

January 01, 2009

What's reason got to do with it? Affect as the foundation of learning

Eliza Bliss-Moreau

University of California - Davis

Lisa Feldman Barrett

Northeastern University

Recommended Citation

Bliss-Moreau, Eliza and Barrett, Lisa Feldman, "What's reason got to do with it? Affect as the foundation of learning" (2009). *Psychology Faculty Publications*. Paper 23. <http://hdl.handle.net/2047/d20000870>

This work is available open access, hosted by Northeastern University.

that, given the obvious need to postulate the existence of a propositional module in order to explain many aspects of human associative learning, an association formation module that does not add explanatory power to the propositional module is entirely redundant.

Still, the claim that any instance of human associative learning (i.e., any change in performance that is a result of the presence of regularity in the relation between events in the world) must by necessity be due to the operation of controlled, propositional reasoning processes, is a strong one. *Prima facie*, Mitchell et al.'s claim seems ill-fitted to the existence of a phenomenon such as evaluative conditioning: If, for example, I have developed a liking for white wine because of spending many pleasant holidays in France, it would seem that this liking for white wine does not need to reflect any knowledge about the relation between white wine and anything else, beyond the fact that white wine makes me think of France. Essentially, the fact that white wine makes me think of France is non-propositional (I cannot be right or wrong for being reminded of France upon smelling white wine). As such, evaluative conditioning effects very strongly appear to result from the mind being carried from one idea or representation to another, without any intermediate processing, much like what is the presumed mode of operation of an association (Fodor 2003).

It is then perhaps not surprising that some of the best evidence for automatic association formation comes from an evaluative conditioning study (Baeyens et al. 1990a). Still, the description of the propositional approach as offered by Mitchell et al. leaves open the possibility that even evaluative conditioning effects, although perhaps resulting from automatic, non-propositional memory retrieval processes (an object "automatically" making you think of something pleasant), do necessitate the conscious, falsifiable establishment, in propositional form, of a link between events (if only of the form "event A co-occurred with event B") at some earlier point in time. That is, the fact that evaluative conditioning effects, at performance, are almost by nature non-propositional (the fact that A makes you think of B is not something that you can subsequently evaluate as correct or wrong), does not preclude that they perhaps only occur if people at some point have consciously noticed some sort of real-world relationship between A and B (such as "A has repeatedly co-occurred with B," a statement which you can obviously evaluate to be true or false).

Does this render the propositional account unfalsifiable? Surely, the fact that performance may reflect automatic memory retrieval of propositional knowledge stored earlier and may moreover reflect propositional knowledge indirectly (such as when stored propositional knowledge about the co-occurrence of two events influences your subsequent evaluation of one of both), does make falsification of the general framework difficult, but not impossible. It would suffice to convincingly demonstrate associative learning about entirely subliminally presented CSs to rule out a role for propositional reasoning altogether. The debate about whether such evidence already exists still seems to be open (see Wiens & Öhman 2002 vs. Shanks & Lovibond 2002).

However, the most important contribution of the propositional approach to associative learning is not to be situated in proving the association formation approach wrong. As Mitchell et al. point out, what is perhaps most important, is that it has provided a new perspective on conditioning, not only in humans (where at least a contribution of reasoning processes to learning has long been acknowledged), but also in animals. This perspective has not only enabled us to unveil the importance of rule learning in animal Pavlovian fear conditioning (Beckers et al. 2006), but also to highlight the parallels between extinction learning and rule learning in terms of context sensitivity and generalisation (Wheeler et al. 2008). As such, the propositional approach has opened up a whole new framework for the understanding and the prediction of human and animal conditioning phenomena,

the impact of which is bound to further increase over the coming years.

And perhaps this is where a caveat about the propositional approach to associative learning, in turn, is warranted. Notwithstanding the impressive amount of evidence that the propositional approach is more veridical than the association formation approach, it seems beyond argument that models developed within the association formation tradition have continuing heuristic value as well. As an example, just recently Leung and Westbrook (2008), in a series of extremely elegant experiments, demonstrated that the degree of additional extinction accrued by a cue exhibiting spontaneous recovery is governed by both individual prediction error of the cue and common prediction error of all cues present during an extinction trial. Does such a finding invalidate the propositional nature of associative learning? Not necessarily (probably not, one might even argue). Still, it is obvious that experiments like these would never have been designed, and these findings never revealed, on the basis of our current understanding of propositional reasoning. As such, it may simply be too early for one truth to govern our inquiries into human and animal associative learning. Keeping our antennas open to discover empirical phenomena in the realm of associative learning and conditioning will probably necessitate a willingness to entertain a variety of models and approaches for some time to come.

What's reason got to do with it? Affect as the foundation of learning

doi:10.1017/S0140525X09000892

Eliza Bliss-Moreau^a and Lisa Feldman Barrett^{b,c}

^aDepartment of Psychiatry and Behavioral Sciences, California National Primate Research Center, University of California—Davis, Davis, CA 95616;

^bDepartment of Psychology, Boston College, Chestnut Hill, MA 02467; and

^cMassachusetts General Hospital/Harvard Medical School, Charlestown, MA 02129.

elblissmoreau@ucdavis.edu barretli@bc.edu

www2.bc.edu/~barretli

Abstract: We propose that learning has a top-down component, but not in the propositional terms described by Mitchell et al. Specifically, we propose that a host of learning processes, including associative learning, serve to imbue the representation of the conditioned stimulus (CS) with affective meaning.

In the target article, Mitchell et al. characterize associative learning phenomena according to the *relationship* established between two previously unrelated stimuli (i.e., a conditioned stimulus [CS] and an unconditioned stimulus [US]). Associative learning, they suggest, occurs when the CS becomes *propositionally related* to the US using effortful, controlled, and rational processing. We believe this view does not account for important questions about how the representations in question (the CS and US) are modified by experience. Furthermore, this view makes assumptions about how stored representations are activated. We suggest that stimulus representations are realized by multimodal states reflecting both exteroceptive and interoceptive information brought online by a combination of top-down (e.g., propositional) and bottom-up (e.g., stimulus-driven) processes. In this view, learning occurs when the multimodal representation of a stimulus acquires an affective component. Propositional change is not necessary for learning.

A growing body of evidence suggests that the human brain captures statistical regularities in sensory-motor patterns and stores them as representations. These representations are used to continuously organize incoming sensations during the process of predicting what those sensations stand for in the world (Bar 2003; 2007; Kveraga et al. 2007). External sensations

always occur in a context of internal sensations from the body. As a result, the sensory-motor pattern that is stored for future use will always include a representation of the interoceptive state of the body. The brain predicts what sensations refer to in the world in part based on prior experiences of how those external sensations have influenced, or changed, internal sensations from the body on prior encounters (cf. Barrett & Bar, in press). These bodily changes are referred to as “affective.” Affective states can be described hedonic (pleasure or displeasure) with some degree of arousal (for a recent review, see Barrett & Bliss-Moreau, in press). These ideas are consistent with a growing body of research demonstrating that knowledge about the world is “embodied,” or grounded, by a network of broadly distributed, diverse, multimodal states which are encoded during the experience of a given stimulus (see Barsalou 2008). What you know about an object is therefore based, in part, on its affective impact in the past.

When it comes to learning, changes in a CS’s meaning can be thought of as the process by which the multimodal representation of the stimulus is changed by any experience. The most fundamental change occurs because the representation of the CS is experienced in a context of affective arousal that is derived from the representation of the US. Any number of relationships between the CS and US could serve to alter the representation of the CS. The CS and US could be paired in time or space, associated semantically, or even explicitly coupled via rule-based learning. In our view, the need to differentiate types of learning in terms of how the relationship between CS and US is established (as exemplified by Mitchell et al.’s model) is eliminated. All learning can be subsumed under the same general, basic mechanism that exists in all organisms that possess the capacity to generate affective responses to stimuli in the environment. Thus, to some extent, any change in the representation of a CS is *affective learning* (Bliss-Moreau et al. 2008).

In typical classical conditioning paradigms, examples of USs include shocks (e.g., Vervliet et al. 2005), very loud noises (e.g., Neumann & Waters 2006), and even sexual arousal (e.g., Hoffmann et al. 2004). These USs act on the nervous system directly to generate a robust affective response in a bottom-up or stimulus-driven way that is automatic and unconscious. Other USs, such as negative words or pictures, have a less robust bottom-up effect on the learner’s nervous system. Instead, such USs have top-down effects because they have propositional meaning. The difference in the bottom-up potency of different USs leads some theorists to believe that different models are required to account for learning phenomena. According to the affective learning perspective, this is not so – changes in affect can and do occur via both bottom-up and top-down processing and therefore with both types of USs. For example, evidence from instructed learning paradigms demonstrates that the representation of a CS can be changed by telling a person that a US will be presented after the CS, even if the US is never presented (e.g., Olsson & Phelps 2004). According to the affective learning perspective, the set of instructions that indicates when the (promised, but never presented) shock will occur sufficiently alters the learner’s affective state so that the interoceptive representation of this affective change is integrated into the representation of the CS. We have demonstrated that people can learn the affective value of other people when presented with propositional information about those people (e.g., seeing the phrase “hit a small child” presented with a picture of Sally) (Bliss-Moreau et al. 2008). In this example, the representation of “hit a small child” has an affective component which is integrated into the representation of Sally.

It is possible that some USs are exclusively experienced either via automatic, effortless associative processing *or* via effortful, controlled propositional processing (but not both), as Mitchell et al. and most dual-process theories suggest (for an extensive review, see Evans 2008). A more likely scenario, however, is that the two types of processing are often active in parallel and

serve to constrain each other to make meaning of a given stimulus in a given context. For example, the sound of gunfire is aversive and may have an automatic effect on the nervous system. But, for a person who has never experienced war, that automatic processing may be constrained by propositional information about the “shoot-em-up” Western movies he or she remembers from childhood, resulting in a relatively neutral experience. For a war vet, the automatic processing may be constrained with propositional information gained in the experience of fighting and killing, resulting in a highly aversive experience. Propositional learning, even for a stimulus that has semantic meaning, is not required.

By focusing on how stimulus representations are changed as a result of internal experience, a whole host of learning phenomena can be united under one principle. Our hope is that by approaching learning from this perspective, the field will generate new hypotheses about the way that people learn about the world.

Learning without thinking

doi:10.1017/S0140525X09000909

R. A. Boakes

School of Psychology (A18), University of Sydney, NSW 2006, Australia.
 bobbo@psych.usyd.edu.au

Abstract: The main conclusion to draw from Mitchell et al.’s article is that it is difficult to disentangle cognitive and learning processes in contingency and causal experiments. More compelling evidence for human associative learning comes from research where, because of the type of events involved, participants are unable or unlikely to think about the relationships between the events.

The conclusion Brewer (1974) drew from his review of conditioning research using human participants came as a great shock. For decades it had been very widely accepted – and not just by the many behaviorists of those times – that the processes described by S-R (stimulus-response) or reinforcement theorists based on animal evidence also provided a basis for at least simple aspects of human behavior in a manner that was independent of belief or awareness. Brewer’s conclusion was that there was no convincing evidence to support this assumption. His alternative account of what goes in a conditioning experiment is captured by the following quote: “The college sophomore does not leave his higher mental processes outside the door when he walks into the experimental room, but he uses them to try to understand what is going on and what he should do about it” (p. 2). Despite many subsequent attempts to show his conclusion to be wrong, Brewer (1974) clearly was correct about the overwhelming influence of “higher mental processes” in determining a participant’s behavior in the kind of conditioning experiment – mainly “conditioning” of autonomic responses or of small movements – that he reviewed.

The article by Mitchell et al. can be seen as a successor to Brewer (1974), in which a similar argument is directed mainly at experiments from the past two decades that have used causal or predictive scenarios in experiments to test principles of associative learning. One similarity between past and present research is the overwhelming use of college students as participants, a population that has been selected on the basis of thoughtfulness and then encouraged to be curious. An odd aspect of too many causal judgment experiments is that, although the researchers want their participants’ higher mental processes to operate in order to understand the sometimes complex instructions, interpretation of the results assumes the absence of any such influence following a participant’s first response. In this respect, many points made by Mitchell et al are salutary, including the important one that associations are not expectancies. When applied to animal data, the absence from associative