

RUNNING HEAD: EC as a Symbolic Phenomenon

**Evaluative conditioning as a symbolic phenomenon:
On the relation between evaluative conditioning, evaluative conditioning via instructions,
and persuasion**

Jan De Houwer & Sean Hughes
Ghent University, Ghent, Belgium

In press. Social Cognition.

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

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

Abstract

Evaluative conditioning (EC) is sometimes portrayed as a primitive way of changing attitudes that is fundamentally different from persuasion via arguments. We provide a new perspective on the nature of EC and its relation to persuasion by exploring the idea that stimulus pairings can function as a symbol that conveys the nature of the relation between stimuli. We put forward the concept of symbolic EC to refer to changes in liking that occur because stimulus pairings function as symbols. The idea of symbolic EC is consistent with at least some current theories of persuasion. It clarifies what EC research can add to the understanding of the origins of our preferences and has implications for how (symbolic and non-symbolic) EC can be established, the boundaries of EC research, and cognitive and functional models of EC.

Keywords: Evaluative conditioning, persuasion, instructions

Evaluative conditioning as a symbolic phenomenon:**On the relation between evaluative conditioning, evaluative conditioning via instructions, and persuasion**

For many years now, researchers have been intrigued by the impact that stimulus pairings seem to have on liking (Martin & Levy, 1978; Razran, 1954; see De Houwer, Thomas, & Baeyens, 2001, and Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010, for reviews). This phenomenon, which is typically referred to as evaluative conditioning (EC), promises to provide a unique contribution to evaluative learning research, that is, research on how likes and dislikes are acquired and can be changed. Acquiring knowledge about the origins of preferences is considered to be a core mission for psychology as a whole because preferences are assumed to guide many aspects of human behavior.

Within the realm of evaluative learning research, EC is often seen as fundamentally different from persuasion via arguments. Whereas the latter requires effortful processing of words and sentences, EC is assumed to be a “primitive means of changing attitudes” (Briñol, Petty, & McCaslin, 2009, p. 287) that requires little cognitive effort and thought. In this paper, we offer a different perspective on EC that highlights the symbolic nature of this phenomenon. More specifically, we argue that the spatio-temporal pairing of stimuli can function as a symbol indicating how those stimuli are related, much like words and sentences can. Such instances of EC are similar to persuasion via arguments in that both phenomena depend on the ability to use symbols. They are different primarily with regard to the proximal event that functions as a symbol (i.e., stimulus pairings or a collections of words or sentences).

In the second part of this paper, we discuss a number of implications of this point of view. First, it reveals that EC research can shed light on a unique symbolic way of influencing liking (i.e., via stimulus pairings). Second, we discuss the implications for how the presence of

(symbolic and non-symbolic) EC can be established. Third, we explore the boundaries of EC research. Finally, we examine the implications for cognitive and functional accounts of EC. In sum, the central message of this paper is that EC is a symbolic phenomenon that shares important properties with other symbolic phenomena such as persuasion via verbal arguments but that also has unique features and thus offers a unique way to change preferences.

Communalities between EC, EC via Instructions, and Persuasion

EC via Instructions

We start our analysis with a recently documented phenomenon known as EC via instructions. It refers to the fact that merely informing participants about a relation between a neutral stimulus and a valenced stimulus is sufficient to change liking, even when those stimuli are never actually presented together. For instance, after reading the sentences “UDIBNON will be followed by positive pictures” and “BAYRAM will be followed by negative pictures”, participants like UDIBNON more than BAYRAM, irrespective of whether liking is assessed using direct measurement procedures (e.g., evaluative ratings) or indirect measurement procedures (e.g., the Implicit Association Test; see De Houwer, 2006; Gast & De Houwer, 2013).

As noted by Gast, Gawronski, and De Houwer (2012), there has been a debate about whether EC via instructions qualifies as EC, that is, as changes in liking that are due to stimulus pairings (De Houwer, 2007). Some have argued that it does (e.g., Gast et al., 2012; see Field, 2006, for a similar argument in the context of fear conditioning via instructions). It seems indisputable that the presentation of instructions is a spatio-temporal event: it involves the presence of physical stimuli such as letters or sounds in a particular space and a particular

period of time.¹ Because instructions typically involve multiple stimuli, one could even say that instructions involve a pairing of stimuli, be it that each pair of stimuli typically occurs only once in a single instruction (see De Houwer, Barnes-Holmes, & Moors, 2013, for a discussion of what constitutes a pairing or, more generally, a regularity in the environment). Structurally speaking, when the instruction involves both neutral and valenced stimuli (e.g., “UDIBNON will be followed by positive stimuli”), one could argue that there is a pairing of a neutral (i.e., UDIBNON) and a valenced stimulus (e.g., “positive”) that could in principle lead to a subsequent change in how the neutral stimulus is evaluated. Hence, one could argue that EC via instructions is an instance of EC.

On the other hand, it seems overly simplistic to argue that instructions have an effect because of their *mere* spatio-temporal properties, that is, merely because it consists of stimuli that co-occur in space and time (also see De Houwer et al., 2013, pp. 639-640). In all likelihood, instructions have an impact on liking because of the symbolic meaning of those instructions, more specifically, what the instructions imply about the nature of the relation between the stimuli. For instance, it seems reasonable to assume that an instruction such as “UDIBNON will be followed by positive pictures” has an effect on liking not simply because “UDIBNON” and “positive” co-occur in the same sentence but because, for an English speaker, the instruction specifies that “UDIBNON” and “positive” are related in a specific manner (i.e., the UDIBNON will precede the presentation of positive pictures). Hence, one could argue that the name “EC via instructions” is inappropriate in that the change in liking does not qualify as an instance of EC.² More importantly for the present purposes, these

¹ Hence, “learning via instructions” and “learning via direct experience” should not be contrasted. Receiving instructions is a physical event that is experienced directly.

² One might argue that the role of symbolic meaning in EC via instructions is restricted to the interpretation of individual words. Understanding the symbolic meaning of a word (i.e., to understand what the word stands for) merely alters the nature of the individual stimulus (e.g., it changes the stimulus “positive” from a string of letters

arguments imply that EC via instructions is an essentially symbolic phenomenon. Participants derive symbolic meaning from a particular proximal event (i.e., an event that occurs here and now in space and time), being the words and sentences of which the instructions are composed. They can do so because they have the capacity to use symbols.

EC as a Symbolic Phenomenon

Now that we have established that EC via instructions is a symbolic phenomenon, we examine whether standard EC might also qualify as a symbolic phenomenon. When examining this question, it is important to realize that once people have learned to use symbols, they can apply this ability not only to proximal stimuli such as words or sentences but also to stimulus pairings (also see Gast et al., 2012). Virtually any proximal event can function as a symbol, that is, something that stands for something else. For instance, in sign language, gestures function in much the same way as words or phrases. Under the right conditions, even a simple wink of the eye can convey as much symbolic meaning as an entire sentence. In more technical terms, a range of proximal events can function as symbols on the basis of which meaning is constructed. The fact that humans use symbols in such a flexible and ubiquitous manner has led to the proposal that humans are essentially symbolic beings (Deacon, 1997).

In line with this idea, we argue that the pairing of stimuli is one of those proximal events that can convey symbolic meaning. More specifically, we propose that the pairing of stimuli can change liking because it symbolizes that the stimuli are related in a certain way. For

into a positive stimulus). Hence, EC via instruction might still be regarded as the effect of the mere proximal pairing of individual stimuli, be it symbolically interpreted stimuli. Although we do not exclude this possibility on an *a priori* basis, it is important to realize that an instruction goes beyond the meaning of the individual words in that it involves some type of relation between the concepts referred to by the individual words. In fact, if such an overarching meaning is absent, the instruction would merely be a list of unrelated words that would not constitute what is commonly labeled an instruction. Note that we do not claim that the change in valence must be a logical consequence of the instruction. For instance, there is no logical reason that UDIBNON is good because it precedes positive pictures. In that sense, the change in liking is irrational. Nevertheless, regardless of whether the change in liking is rational or irrational, it is likely to depend on the symbolic meaning of the instruction as a whole rather than on the symbolic meaning of the individual words as such.

instance, the spatio-temporal co-occurrence of a neutral stimulus and a valenced stimulus might be taken to indicate that the two stimuli are similar. In this case, stimulus pairings function in much the same way as the expression “is similar to” in the instruction “A is similar to B”, that is, they symbolize that the two stimuli share certain properties. As is the case with words and other proximal events that function as symbols, the meaning that is conveyed by stimulus pairings is not entirely fixed. Depending on the context, stimulus pairings might even symbolize different types of relations (e.g., similarity, opposition, hierarchy). We refer to an instance of EC as symbolic EC when the change in liking is due to the fact that the stimulus pairings function as a symbol for the relation between the stimuli.³

Many before us have pointed out that human and nonhuman animals often (if not always) respond on the basis of how they construct their environment (i.e., the meaning that they attribute to their environment) rather than to the physical environment itself (e.g., Bruner & Goodman, 1947). Our claims fit within this general constructivist approach but are novel in that they focus on a specific type of meaning construction (i.e., symbolic meaning construction) that is applied to a specific type of proximal event (i.e., stimulus pairings). Note that meaning construction can also be non-symbolic, for instance, when stimuli function as signals. Signals differ from symbols in that they do not stand for something else but merely announce future events or trigger future actions in a dynamic or causal manner (Levi, 2007). Interestingly, it has been argued that only humans can use symbols (e.g., Deacon, 1997). This would imply that only humans who have developed the ability to use symbols can display symbolic EC. Also note that we do not commit to any specific theory about the development of,

³ One could argue the EC qualifies as symbolic also when the individual stimuli function as symbols (e.g., words that denote neutral or valenced concepts) but the spatio-temporal pairing does not. In this paper, however, we aim to highlight the fact that stimulus pairings can function as a symbol and to examine the implications of this idea. Therefore, we restrict our definition of symbolic EC to those instances of EC in which stimulus pairings function as symbols.

or mental processes underlying, the ability to use symbols. For instance, our arguments do not depend on whether this ability is mediated by propositional representations. Our arguments are thus situated at a very abstract level of analysis and do not have to be tied down to existing theoretical debates. By operating at a high level of abstraction, we hope to obtain new insights and reach consensus that might not be achievable when operating at the level of specific theories about learning and liking. We will return to this issue near the end of the paper.

Symbolic EC and Persuasion

In those cases where stimulus pairings function as symbols, EC has a lot in common with persuasion via arguments. Both phenomena depend on the use of symbols. Although the two phenomena involve different types of proximal events as symbols (stimulus pairings versus words or sentences), they are both symbolic in nature. This perspective contrasts with the widespread view that EC is fundamentally different from persuasion via arguments. For instance, within the Elaboration-Likelihood Model (e.g., Petty & Cacioppo, 1986), EC is seen as a primitive process that is part of the peripheral route of attitude change. This peripheral route is assumed to operate separately from a central route that is responsible for the processing of arguments. In contrast to the Elaboration-Likelihood model, we argue that stimulus pairings function in much the same way as arguments: the effect of both types of proximal events depend on the symbolic meaning that is constructed on the basis of those events.

From the perspective of the Heuristic-Systematic Model (Chaiken, Liberman, & Eagly, 1989), one might argue that stimulus pairings can function as heuristic cues. Such cues can lead to attitude change via “simple rules of inference linking elements of the persuasion setting (heuristic cues) with evaluative inferences” (Bohner, 2008, p. 167). It is not unreasonable to assume that that people have heuristic rules that link stimulus pairings to valence (e.g., “stimuli that co-occur, are similar in valence” or “stimuli that are similar with regard to their

spatio-temporal presence are likely to be similar also in valence”). These heuristic rules are not logically sound or always correct, but in some situations, people might turn to them (either intentionally or unintentionally) as shortcuts for determining the evaluative properties of objects. Provided that the application these heuristic rules (at least sometimes) involves treating the stimulus pairings as a symbolic event, one could argue that the Heuristic-Systematic Model of persuasion is compatible with the idea the stimulus pairings can function as symbols. In fact, we believe that some symbolic instances of EC might well be mediated by the application of these types of heuristic rules.

Finally, the idea of symbolic EC is compatible also with the unimodel of persuasion (e.g., Kruglanski & Thompson, 1999). As noted by Bohner (2008, p. 172), the unimodel postulates that “persuasive evidence can be presented in an unlimited number of forms and variations”, including forms that do not involve words or sentences. Hence, even events such as stimulus pairings might function as persuasive evidence in much the same way as words and sentences do. Whereas the Heuristic-Systematic Model puts forward a heuristic mode of processing that is separated from a more effortful systematic mode of processing, the unimodel retains only one mode of processing that is called upon by all persuasive evidence. As an idea that is situated at a high level of abstraction, symbolic EC is agnostic about the type and number of processes on which it is based. The main point that we want to make in this section is that symbolic EC is closely related to persuasion and compatible with several current theories of persuasion. As far as we know, EC has never before been linked to persuasion in this manner, perhaps because EC has typically been thought of as a primitive, non-symbolic phenomenon.

Implications

In the second section of this paper, we discuss the implications of our analysis for a number of important questions in EC research. First, we try to shed new light on the question of

why EC research is important. Second, we explore ways to establish the presence of (symbolic and non-symbolic) EC. Third, we discuss the boundaries of EC research. Finally, we briefly explore the implications of our ideas for mental and functional theories of EC.

What is the added value of EC research?

The idea that EC is a symbolic phenomenon could be perceived as a threat for EC research. Because it points to an important communality between EC and persuasion, it raises the question of what EC research can offer that cannot be learned from research on persuasion via arguments. Some might argue that the only hope for the survival of EC research is that at least some instances of EC do not depend on the ability to use symbols. In these cases of non-symbolic EC, stimulus pairings would function as a mere proximal cause of changes in liking, that is, as a proximal cause whose impact is does not depend on the ability to use symbols. If it is indeed the case that certain instances of EC are non-symbolic, then studying these instances could offer insights in the origins of preferences that cannot be gained from studying the effects of persuasive arguments on liking.

This brings us to the second central point of our paper (the first one being that many if not all instances of EC are symbolic in nature). We believe that EC research can offer more than the hope of uncovering non-symbolic instances of EC. The idea that many if not all instances of EC are symbolic offers new opportunities for EC research. All instances of symbolic meaning construction involve a proximal event on the basis of which meaning is constructed. Importantly, the consequences of symbolic meaning construction for changes in liking might depend on the nature of the proximal event. For instance, stimulus pairings might have different effects on liking than sentences that describe the relation between stimuli, even when the effect of both types of proximal events depends on the ability to use symbols.

Because of its focus on stimulus pairings as a source of liking, EC research can provide unique

information about the way in which *symbolic meaning construction on the basis of stimulus pairings* gives rise to changes in liking. In this way, EC research contributes in a unique way to our understanding of how likes and dislikes are acquired and how they can be changed.

From this angle, the importance of EC research depends on the extent to which the properties of changes in liking vary as a function of the nature of the proximal event that gives rise to symbolic meaning construction. Ultimately, whether and how the nature of the proximal event matters in this respect can be determined only via empirical research. Nevertheless, there are reasons to believe that it might matter. For instance, a persuasive message such as “Bob is a great guy” could sometimes lead to reactance. That is, people might dislike Bob because they feel coerced into liking Bob. Instead, simply pairing Bob with positive pictures could lead to interpretations of the situation that people are less likely to experience as imposed upon them because they are self-generated (e.g., “a picture of Bob predicts positive images”) but that might still influence their preferences in subtle, perhaps even irrational ways. In fact, it would be interesting to examine whether symbolic EC depends on factors that are known to influence persuasion via arguments (e.g., information about the intentions of the experimenter). Regardless of the merit of this specific argument, it seems worthwhile to at least consider the possibility that stimulus pairings provide a unique way of conveying meaning and thus a unique way of influencing liking. EC research can document this potentially unique route to liking and can thus contribute in a unique way to evaluative learning research.

In sum, EC research is potentially interesting not only because it might reveal a primitive, non-symbolic way of influencing liking but also because it might reveal a unique symbolic way of influencing liking.

How to establish the presence of (symbolic and non-symbolic) EC?

In order to realize this potential, EC researchers need to be able to determine when EC

has taken place and, preferably, when such an instance of EC qualifies as symbolic or non-symbolic. We believe that the definition of EC as put forward by De Houwer (2007) remains useful to achieve this goal. De Houwer (2007) argued that EC is best conceived of an effect, that is, as the impact of stimulus pairings on liking. The central point of the present paper is that, at an abstract level of analysis, stimulus pairings could influence liking because of two distinct functions that stimulus pairing can have. First, they can be a mere proximal cause of changes in liking. This is the case when the impact of the pairings merely depends on their spatio-temporal properties. Second, stimulus pairings can have an effect because they are interpreted as a symbol for the relation between the stimuli, something that can happen only in organisms that are able to treat proximal events as symbols.⁴ However, regardless of whether a specific instance of EC qualifies as symbolic or non-symbolic, in order to establish the presence of EC, evidence is required that the observed change in liking is due to stimulus pairings rather than other proximal events. Once such evidence has been provided, one can try to determine whether that instance of EC qualifies as symbolic or non-symbolic.

Establishing stimulus pairings as a cause of changes in liking can be achieved by manipulating those pairings while eliminating or controlling for other potential proximal events such the repeated presentation of one stimulus, relations between responses and stimuli, and the presentation of other stimuli such as words or sentences that can give rise to meaning construction (see De Houwer, 2007; De Houwer, 2011). First, it is relatively easy to control for the impact of stimulus frequency on liking (i.e., mere exposure effects; see Bornstein, 1989, for a review). For instance, one could vary on a within subjects basis whether a stimulus is paired with a positive or a negative event while keeping the number of stimulus presentations

⁴ In principle, stimulus pairings could function also in other ways, for instance, as a signal. For the sake of simplicity, we will not elaborate on this possibility and focus on whether stimulus pairings do (as in symbolic EC) or do not (as in non-symbolic EC) function as symbols.

constant. Second, effects of response-outcome relations can be eliminated by adhering to a strict Pavlovian procedure in which the presence of events is under the control of the experimenter. Third, one can eliminate meaning construction on the basis of words or sentences by not including such stimuli. Given that most studies on EC do eliminate or control for all these alternative proximal causes or cues in this way, it seems safe to conclude that EC research can document the ways in which stimulus pairings influence liking.

It seems less straightforward, however, to demonstrate the stimulus pairings do not function as a symbol. First, one cannot simply exclude symbolic meaning construction as a factor in EC by verifying the absence of stimuli such as words or instructions. As we noted earlier, symbolic meaning can be constructed on the basis of a wide variety of proximal events, including the fact that two stimuli co-occur in space and time. Second, inferring the absence of symbolic meaning construction from the absence of evidence for this type of meaning construction can also be problematic. As is always the case when trying to demonstrate the absence of something, one could argue that the absence of evidence is due to a failure to use appropriate tests (i.e., test that are sufficiently valid, comprehensive, and sensitive). Given these problems, we foresee a difficult and lengthy debate about whether EC can be found in the absence of symbolic meaning construction.

On the other hand, we believe that it should be possible to establish that stimulus pairings do function as a symbol. One approach is to keep constant the spatio-temporal properties of the pairings while varying the symbolic meaning that can be constructed on the basis of those pairings. Consider the work of Zanon, De Houwer, and Gast (2012) who paired a first compound of two stimuli (AB) with a positive outcome and a second compound of two other neutral stimuli (CD) with a negative outcome. These target pairings were intermixed with filler trials during which other stimuli were paired with the positive or the negative outcome.

Importantly, the filler trials were constructed in line with one of two rules. The first rule stipulated that a stimulus is always followed by the same (positive or negative) outcome regardless of whether it is presented on its own (e.g., K-positive; L-positive; M-negative; N-negative) or in compound with another stimulus (e.g., KL-positive; MN-negative). The second rule specified that the same stimulus is followed by a different outcome depending on whether it is presented on its own (e.g., K-positive; L-positive; M-negative; N-negative) or in compound with another stimulus (e.g., KL-negative; MN-positive). Even though the filler trials do not change the spatio-temporal properties of the target pairings (AB-positive; CD-negative) in any way, they could change the symbolic meaning of those pairings. According to the first rule, observing a target pairing points at an unambiguous relation between stimuli and outcomes (i.e., a particular stimulus will always be followed either by a positive or a negative event; e.g., AB-positive implies that A on its own will also be followed by the positive outcome). According to the second rule, however, the target pairings imply that stimuli and outcomes are related in a more complex manner (i.e., the same stimulus is followed by a different event depending on whether it is presented alone or in combination with another stimulus; e.g., AB-positive implies that A on its own will be followed by the negative outcome). Results revealed a standard EC effect for the target stimuli (e.g., A more positive than C) but only when the filler trials followed the first rule. The moderating impact of the filler trials cannot be due to changes in the spatio-temporal properties of the target pairings (AB-positive; CD-negative) simply because these properties were kept constant. The filler trials could, however, have influenced the symbolic meaning that the stimulus pairings conveyed about the nature of the relation between the stimuli and outcomes.

A similar approach can be used to test whether proximal events other than stimulus pairings give rise to symbolic meaning construction. That is, one could manipulate the

symbolic meaning of the events while keeping the spatio-temporal properties of the stimulus pairings constant. Consider the instruction “Bob loves kittens”. The role of the word “loves” can be examined by replacing it with another similarly valenced word such as “flower” that does not imply the same relation between Bob and kittens. One could also change the position of the word “loves” so that its relational function is altered (e.g., “loves Bob kittens”). If these manipulations moderate the change in liking, it seems safe to conclude that the change in liking is not merely a function of the spatio-temporal co-occurrence of words (i.e., the fact that “Bob” and “kittens” appear in the same sentence) but depends on the symbolic meaning of the words, that is, on what the word “loves” implies about the relation between “Bob” and “kittens”.⁵

What are the boundaries of EC research?

As noted in the previous section, we believe it should be possible to determine whether stimulus pairings influence liking (which establishes the presence of EC), whether they do so by functioning as symbols (which establishes the presence of symbolic EC), and whether proximal events other than stimulus pairings function as symbols. In this section, we consider more complex situations in which multiple proximal events interact in producing changes in liking. Although such changes in liking raise questions about the exact boundaries between different evaluative learning phenomenon, we argue that they do not call in to question the view that EC research is directed at understanding the effects of stimulus pairings on learning (De Houwer, 2007).

Let us consider the work of Zanon, De Houwer, Gast, and Smith (2014) who not only paired stimuli repeatedly but also provided instructions about the relational implications of the stimulus pairings (for related studies, see Fiedler & Unkelbach, 2011; Gast & De Houwer,

⁵ Note that the absence of an effect of these manipulations does not demonstrate that the change in liking is due to mere spatio-temporal stimulus pairings. Such a conclusion would require the assumption the manipulations provide a valid, comprehensive, and sensitive test of symbolic meaning construction on the basis of cues other than stimulus pairings.

2012; Moran & Bar-Anan, 2013; Peters & Gawronski, 2011). All participants experienced a learning phase during which one non-word (e.g., LOKANTA) was paired repeatedly with a well-known positive word (HAPPY) whereas another non-word (e.g., FEVKANI) repeatedly co-occurred with a well-known negative word (SAD). Half of the participants were informed either before or after the stimulus pairings that a non-word had the same meaning as the word it was paired with (i.e., synonym condition) whereas the other participants were told that paired stimuli had an opposite meaning (i.e., antonym condition). Afterwards, (implicit) evaluations of the non-words were in line with the valence of the word they were paired but only when the paired stimuli were described as synonyms. When the antonym instructions were provided before the stimulus pairings, (implicit) evaluations of the non-words were even opposite to the valence of the word with which they were paired.

The results of Zanon et al. (2014) clearly reveal an impact of the relational meaning implied by the instructions (i.e., whether paired stimuli were said to be synonyms or antonyms). Because a proximal event other than stimulus pairings determined the changes in liking, one could argue that the observed effects do not qualify as instances of EC. On the other hand, the instructions of Zanon et al. did not directly specify the relation between stimuli but merely provided information about the relational implications of stimulus pairings. That is, the instructions did not specify whether two particular stimuli were synonyms or antonyms (e.g., “LOKANTA is a synonym of HAPPY”) but only whether *paired stimuli* were synonyms or antonyms. As such, the changes in liking could be attributed to the symbolic meaning constructed on the basis of stimulus pairings (e.g., “the fact that LOKANTA is paired with HAPPY implies that it is a synonym of HAPPY”). The instructions merely changed the relational implications of the stimulus pairings. If one would define EC as all changes in liking that are due to stimulus pairings (i.e., impact of stimulus pairings as a sufficient inclusion

criterion) rather than as changes in liking that are *exclusively* due to stimulus pairings as a proximal cause (i.e., impact of other proximal events as a sufficient exclusion criterion), then the effects observed by Zanon et al. would qualify as instances of EC. Hence, in some cases where the effects of stimulus pairings interact with those of other proximal events, it might be difficult to decide whether observed changes in liking qualify as instances of EC. In the case of the effects observed by Zanon et al., the uncertainty follows from the fact that there is no consensus yet about whether stimulus pairings must be the only source of changes in liking (i.e., stimulus pairings as an exclusive cause) in order to say that the change in liking qualifies as an instance of EC.

Fortunately, definitions do not have to be entirely coherent or decisive in order to be useful. More generally, scientific definitions are not true or false, they are just tools that are there to help researchers to organize their work and their understanding of the world. From this perspective, it does not matter that much whether it is always perfectly clear when a definition applies. For instance, defining EC as the impact of stimulus pairings on liking continues to provide a very useful demarcation of EC research even when there continue to be debates about whether some changes in liking are instances of EC. Regardless of the outcome of these debates, effects such as those of Zanon et al. (2014) that focus on the way stimulus pairings interact with other proximal events to produce changes in liking most certainly help us to understand EC, that is, the way in which stimulus pairings influence preferences.

Even when a particular effect does not qualify as an instance of EC, it might still help us understand EC. Considered instances of so-called EC via instructions. Instructions about stimulus pairings provide a unique subclass of instructions because they refer to the proximal event that is the focus of EC research (i.e., stimulus pairings). Studying the similarities and differences between EC and the effects of instructions about stimulus pairings could thus

provide important information about the processes underlying both phenomena. Moreover, it could constitute a unique avenue for documenting the added value of experiencing stimulus pairings as compared to being merely instructed about the pairings (see Raes, De Houwer, De Schryver, Brass, & Kalisch, 2014, for research on this topic in the context of fear conditioning). The merit of these contributions does not depend on what effects qualify as “true” instances of EC. As long as there is agreement about the fact that the mission of EC research is to study the effect of stimulus pairings on preferences (De Houwer, 2007), all phenomena that can increase our understanding of the impact of stimulus pairings on liking should be of interest to EC researchers, regardless of whether all of these phenomena are called EC.

What are the implications for cognitive and functional theories of EC?

In recent years, EC research has been dominated by the debate between association formation theories and propositional theories of EC. In essence, association formation models postulate that stimulus pairings set in motion a mental mechanism that, more or less automatically, forms associations between representations in memory (see Hofmann et al., 2010, for a review of these models). During the last decade, alternative theories were formulated according to which EC is mediated by the formation of propositional representations (De Houwer, 2014; Mitchell, De Houwer, & Lovibond, 2009). Unlike associations, propositions are statements about the world that can specify different ways in which stimuli are related (e.g., A predicts B, A causes B). Unlike association formation, the formation of propositions is thought to be a relatively non-automatic processes.

During the analysis that we presented in this paper, we intentionally did not refer to this debate. Instead, we used abstract terms such as “symbolic meaning construction” that were not phrased in terms of associations and propositions. We deliberately opted for this approach because we believe that our analysis is situated at a higher level of analysis than the debate

between associative and propositional theories (see De Houwer & Moors, 2015). On the one hand, the idea that stimulus pairings can function as symbols fits better with propositional models of EC than with existing association formation models of EC because the relational nature of symbols (i.e., something that stands for something else) seems to require representations that can encode relational information (i.e., propositions). On the other hand, we do not want to exclude the possibility that symbol use can somehow be explained on the basis of the principle of association formation. Association formation models that can account for the use of stimulus pairings as symbols would compete with proposition formation models in explaining symbolic EC. Importantly, however, such association formation would not question the fact that stimulus pairings can function as symbols in EC. We therefore hope that our arguments and ideas about the symbolic nature of EC will survive the current debate between propositional and associative theories of EC and that researches might reach consensus over the existence of symbolic EC even when they would not agree about whether symbolic EC is mediated by associative or propositional processes.

In a recent paper, Hughes, De Houwer, and Barnes-Holmes (in press) pointed out that EC cannot only be studied from a cognitive perspective but also from a functional-analytic perspective. Without going into too much detail, functional-analytic explanations aim to understand behavior in terms of behavioral principles, that is, general ways in which events in the environment influence behavior. Within this approach, Hughes et al. argued that instances of EC might qualify as instances of a functional phenomenon known as Arbitrarily Applicable Relational Responding (AARR). In simple terms, AARR refers to patterns of behaviors that can be described as “acting as if” stimuli are related (see Hughes et al., in press, and Hughes & Barnes-Holmes, 2016a, 2016b). EC would qualify as an instance of AARR if the pairing of stimuli functions as a cue for people to act as if the paired stimuli are related with regard to their

valence. According to Relational Frame Theory (Hayes, Barnes-Holmes, & Roche, 2001), AARR depends not only on proximal events such as stimulus pairings but also on events in the past (i.e., distal events) that gave rise to symbolic behavior. This brief description already clarifies that there are many parallels between the analysis that we provided in our paper and the argument that (some instances of) EC might qualify as instances of AARR. Research on AARR could thus provide direction to research on understanding symbolic instances of EC. However, during our analysis, we did not explicitly refer to AARR or Relational Frame Theory because we believe that our arguments remain valid even if there would not be a perfect fit between EC and AARR or if the ideas put forward in RFT might not stand the test of time. Hence, also at the functional-analytic level, we hope that our arguments about the symbolic nature of EC will survive future developments in research on AARR and RFT.

Conclusion

We put forward the idea that EC is a symbolic phenomenon. It implies that stimulus pairings change liking because they function as symbols for the nature of the relation between stimuli. Rather than seeing this idea as a threat to EC research, we believe that it highlights new ways in which the study of EC could contribute to evaluative learning research. More specifically, EC studies could reveal how symbolic meaning that is constructed on the basis of stimulus pairings gives rise to changes in liking. Our analysis also has implications for how (symbolic and non-symbolic) EC can be established and is consistent with the idea that the mission of EC research is to elucidate the impact of stimulus pairings on liking. It is related to current cognitive and functional models but supersedes debates at that level, thus increasing the chances of consensus amongst research.

As a final note, we would like to point out that our analysis reveals that symbolic abilities are not a nuisance factor when trying to understand EC.⁶ Instead, the fact that humans are symbolic beings should be a crucial factor in EC research. Once human beings have acquired the ability to use symbols, why would they not make maximal use of this of this ability for making sense of their environment? Why would only words or sentences function as symbols? All kinds of events in the environment can function as symbols, including the pairing of stimuli. Given human propensity to constantly construct symbolic meaning on the basis of all kinds of proximal events, we would not be surprised that symbolic instances of EC are the rule rather than the exception. Of course, there is still merit in looking for non-symbolic instances of EC. Knowledge about these instances of EC would unique insights in the origins of our likes and dislikes and a unique tool for influencing those likes and dislikes. But we can also learn a lot from symbolic instances of EC. By putting symbolic EC effects on the agenda, we hope to stimulate future EC research and to maintain EC as an important topic in psychology.

⁶ We thank Dermot Barnes-Holmes for pitching this idea during one of our many stimulating discussions.

References

- Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968-1987. *Psychological Bulletin*, *106*, 265-289.
- Bruner, J. S., & Goodman, C. C. (1947). Value and need as organizing factors in perception. *Journal of Abnormal and Social Psychology*, *42*, 33-44.
- Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information Processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 212–252). New York: Guilford.
- Deacon, T. W. (1997). *The Symbolic Species: The Co-evolution of Language and the Brain*. New York: Norton.
- De Houwer, J. (2006). Using the implicit association test does not rule out an impact of conscious propositional knowledge on EC. *Learning and Motivation*, *37*, 176-187.
- De Houwer, J. (2007). A conceptual and theoretical analysis of EC. *The Spanish Journal of Psychology*, *10*, 230-241.
- De Houwer, J. (2011). Evaluative conditioning: Methodological considerations. In Klauer, K. C., Stahl, C., & Voss, A. (Eds.) *Cognitive methods in social psychology* (pp. 124-147). New York: Guilford.
- De Houwer, J. (2014). Why a propositional single-process model of associative learning deserves to be defended. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual processes in social psychology* (pp. 530-541). NY: Guilford.
- De Houwer, J., Barnes-Holmes, D., & Moors, A. (2013). What is learning? On the nature and merits of a functional definition of learning. *Psychonomic Bulletin & Review*, *20*, 631-642.
- De Houwer, J., & Moors, A. (2015). Levels of analysis in social psychology. In B. Gawronski

- & G. Bodenhausen (Eds.), *Theory and Explanation in Social Psychology* (pp. 24-40). NY: Guilford.
- De Houwer, J., Thomas, S., & Baeyens, F. (2001) Associative learning of likes and dislikes: A review of 25 years of research on human EC. *Psychological Bulletin*, *127*, 853-869.
- Field, A. P. (2006). Is conditioning a useful framework for understanding the development and treatment of phobias? *Clinical Psychology Review*, *26*, 857-875.
- Fiedler, K., & Unkelbach, C. (2011). Evaluative conditioning depends on higher order encoding processes. *Cognition and Emotion*, *25*, 639–656.
- Gast, A., & De Houwer, J. (2012). Evaluative conditioning without directly experienced pairings of the conditioned and the unconditioned stimuli. *The Quarterly Journal of Experimental Psychology*, *65*, 1657-1674.
- Gast, A., & De Houwer, J. (2013). The influence of extinction and counterconditioning instructions on evaluative conditioning effects. *Learning and Motivation*, *44*, 312-325.
- Gast, A., Gawronski, B., & De Houwer, J. (2012). Evaluative Conditioning: Recent developments and future directions. *Learning and Motivation*, *43*, 79-88.
- Hayes, S. C., Barnes-Holmes, D., & Roche, B. (Eds.). (2001). *Relational Frame Theory: A Post-Skinnerian account of human language and cognition*. New York: Plenum Press.
- Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin*, *136*, 390-421.
- Hughes, S., De Houwer, J. & Barnes-Holmes, D. (in press). The Moderating Impact of Distal Regularities on the Effect of Stimulus Pairings: A Novel Perspective on Evaluative Conditioning. *Experimental Psychology*.
- Hughes, S., & Barnes-Holmes, D. (2016). Relational frame theory: The basic account. In R. D. Zettle, S. C. Hayes, D. Barnes-Holmes, & A. Biglan (Eds), *The Wiley handbook of*

- contextual behavioral science* (pp. 129-178), West Sussex, UK: Wiley-Blackwell.
- Hughes, S., & Barnes-Holmes, D. (2016). Relational frame theory: Implications for the study of human language and cognition. In R. D. Zettle, S. C. Hayes, D. Barnes-Holmes, & A. Biglan (Eds), *The Wiley handbook of contextual behavioral science* (pp. 179-226), West Sussex, UK: Wiley-Blackwell.
- Kruglanski, A. W., & Thompson, E. P. (1999). Persuasion by a single route: A view from the unimodel. *Psychological Inquiry*, *10*, 83–109.
- Levi, J. M. (2007). Symbols. In W. Darity (Ed.), *International Encyclopedia of the Social Sciences* (Volume 8, pp. 249-253). New York: Macmillan.
- Martin, I. & Levey, A. B. (1978). Evaluative conditioning. *Advances in Behaviour Research and Therapy*, *1*, 57-102.
- Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, *32*, 183-198.
- Moran, T., & Bar-Anan, Y. (2013). The effect of object-valence relations on automatic evaluation. *Cognition and Emotion*, *4*, 743-752.
- Peters, K. R., & Gawronski, B. (2011). Are we puppets on a string? Comparing the impact of contingency and validity on implicit and explicit evaluations. *Personality and Social Psychology Bulletin*, *37*, 557-569.
- Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, *19*, 124–203.
- Raes, A. K., De Houwer, J., De Schryver, M., Brass, M., & Kalisch, R. (2014). Do CS-US pairings actually matter? A within-subject comparison of instructed fear conditioning with and without actual CS-US pairings. *PLoS ONE*, *9*(1): e84888.
- Razran, G. (1954). The conditioned evocation of attitudes (cognitive conditioning?). *Journal of*

Experimental Psychology, 48, 278-282.

Zanon, R., De Houwer, J., & Gast, A. (2012). Context effects in evaluative conditioning of implicit evaluations. *Learning & Motivation*, 43, 155-165.

Zanon, R., De Houwer, J., Gast, A., & Smith, C. T. (2014). When does relational information influence evaluative conditioning? *Quarterly Journal of Experimental Psychology*, 67, 2105-2122.

Author Note

Jan De Houwer and Sean Hughes, Ghent University, Belgium. The preparation of this paper was made possible by Grant BOF09/01M00209 of Ghent University and the Belgian Science Policy under the Interuniversity Attraction Poles program (IAP7/33). We thank Dermot Barnes-Holmes for the numerous discussions on topics related to this paper. Correspondence should be addressed to Jan De Houwer, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium. Electronic mail can be sent to Jan.DeHouwer@UGent.be .