*, 230–241.
- De Houwer, J. (2009a). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior*, *37*, 1–20. doi:10.3758/LB.37.1.1
- De Houwer, J. (2009b). How do people evaluate objects? A brief review. *Social and Personality Psychology Compass*, *3*, 36–48. doi:10.1111/j.1751-9004.2008.00162.x
- 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. doi:10.1177/1745691611400238
- De Houwer, J. (in press). A propositional model of implicit evaluation. *Social Psychology and Personality Compass*.
- 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. doi: 10.3758/s13423-013-0386-3
- De Houwer, J., & Moors, A. (2010). Implicit measures: Similarities and differences. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 176–193). New York, NY: Guilford Press.
- De Houwer, J., & Moors, A. (2013). How to define and examine implicit processes?. In R. Proctor & J. Capaldi (Eds.), *Implicit and explicit processes in the psychology of science* (pp.183–198). New York, NY: Oxford University Press.
- De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, *135*, 347–368. doi:10.1037/a0014211
- 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. doi:10.1037/0033-2909.127.6.853
- Deutsch, R., & Gawronski, B. (2009). When the method makes a difference: Antagonistic effects on “automatic evaluations” as a function of task characteristics of the measure. *Journal of Experimental Social Psychology*, *45*, 101–114. doi:10.1016/j.jesp.2008.09.001
- Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin*, *23*, 316–326. doi:10.1177/0146167297233009
- Eagly, A. H., & Chaiken, S. (2007). The advantages of an inclusive definition of attitude. *Social Cognition*, *25*, 582–602. doi:10.1521/soco.2007.25.5.582
- Ebert, I. D., Steffens, M. C., von Stülpnagel, R., & Jelenec, P. (2009). How to like yourself better, or chocolate less: Changing implicit attitudes with one IAT task. *Journal of Experimental Social Psychology*, *45*, 1098–1104. doi:10.1016/j.jesp.2009.06.008
- Egan, L. C., Santos, L. R., & Bloom, P. (2007). The origins of cognitive dissonance: Evidence from children and monkeys. *Psychological Science*, *18*, 978–983. doi:10.1111/j.1467-9280.2007.02012.x
- Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. *Advances in Experimental Social Psychology*, *23*, 75–109. doi:10.1016/S0065-2601(08)60318-4
- Fazio, R. H. (2007). Attitudes as object-evaluation associations of varying strength. *Social Cognition*, *25*, 603–637. doi:10.1521/soco.2007.25.5.603
- 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. doi:10.1037/0022-3514.69.6.1013

- 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. doi:10.1037/0022-3514.50.2.229
- Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row Peterson.
- Gardner, H. (1987). *The mind's new science: A history of the cognitive revolution*. New York: Basic Books.
- Gawronski, B. (Ed.). (2007). What is an attitude? [Special issue]. *Social Cognition*, *25*(5).
- 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. doi:10.1037/0033-2909.132.5.692
- Gawronski, B., & Bodenhausen, G. V. (2011). The associative-propositional evaluation model: Theory, evidence, and open questions. *Advances in Experimental Social Psychology*, *44*, 59–127. doi:10.1016/B978-0-12-385522-0.00002-0
- Gawronski, B., & Bodenhausen, G. V. (in press). Theory evaluation. In B. Gawronski & G. V. Bodenhausen (Eds.), *Theory and explanation in social psychology*. New York, NY: Guilford Press.
- Gawronski, B., Bodenhausen, G. V., & Becker, A. P. (2007). I like it, because I like myself: Associative self-anchoring and post-decisional change of implicit evaluations. *Journal of Experimental Social Psychology*, *43*, 221–232. doi:10.1016/j.jesp.2006.04.001
- Gawronski, B., Cunningham, W. A., LeBel, E. P., & Deutsch, R. (2010). Attentional influences on affective priming: Does categorisation influence spontaneous evaluations of multiply categorisable objects?. *Cognition and Emotion*, *24*, 1008–1025. doi:10.1080/02699930903112712
- Gawronski, B., Deutsch, R., LeBel, E. P., & Peters, K. R. (2008). Response interference as a mechanism underlying implicit measures: Some traps and gaps in the assessment of mental associations with experimental paradigms. *European Journal of Psychological Assessment*, *24*, 218–225. doi:10.1027/1015-5759.24.4.218
- Gawronski, B., & LeBel, E. P. (2008). Understanding patterns of attitude change: When implicit measures show change, but explicit measures do not. *Journal of Experimental Social Psychology*, *44*, 1355–1361. doi:10.1016/j.jesp.2008.04.005
- Gawronski, B., Rydell, R. J., Vervliet, B., & De Houwer, J. (2010). Generalisation versus contextualisation in automatic evaluation. *Journal of Experimental Psychology: General*, *139*, 683–701. doi:10.1037/a0020315
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27. doi:10.1037/0033-295X.102.1.4
- Greenwald, A. G., & Nosek, B. A. (2009). Attitudinal dissociation: What does it mean?. In R. E. Petty, R. H. Fazio, & P. Brinol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 65–82). Hillsdale, NJ: Erlbaum.
- Hayes, S. C. (1995). Why cognitions are not causes. *The Behavior Therapist*, *18*, 59–60.
- Hayes, S. C., Barnes-Holmes, D., & Wilson, K. G. (2012). Contextual behavioral science: Creating a science more adequate to the challenge of the human condition. *Journal of Contextual Behavioral Science*, *1*, 1–16. doi:10.1016/j.jcbs.2012.09.004
- Hempel, C. G. (1970). *Aspects of scientific explanation and other essays in the philosophy of science*. New York: Free Press.
- Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin*, *136*, 390–421. doi:10.1037/a0018916
- Hughes, S., Barnes-Holmes, D., & De Houwer, J. (2011). The dominance of associative theorising in implicit attitude research: Propositional and behavioral alternatives. *The Psychological Record*, *61*, 465–498.
- Hume, D. (1987). *A treatise of human nature* (2nd ed.). Oxford, UK: Clarendon Press. (Original work published 1739)

- Ito, T. A., & Cacioppo, J. T. (2007). Attitudes and mental and neural states of readiness: Using physiological measures to study implicit attitudes. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes* (pp. 125–158). New York: Guilford Press.
- Jacoby, J., & Sassenberg, K. (2011). Interactions do not only tell us when, but can also tell us how: Testing process hypotheses by interaction. *European Journal of Social Psychology, 41*, 180–190. doi:10.1002/ejsp.762
- Jones, C. R., Fazio, R. H., & Olson, M. A. (2009). Implicit misattribution as a mechanism underlying evaluative conditioning. *Journal of Personality and Social Psychology, 96*, 933–948. doi:10.1037/a0014747
- Kaiser, F. G., Byrka, K., & Hartig, T. (2010). Reviving Campbell's paradigm for attitude research. *Personality and Social Psychology Review, 14*, 351–367. doi:10.1177/1088868310366452
- Kant, I. (1965). *Critique of pure reason*. London: Macmillan. (Original work published 1781)
- Krosnick, J. A., Judd, C. M., & Wittenbrink, B. (2005). The measurement of attitudes. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 21–76). Mahwah, NJ: Erlbaum.
- Maio, G. R. & Olson, J. M. (Eds.). (2000). *Why we evaluate: Functions of attitudes*. Mahwah, NJ: Erlbaum.
- Lieberman, M. D., Ochsner, K. N., Gilbert, D. T., & Schacter, D. L. (2001). Do amnesics exhibit cognitive dissonance reduction? The role of explicit memory and attention in attitude change. *Psychological Science, 12*, 135–140. doi:10.1111/1467-9280.00323
- Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences, 32*, 183–198. doi:10.1017/S0140525X09000855
- Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin, 132*, 297–326. doi:10.1037/0033-2909.132.2.297
- Olson, J. M., Vernon, P. A., Harris, J. A., & Jang, K. (2001). The heritability of attitudes: A study of twins. *Journal of Personality and Social Psychology, 80*, 845–860. doi:10.1037/0022-3514.80.6.845
- Ottati, V. C., Coats, S., Mae, L., DeCoster, J., & Smith, E. R. (2010). Implicit evaluation of familiar and novel concepts presented at low levels of conscious detectability. *The American Journal of Psychology, 123*, 15–27.
- Perugini, M., Richetin, J., & Zogmaister, C. (2012). The formation of implicit and explicit attitudes for neutral and valenced stimuli using the self. *Learning & Motivation, 43*, 135–143. doi:10.1016/j.lmot.2012.02.002
- Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology, 90*, 751–783. doi:10.1037/0022-3514.90.5.751
- Petty, R. E., & Briñol, P. (2010). Attitude structure and change: Implications for implicit measures. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 335–352). New York: Guilford Press.
- Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of persuasion. *Advances in Experimental Social Psychology, 19*, 123–205. doi:10.1016/S0065-2601(08)60214-2
- Pleyers, G., Corneille, O., Yzerbyt, V., & Luminet, O. (2009). Evaluative conditioning may incur attentional costs. *Journal of Experimental Psychology: Animal Behavior Processes, 35*, 279–285.
- Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data?. *Trends in Cognitive Sciences, 10*, 59–63. doi:10.1016/j.tics.2005.12.004
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers, 36*, 717–731. doi:10.3758/BF03206553
- Prestwich, A., Perugini, M., Hurling, R., & Richetin, J. (2010). Using the self to change implicit attitudes. *European Journal of Social Psychology, 40*, 61–71. doi:10.1002/ejsp.610
- Quine, W. V. O., & Ullian, J. S. (1978). *The web of belief* (2nd ed.). New York: McGraw-Hill.

- Ranganath, K. A., Smith, C. T., & Nosek, B. A. (2008). Distinguishing automatic and controlled components of attitudes from direct and indirect measurement methods. *Journal of Experimental Social Psychology*, *44*, 386–396. doi:10.1016/j.jesp.2006.12.008
- Reis, H. T. (2008). Reinvigorating the concept of situation in social psychology. *Personality and Social Psychology Review*, *12*, 311–329. doi:10.1177/1088868308321721
- Ross, L., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology*. New York: McGraw Hill.
- Schuldt, J., Konrath, S., & Schwarz, N. (2011). “Global warming” or “climate change”? Whether the planet is warming depends on question wording. *Public Opinion Quarterly*, *75*, 115–124. doi:10.1093/poq/nfq073
- Schwarz, N. (1990). Feelings as information: Informational functions of affective states. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Volume 2, pp. 527–561). New York: Guilford Press.
- Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition*, *25*, 638–656. doi:10.1521/soco.2007.25.5.638
- Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*, 513–523. doi:10.1037/0022-3514.45.3.513
- Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century.
- Smith, E. R., & Queller, S. (2001). Mental representations. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of social psychology: Intraindividual processes* (pp. 111–133). Oxford, UK: Blackwell.
- Sosa, E. & Tooley, M. (Eds.). (1993). *Causation*. Oxford, UK: Oxford University Press.
- Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: Why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology*, *89*, 845–851. doi:10.1037/0022-3514.89.6.845
- Sweldens, S., Van Osselaer, S., & Janiszewski, C. (2010). Evaluative conditioning procedures and the resilience of conditioned brand attitudes. *Journal of Consumer Research*, *37*, 473–489. doi:10.1086/653656
- Teige-Mocigemba, S., & Klauer, K. C. (2013). On the controllability of evaluative-priming effects: Some limits that are none. *Cognition & Emotion*, *27*, 632–657. doi:10.1080/02699931.2012.732041
- Vantomme, D., Geuens, M., De Houwer, J., & De Pelsmacker, P. (2005). Implicit attitudes toward green consumer behaviour. *Psychologica Belgica*, *45*, 217–239. doi:10.5334/pb-45-4-217
- Wilson, T. D., & Dunn, E. W. (2004). Self-knowledge: Its limits, value, and potential for improvement. *Annual Review of Psychology*, *55*, 493–518. doi:10.1146/annurev.psych.55.090902.141954
- Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, *107*, 101–126. doi:10.1037/0033-295X.107.1.101
- Wyer, R. S. (2007). Principles of mental representation. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (2nd ed, pp. 285–307). New York: Guilford Press.
- Zanon, R., De Houwer, J., & Gast, A. (2012). Context effects in evaluative conditioning of implicit evaluations. *Learning & Motivation*, *43*, 155–165. doi:10.1016/j.lmot.2012.02.003