

RUNNING HEAD: EVALUATIVE LEARNING

**Learning to Like or Dislike:
Revealing Similarities and Differences Between Evaluative Learning Effects**

Jan De Houwer & Sean Hughes

Ghent University, Belgium

In press. Current Directions in Psychological Science

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

Researchers study phenomena such as the mere exposure effect, evaluative conditioning, and persuasion to learn more about the way in which likes and dislikes can be formed and changed. Often these phenomena are studied in isolation. We review and integrate conceptual analyses that highlight ways to relate these different phenomena and that reveal new avenues for research on evaluative learning. At the core of these analyses lies the idea that evaluative learning can be defined as changes in liking that are due to regularities in the environment. We discuss how this definition allows one to distinguish different types of evaluative learning on the basis of the nature of regularities (e.g., regularities in the presence of one stimulus vs. in the presence of two stimuli) and the function of regularities (i.e., symbolic vs. non-symbolic).

Keywords: evaluative learning, attitudes, evaluative conditioning, mere exposure, persuasion, complex learning

Learning to Like or Dislike:

Revealing Similarities and Differences Between Evaluative Learning Effects

Although we arrive into this world with a small collection of in-built preferences (e.g., for sugary and fatty foods) most of our likes and dislikes are shaped during our lifetimes. It often takes time to acquire them (e.g., many start off disliking and then later savoring the taste of beer) and some can be quite exotic or unexpected (e.g., when we come to like foul smelling cheeses or like being violently shaken on rollercoasters). Because our likes and dislikes are assumed to shape and guide our behavior, psychologists have invested much effort into examining the different ways in which likes and dislikes can be acquired or changed. Currently, research on this topic is heavily fragmented, with different groups of researchers examining different types of evaluative learning effects (e.g., the mere exposure effect, evaluative conditioning, persuasion). Over the past years, we have steadily developed a new way of thinking about evaluative learning (De Houwer, 2007; De Houwer, Barnes-Holmes, & Moors, 2013a; De Houwer & Hughes, 2016, 2020). In this paper, we bring those ideas together for the first time. In doing so, we hope to reveal ways of transcending the partitions that are present in the literature on evaluative learning and thereby inspire new research. More specifically, we define and relate different types of evaluative learning in terms of the environmental regularities that are involved, as well as the function that these regularities have.

On the Definition of (Evaluative) Learning

The concept of an “environmental regularity” is meant to capture the idea that events in the environment can occur in an orderly manner. It encompasses any state in the environment that involves more than one stimulus or behavior at one point in time (De Houwer, Barnes-Holmes, & Moors, 2013a). For instance, the fact that a stimulus is presented repeatedly, or the fact that two stimuli are presented together in space and time, both qualify as

environmental regularities. Learning can then be defined as an effect, that is, as changes in behavior that are due to regularities in the environment (De Houwer et al., 2013a; De Houwer & Hughes, 2020). This definition captures the intuition that learning can be thought of as ontogenetic adaptation, that is, adaptation to the environment during the lifetime of the organism. Because it is agnostic with regard to the mental mechanisms that mediate learning, it allows for maximal theoretical freedom and avoids the problems inherent to confounding to-be-explained effects (e.g., conditioning as the impact of stimulus pairings on behavior) with explanatory mechanisms (e.g., conditioning as the formation of associations in memory; De Houwer, 2011; Hempel, 1970).

Evaluative learning differs from other types of learning in that it focuses on changes in evaluative behavior rather than other types of behavior (De Houwer, 2007; see De Houwer, Gawronski, & Barnes-Holmes, 2013b, pp. 258-260, for a discussion of how evaluative behavior differs from other behavior). Introducing the concept of an environmental regularity allows us to distinguish between different types of evaluative learning effects on the basis of (a) the type of regularity that produces a change in behavior (i.e., what kind of order can be discerned in the spatio-temporal presence of events) and (b) the function of the regularity (i.e., the role it plays in determining behavior).

Before we describe these distinctions in more detail, please note that we do not intend to make ontological claims (i.e., claims about what is ‘true’ in some absolute sense). First, we realize that it is notoriously difficult to reach agreement about what is the “true” or “best” definition of a (psychological) phenomenon. Our definition of evaluative learning is best conceived of as a working definition. Second, our definition implies that all statements about instances of learning are hypothetical claims about the environmental causes of behavior (i.e., that the change in behavior is due to a regularity). Because causality cannot be observed directly, also (evaluative) learning cannot be observed but can only be inferred from

observations. Rather than making ontological claims about learning in general or specific instances of learning, we hope to provide useful conceptual tools and analyses. In doing so, we aim to shed new light on the relation between different types of learning (i.e., it has heuristic value) and reveal new opportunities for research (i.e., it has generative value).

Distinctions Based on the Type of Regularity

Although there are also other types of regularities, within the psychology of learning, typically three types of regularities are distinguished (Bouton, 2016; De Houwer, 2007; De Houwer et al., 2013a; De Houwer & Hughes, 2020): (1) regularities in the presence of one stimulus; (2) regularities in the presence of two stimuli; and (3) regularities in the presence of behavior and stimuli. In simple types of learning, behavior changes as the result of a single regularity in the environment. In complex types of learning, multiple regularities jointly influence behavior (see De Houwer & Hughes, 2020, Chapter 4).

When applied to evaluative learning, these ideas allow us to define and relate three known types of evaluative learning effects (De Houwer, 2007). A first subclass of simple evaluative learning effects encompasses changes in liking that are due to regularities in the presence of one stimulus. A well-known example is the mere exposure effect, which typically refers to an increase in liking of a stimulus that is due to the repeated presentation of that stimulus (Moreland & Topolinski, 2010; Zajonc, 1968). The regularity that is assumed to cause the change in liking involves only one stimulus and refers to the fact that this stimulus is presented repeatedly.

A second subclass encompasses changes in liking that are due to regularities in the presence of two stimuli. Imagine a situation in which a neutral brand name is repeatedly presented together with positive pictures (e.g., smiling faces) and another brand name is paired with negative pictures (e.g., a snarling dog). Research shows that typically, the first brand name is afterwards liked more than the second one (e.g., Pleyers, Corneille, Luminet,

& Yzerbyt, 2007). This acquired difference in liking cannot be due to the mere repeated presentation of the brand names because both brand names have been presented equally often. Instead, it is probably due to the fact that the first brand name co-occurred with positive pictures whereas the second one co-occurred with negative pictures. If this is the case, then the change in liking qualifies as an effect of regularities in the presence of two stimuli. Such changes are often referred to as evaluative conditioning effects (e.g., Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010).¹

The third subclass of simple evaluative learning effects involves changes in liking that result from regularities in the presence of behavior and stimuli. In studies on approach-avoid learning, for instance, participants could be asked to approach one neutral brand name and avoid another one. Assume that, as a result of this procedure, the first brand name is afterwards liked more than the second one. This change in liking could be conceived of as an effect of a regularity in the presence of a neutral stimulus (brand name) and a positive (approach) or negative (avoid) behavior. Such effects have been referred to as operant evaluative conditioning (De Houwer, 2007; Eder, Krishna, & Van Dessel, 2019).

The idea that different types of evaluative learning involve different types of regularities not only sheds light on the relation between known evaluative learning effects but also generates ideas about new ways of changing liking. For instance, in our lab we demonstrated several new types of operant evaluative conditioning and complex evaluative learning by varying the nature of the regularities and the way in which they are related (e.g., Hughes, De Houwer, & Perugini, 2016).

¹ One could argue that changes in liking in mere exposure studies are not due to the repeated presentation of the stimuli but to the fact that stimuli are paired with a positive event (i.e., the absence of a negative stimulus; see Zajonc, 2001, pp. 225-226), in which case those changes would not qualify as instances of mere exposure effects but as instances of evaluative conditioning. Although we do not see strong arguments in favor of this alternative analysis, it illustrates that claims about (types of) learning are always hypotheses about the causes of changes in behavior. More specifically, claims about learning rest on an analysis by the researcher who makes assumptions about what constitutes a stimulus, a response, a regularity, and an effect. For us, the quality of the analysis is not determined by whether it is “true” but by whether it is useful (i.e., has heuristic and predictive value).

Distinctions Based on the Function of Regularities

We recently argued that regularities can change liking either in a non-symbolic or in a symbolic manner (De Houwer & Hughes, 2016). In humans, events in the environment often influence behavior not because of their mere physical characteristics but because of their symbolic meaning, that is, what they stand for (Deacon, 1997). According to this symbolic perspective, early on in their development, humans gain access to a symbolic learning pathway, one that enables them to relate stimuli in a diverse number of ways. Once this ability is acquired, they can come to treat virtually any proximal event in the environment as a symbolic cue. These cues can take many forms: from words and sentences, to musical or mathematical notation, the signs used in sign language, or even gestures such as a wink of the eye. Different types of physical events (e.g., the presence of the sentence “I love you”, the sign language sign for “I love you”, the icon of a red heart in a text message, or a wink to your partner) can have the same symbolic meaning, that is, stand for the same thing.

In an earlier paper, we put forward the idea that also the mere spatio-temporal properties of events might function as a symbolic cue (De Houwer & Hughes, 2016). For instance, the mere fact that a neutral brand and a positive picture occur together in space and time could function as a symbol for the similarity between those stimuli, much like the word “SIMILAR” or the sign language sign for “similar” does. As a result, people start treating the neutral brand as similar to the positive stimulus, which amongst other things, includes responding in positive ways to the brand. Such a change in liking would still qualify as an instance of evaluative conditioning (i.e., a change in liking due to the pairing of the stimuli) but it would be a symbolic instance of evaluative conditioning in that the pairing of stimuli functions as a symbolic cue for how these stimuli are related (see De Houwer & Hughes, 2016, for more details, and De Houwer & Hughes, 2017, 2020, for more technical analyses). More generally, symbolic evaluative learning can be defined as changes in liking that are due

to regularities that function as symbolic cues.²

The idea of symbolic evaluative learning provides a new perspective on how persuasion fits within the realm of evaluative learning. Persuasive messages involve spoken or written sentences (e.g., “Brand A is good”) that are directed at changing beliefs and feelings (e.g., Böhner, Erb, & Siebler, 2008). Because a sentence consists of more than one stimulus at one point in time, it could be regarded as a regularity.³ Hence, its impact on evaluative responses (e.g., the fact that you start liking Brand A) can be conceived of as an instance of evaluative learning. Moreover, persuasion is symbolic in nature because it depends on the symbolic meaning of sentences (i.e., what they stand for).

An important difference between persuasion and other types of symbolic evaluative learning is the nature of the regularity that functions as a symbol. For instance, whereas in symbolic evaluative conditioning, the mere fact that two stimuli are present together in space and time functions as a symbol (e.g., the fact that the name of Brand A appears on a screen together with a positive word), in persuasion, it is the sentence as a whole that functions as a symbol (e.g., a sentence in which the word “is” is present at a particular location together with the name of Brand A and the word “good”; see De Houwer & Hughes, 2016). A second difference is that persuasive messages typically originate from another person (i.e., the sender) who clearly intends to influence someone’s beliefs and feelings (i.e., the receiver). Hence, persuasive messages often evoke reactance. This is less likely to be a problem when other regularities (e.g., stimulus pairings) function as symbolic cues because it is often less

2 As noted by De Houwer and Hughes (2016), the symbolic perspective on learning does not require a commitment to a specific mechanism via which people are able to respond to symbols. Also, the main contribution of the symbolic perspective on learning is not the well-known idea that people can respond on the basis of symbolic meaning, but that this ability can be applied also to mere spatio-temporal properties of events (e.g., the fact that two stimuli repeatedly occur together on a screen).

3 As with other aspects of our analysis of evaluative learning, the proposal that sentences qualify as regularities is not an ontological claim but simply one possible way of analyzing these events. Some might argue that a single stimulus (e.g., the word “LOVE” or the drawing of a heart) can have the meaning of a sentence. We believe, however, that this is possible only if the context relates that single stimuli to other stimuli (e.g., makes apparent that it is Person A who is the lover and Person B is the beloved) which implies that the symbolic meaning is effectively conveyed by multiple stimuli.

clear whether they originate from a sender who has the intention to influence the receiver. Moreover, whereas persuasive messages often blatantly convey the information that the receiver is meant to receive (e.g., “Brand A is good”), the symbolic meaning of other regularities might be less constrained (e.g., “Brand A is somehow similar to the positive words it co-occurs with”).

Much of our recent research is related to the topic of symbolic evaluative learning. For instance, we reasoned that if stimulus pairings influence liking because they function as a symbol, then changes in liking should depend not on the pairings as such but on the symbolic meaning they convey. Hence, we studied the effects of sentences that describe the mere spatio-temporal properties of events (e.g., the sentence “the neutral stimulus and the positive stimulus will appear together on a screen”) and observed that these effects are very similar to the effects of the spatio-temporal events themselves (e.g., actually presenting a neutral and positive stimulus together on a screen; see De Houwer, Van Dessel, & Moran, 2020, for a review).

We also applied this reasoning to observational evaluative learning. Observing others interact in positive or negative ways with objects in the environment is a vital pathway for forming and changing likes and dislikes. It is possible that (some of) these effects arise because the observed interactions are regularities that function as symbolic cues. In line with this idea, recent studies in our lab indicate that sentences describing how others interact with objects, have similar effects as actually observing those interactions (Kasran, Hughes, & De Houwer, 2020).

In still other ongoing work, we explore the idea that when stimuli share a feature, people treat this as a symbolic cue as well. Imagine a task in which on each trial, three words are presented on a computer screen: one neutral, one positive, and one negative. On half of the trials, a first neutral word always appears in the same color as the positive word. On the

other trials, a second neutral word is always presented in the same color as the negative word. Our studies show that afterwards, participants prefer the first neutral word over the second one (Hughes, De Houwer, Mattavelli, & Hussey, in press). It is unlikely that these changes in liking are due to the mere pairing of stimuli because each neutral word occurs equally often with positive and negative words. Instead, we argued that the sharing of a feature symbolizes similarity, much like the word “SIMILAR” does. Hence, when a neutral and a positive word have the same color, people treat this as a cue indicating that those words share also other features such as their valence, which results in positive responses toward the neutral word (see Hughes et al., in press, for other evidence and studies that exclude alternative explanations). In fact, one could argue that the pairing of stimuli implies a sharing of features, more specifically shared spatio-temporal features (e.g., the fact that two stimuli are present at the same time and location). From this perspective, evaluative conditioning effects are just one subset of a broader class of shared features effects.

The recent papers of Kasran et al. (2020) and Hughes et al. (2020) already illustrate the extent to which the idea of symbolic evaluative learning generates entirely new ideas for research. Other examples include studies on the context dependency of evaluative learning. Given that symbolic meaning is heavily context dependent (e.g., the wink of an eye can convey mutual understanding in one context and romantic interest in another), one would predict that also the effect of regularities on liking (e.g., evaluative conditioning, mere exposure) depends on context (see Hughes, Ye, & De Houwer, 2019, for initial evidence supporting this prediction). Likewise, because only verbal humans are symbolic, there might be fundamental differences between evaluative learning in (verbally able) humans versus nonhuman animals. Finally, if not only persuasion but also other types of evaluative learning in verbal humans are symbolic, does this mean that variables that are known to moderate persuasion (e.g., diagnosticity) also moderate other types of evaluative learning such as

evaluative conditioning and observational evaluative learning?

Conclusion

Defining evaluative learning as effects of regularities on liking not only allows us to define and relate different types of evaluative learning but also reveals new avenues for research on the ways in which liking can be changed. We thus hope that the conceptual work reviewed in this paper will help researchers to better understand and control what people like and dislike.

Recommended Readings

De Houwer, J. (2007). A conceptual and theoretical analysis of evaluative conditioning. *The Spanish Journal of Psychology, 10*, 230-241.

This was the first paper in which different types of evaluative learning were related in terms of the nature of the regularity that produces the change in liking.

De Houwer, J., & Hughes, S. (2016). Evaluative conditioning as a symbolic phenomenon: On the relation between evaluative conditioning, evaluative conditioning via instructions, and persuasion. *Social Cognition, 34*, 480-494.

This paper introduces the idea of symbolic evaluative conditioning and discusses its relation to persuasion.

De Houwer, J., & Hughes, S. (2020). *The psychology of learning: A functional-cognitive introduction*. Boston, MA: MIT Press.

This book provides a detailed discussion of the definition of learning, the different ways in which learning can be studied, and the concept of complex learning.

References

- Böhner, G., Erb, H.-P., & Siebler, F. (2008). Information processing approaches to persuasion: Integrating assumptions from the dual- and single-processing perspectives. In W. D. Crano & R. Prislin (Eds.), *Attitudes and attitude change* (pp.161-188). New York: Psychology Press.
- Bouton, M. E. (2016). *Learning and Behavior: A Contemporary Synthesis* (2nd ed.). Sunderland, MA: Sinauer.
- Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain*. New York: Norton.
- De Houwer, J. (2007). A conceptual and theoretical analysis of evaluative conditioning. *The Spanish Journal of Psychology*, *10*, 230-241.
- De Houwer, J. (2011). Why the cognitive approach in psychology would profit from a functional approach and vice versa. *Perspectives on Psychological Science*, *6*, 202-209.
- De Houwer, J., Barnes-Holmes, D., & Moors, A. (2013a). What is learning? On the nature and merits of a functional definition of learning. *Psychonomic Bulletin & Review*, *20*, 631-642.
- De Houwer, J., Gawronski, B., & Barnes-Holmes, D. (2013b). A functional-cognitive framework for attitude research. *European Review of Social Psychology*, *24*, 252-287.
- De Houwer, J., & Hughes, S. (2016). Evaluative conditioning as a symbolic phenomenon: On the relation between evaluative conditioning, evaluative conditioning via instructions, and persuasion. *Social Cognition*, *34*, 480-494.
- De Houwer, J., & Hughes, S. (2017). Environmental regularities as a concept for carving up the realm of learning research: Implications for Relational Frame Theory. *Journal of Contextual Behavioral Science*, *6*, 343-346.

- De Houwer, J., & Hughes, S. (2020). *The psychology of learning: A functional-cognitive introduction*. Boston, MA: MIT Press.
- De Houwer, J., Van Dessel, P., & Moran, T. (2020). Attitudes beyond associations: On the role of propositional representations in stimulus evaluation. *Advances in Experimental Social Psychology*, *61*, 127-183.
- Eder, A. B., Krishna, A., & Van Dessel, P. (2019). Operant evaluative conditioning. *Journal of Experimental Psychology: Animal Learning and Cognition*, *45*, 102-110.
- Hammerl, M., & Grabitz, H.-J. (1996). Human evaluative conditioning without experiencing a valued event. *Learning and Motivation*, *27*, 278–293.
- 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.
- Hughes, S., & Barnes-Holmes, D. (2014). Associative concept learning, stimulus equivalence, and relational frame theory: Working out the similarities and differences between human and non-human behavior. *Journal of the Experimental Analysis of Behavior*, *101*, 156-160.
- Hughes, S., De Houwer, J., Mattavelli, S., & Hussey, I. (in press). The Shared Features Principle: If two objects share a feature, people assume those objects also share other features. *Journal of Experimental Psychology: General*.
- Hughes, S., De Houwer, J., & Perugini, M. (2016). Expanding the boundaries of evaluative learning research: How intersecting regularities shape our likes and dislikes. *Journal of Experimental Psychology: General*, *145*, 731-754.
- Hughes, S., Ye, Y., & De Houwer, J. (2019). Evaluative conditioning effects are modulated by the nature of contextual pairings. *Cognition and Emotion*, *33*, 871–884.

- Kasran, S., Hughes, S., & De Houwer, J. (2020). Social learning of preferences: Comparing the effects of actual observations and verbal instructions. Manuscript in preparation.
- Pleyers, G., Corneille, O., Luminet, O., & Yzerbyt, V. (2007). Aware and (dis)liking: Item-based analyses reveal that valence acquisition via evaluative conditioning emerges only when there is contingency awareness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 130–144.
- Moreland, R. L., & Topolinski, S. (2010). The mere exposure phenomenon: A lingering melody by Robert Zajonc. *Emotion Review*, *2*, 329–339.
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology Monographs*, *9*, 1-27.
- Zajonc, R. B. (2001). Mere Exposure: A Gateway to the Subliminal. *Current Directions in Psychological Science*, *10*, 224-228.

Author Note

Jan De Houwer and Sean Hughes; Ghent University, Ghent, Belgium. The preparation of this chapter was made possible by Methusalem Grant BOF16/MET_V/002 of Ghent University. We thank Olivier Corneille and Yoav Bar-Anan for their helpful comments on an earlier draft. 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.