(main)stream is sometimes necessary to reach the source; there are several reasons to believe that the authors are heading in the wrong direction.

In principle, the proposal that the formation of propositions is necessary for learning to occur—a process suberved by protein synthesis—is not different from the absurd claim that successful protein syntheses in the digestive systems must also be preceded by propositional representations. In fact, given our current knowledge of biological systems, it is highly unlikely that humans, not to speak about other animals, are conscious or in control of the majority of processes underlying learning and memory formation. Nonetheless, they may be both (i.e., conscious and in control) in regards to the expression of the acquired learning.

In an attempt to avoid linking their proposal to established biological principles of learning, the authors point towards the distinction between the psychological and the neural level of explanation and argue that their thesis applies only to the former. However, this illustrates a general problem inherent in the presented approach. By limiting the phenomena under investigation by either defining out of existence critical aspects of associative learning (e.g., their biological principles) or neglecting several lines of existing research (e.g., fear conditioning and lesions studies in humans), the authors end up proving little more than their assumptions. It is also noteworthy that the notoriously problematic terms “conscious” and “awareness,” although central to the argument, are not explicated (except that the reader is reassured that “Aplysia do not have conscious beliefs”; target article, sect. 6.3, para. 3).

Here, we highlight three more specific problems with the approach presented by Mitchell et al.

1. The notion of a unitary type of associative learning resting on conscious awareness sits very uncomfortably with established ideas in evolutionary biology. Evolution is commonly conceived as a slow accumulative process, building layer upon layer of brain tissue that incorporates successful adaptations at one level into more complex functions at higher levels. As a consequence, we share many behavioral systems and their associated neural circuitry with our primitive predecessors, unlikely candidates for using awareness as their primary principle of learning. Nor does it seem a likely evolutionary feat to have reorganized the human brain for exclusive use of this principle to modify behavior. Rather, from the evolutionary perspective, many different forms of learning would be expected, as elaborated by, for example, Gregory Razran (1971). The MacLean (1993) concept of a “triune brain” is one, often discussed, example of layered forms of learning would be expected, as elaborated by, for example, Gregory Razran (1971). The MacLean (1993) concept of a “triune brain” is one, often discussed, example of layered evolution of this kind, which directly implies that there are at least three levels of behavioral organization, each of which may incorporate associative learning: one concerned with reflexes and instincts (brain stem and striatum), a second that incorporates emotion and autonomic control (the limbic brain), and a third level concerned with instrumental behavior and cognition (thalamus and the cerebral cortex).

2. Related to the lack of compatibility with evolutionary thinking is the omission of several lines of research within the neurosciences. A contemporary version of the MacLean concept is the model of rodent fear conditioning by LeDoux (1996), Davis (1992), Fanselow (1994), Maren and Quirk (2004), among others, which has been confirmed in human brain imaging studies (Morris et al. 1998). Because this model posits that the input to and output from the central hub in the fear network do not necessarily have to go through the cortex, it strongly implies that the fear network and its modification through fear conditioning are independent of conscious awareness. Therefore, this model (and its elaboration for human fear conditioning by Öhman & Mineka 2001) clearly implies two levels of learning that are partially independent but also interacting. This model provides an articulated version of dual-process theory that integrates neuroscience and behavior and is now supported by a host of both behavioral (Hamm & Vaitl 1996; Öhman & Soares 1998) and imaging work (Crichtley et al. 2002), showing that conscious awareness of the associated stimuli or their contingency is not necessary for learning to be acquired and expressed.

Providing further support for the independence of (at least) two kinds of learning is the work on patients with lesions on the hippocampus, a structure known to be critical for the formation of declarative memories. Following fear conditioning, these patients fail to report the contingency between two associated stimuli (e.g., a neutral tone or image and an aversive shock) in a fear conditioning paradigm, but they show a normal conditioned response as measured by the skin conductance response, SCR (Bechara et al. 1995). In contrast, patients with lesions to the amygdala, a key player in the brain’s fear network and known to be necessary for the implicit expression of learned fear, display the opposite response pattern with intact declarative memory, but an impaired conditioned response (spared conditioned response or SCR) (Bechara et al. 1995; Weike et al. 2005). These findings show a striking dissociation between explicitly (propositionally) and implicitly (SCR) expressed emotional learning. Further support along the same lines is the demonstration of fear conditioning to unseen visual stimuli in a cortically blind patient with bilateral lesions to the primary visual cortex (Hamm et al. 2003).

Taken together, the findings listed above should make it clear that the psychological and neural levels of explanations are tightly coupled and that psychological models of learning can benefit tremendously by drawing from what is known in the neurosciences.

3. In their argumentation against the “dual-process” theory of learning, Mitchell et al. build a straw man around the claim that the SCR as a measure of learning is unaffected by conscious and controlled cognitions. This purported claim, ascribed to the dual-process camp, is then used to refute that emotional learning can occur without conscious and controlled reasoning. Indeed, some work has shown that the SCR can be used to index learned responses to both consciously and non-consciously perceived stimuli (Öhman & Mineka 2001; Olsson & Phelps 2004). However, the authors neglect the literature which claims that the potentiation of the startle reflex may be less affected by propositional and declarative processes. Whereas the SCR has been shown to be more sensitive to cognitive processes, such as propositional reasoning, which is likely to be cortically mediated, the potentiating of the startle reflex and eyeblink conditioning draw mainly on subcortical and cerebellar mechanisms, respectively (Clark & Squire 1998; Davis 2006). This is supported by the accumulating evidence of a dissociation between the startle response and SCR, in which the SCR tracks conscious awareness of stimuli contingencies, whereas the startle response tracks non-conscious learning (Hamm & Vaitl 1996; Weike et al. 2007).

To sum up, the propositional account of associative learning proposed by Mitchell et al. may be parsimonious, but it critically lacks the compatibility with current evolutionary biology and neuroscience. Cultivating this form of “biological blindness” will not advance our understanding of behavioral phenomena, such as associative learning.

There is more to thinking than propositions
doi:10.1017/S0140525X09001113

Derek C. Penn,a Patricia W. Cheng,a Keith J. Holyoak,a John E. Hummel,b and Daniel J. Povinellib

aDepartment of Psychology, University of California, Los Angeles, Los Angeles, CA 90095; bCognitive Evolution Group, University of Louisiana at Lafayette, New Iberia, LA 70560; Department of Psychology, University of Illinois, Champaign, IL 61820.
dcppenn@ucla.edu http://www.cognitiveevolutiongroup.org/
cheng@lifesci.ucla.edu http://reasoninglab.psych.ucla.edu/
Propositions are not just structured relations. Mitchell et al. briefly mention that propositions specify "the way in which events are related" (sect. 1, para. 5). While true, this is just the first step down a long path towards full-fledged propositions. There are many other critical features of propositions that Mitchell et al. omit, such as the capacity to systematically represent types, variables, roles, and higher-order relations, and to perform rule-governed operations over these representations in an inferentially coherent fashion (Hummel & Holyoak 1997).

Crucially, these propositional forms do not form a package by nomological necessity (cf. Fodor & Pylyshyn 1988). In our view, non-human animals approximate certain features of propositions and not others (Penn et al. 2008). By reducing propositions to structured relations, Mitchell et al. gloss over all the interesting computational and comparative challenges.

Causal learning is not monolithic. We agree that causal learning is not purely associative in either human or nonhuman animals (Cheng 1997; Penn & Povinelli 2007; Waldmann & Holyoak 1992). But this does not mean that all kinds of causal learning or all kinds of animals employ the same degree of propositional sophistication (Penn et al. 2008). There is good evidence that nonhuman animals employ structured representations and are capable of first-order causal inferences. But this does not mean that rats employ the "higher-order reasoning processes" employed by humans (see Penn & Povinelli 2007). In the case of Beckers et al. (2006), for example, the rats' inferences can be modeled as a kind of sequential causal learning that does not require higher-order relational representations (Lu et al. 2008).

Rescorla-Wagner is not propositional. Mitchell et al. claim that the Rescorla-Wagner model can be thought of as a "simple mathematical model of propositional reasoning" (sect. 6.1, para. 1). Yes, the Rescorla-Wagner model could be implemented symbolically but that does not make it a model. The Rescorla-Wagner model assumes a linear generating function and lacks a representation of causal power (Cheng 1997). Ironically, Mitchell et al. miss why associationist theories fail as rational models of causal reasoning.

PDP models are better and worse than Mitchell et al. claim. Mitchell et al. have an idiosyncratic view of parallel distributed processing (PDP) models. On the one hand, they claim that traditional PDP models can account for the propositional capabilities of humans. On the other hand, they claim that "a single node in a PDP model does not represent anything" (sect. 6.2, para. 3). They are wrong on both accounts. PDP models are incapable of representing the structured relations that Mitchell et al. claim are the sine qua non of learning (Hummel & Holyoak 1992). But this does not mean they represent nothing at all. Every node in a PDP network has some equivalence class, and this equivalence class is precisely what it represents. Just because this equivalence class does not correspond to something one can point at does not mean it does not exist.

Darwin and neuroscientists are not all wrong. Mitchell et al. admit that the neuroscientific evidence provides little support for their claim that all learning is propositional. But they dismiss this evidence as inconclusive. They take a similarly dismissive attitude towards the comparative evidence. They admit that *Aplysia* do not employ propositions. But they have no evolutionary explanation for what happened to the *Aplysia*’s primordial associative learning mechanisms in more sophisticated creatures such as rats and humans. Did all this pre-propositional baggage simply shrivel up and die? We think not.

Many nonhuman animals have cognitive capabilities that go far beyond the automatic formation of simple links. But the degree to which propositional mechanisms are employed differs between different kinds of animals and between different kinds of cognitive process within a given individual. There are many ways in which humans are still very much like *Aplysia*. There are other ways in which we are unique.
The computational nature of associative learning

Abstract: An attentional-associative model (Schmajuk et al. 1996), previously evaluated against multiple sets of classical conditioning data, is applied to causal learning. In agreement with Mitchell et al.’s suggestion, according to the model associative learning can be a conscious, controlled process. However, whereas our model correctly predicts blocking following or preceding subtractive training, the propositional approach cannot account for those results.

In their target article, Mitchell et al. point out that, in contrast to the propositional approach, associative models cannot explain some causal learning results. Here we show that an attentional-associative model, previously evaluated against multiple sets of classical conditioning data, provides explanations for causal learning experiments.

An attentional-associative model of conditioning. Schmajuk et al. (1996; henceforth SLG) proposed a neural network model of classical conditioning (see also, Larrauri & Schmajuk 2005; Schmajuk & Larrauri 2006). The network incorporates (a) an attentional mechanism regulated not only by novelty (difference between actual and predicted magnitude) of the unconditioned stimulus (US) as in the Pearce and Hall (1980) model, but also by novelty of the conditioned stimuli (CSs) and the context (CX); (b) a network in which associations are controlled by a modified, moment-to-moment (vs. trial-to-trial) constrained version of the Rescorla and Wagner (1972) competitive rule; and (c) feedback from the associative network to the input. The attentional mechanism was designed to explain latent inhibition (Lubow & Moore 1959), and the feedback loop was included to describe inferential processes such as sensory preconditioning.

Gray et al. (1997) showed that the SLG model also describes automatic (or unconscious) and controlled (or conscious) processing (Pearce & Hall 1980; Schneider & Shiffrin 1977). In the framework of the model, stimulus X might be processed in controlled or conscious mode when environmental novelty and the representation of X, Xn, are large; and in automatic or non-conscious mode when novelty and Xn are small. Therefore, in agreement with Mitchell et al.’s position, the SLG model suggests conditioning occurs mostly consciously. However, according to the model, in the case of latent inhibition, a pre-exposed X with a small Xn remains unconscious. Therefore, in line with Mitchell et al.’s reference to the effects of masking on learning processes, the SLG model suggests that X pre-exposure reduces conscious processing of the X but conditioning still occurs at a slower pace.

Causal learning. Several studies on causal learning were concerned with the effect of additivity information on blocking and backward blocking (e.g., Beckers et al. 2005). Blocking refers to the fact that a potential cause X is not considered a cause of a given outcome (OUT, represented by “+) when it is presented together with another potential cause A, if A had been previously shown to be a cause of that US (A+, AX+). Two potential causes, G and H, are additive if, when presented together OUT is equal to the sum of their OUTs when presented separately (this is represented as G+, H+, GH+). When the joint OUT of G and H is less than the sum of their individual OUTs, the causes are subtractive (G+, H+, GH+). Beckers et al. (2005) demonstrated that additivity pre-training resulted in stronger blocking than subtractivity pretraining (Experiment 2); additivity pre-training resulted in stronger backward blocking than subtractivity pre-training (Experiment 3); additivity post-training resulted in stronger blocking than subtractivity post-training (Experiment 4); and blocking is stronger when OUT is weaker than the maximum OUT experienced by the subjects (Experiment 1). According to Beckers et al. (2005), their results can be explained in inferential terms: blocking is not present if either the additivity premise or the submaximal premise is not satisfied. In the following paragraphs, we describe how the model addresses two of these experimental results.

Additivity training preceding blocking. Like the Rescorla-Wagner model, the SLG model explains blocking because, at the time of the presentation of X, A already predicts the OUT. According to the model, the compound stimulus (C) activated by G and H and associated with OUT during pre-training is fully activated by A and X. This association, together with the blocking stimulus A, contributes to predict the OUT, thereby increasing blocking. Because the C-OUT association acquired during pre-training is stronger in the additive than in the subtractive case, blocking is stronger in the former than in the latter case (see Fig. 1, Left Panels).

We assumed generalization between compounds GH and AX to be strong based in Young & Wasserman’s (2002) experimental data showing that generalization between elements is much smaller than generalization between compounds. In addition, the model implements generalization among elements and between compounds through the presence of a common contextual stimulus that is always active.

Additivity training following blocking. As Mitchell et al. correctly observe in the target article, in the absence of pre-training, the C compound is already associated with OUT during posttraining and, therefore, increased C-OUT associations cannot be used to explain increased blocking. Interestingly, the SLG model provides an attentional interpretation for the result. In terms of the model, during the AX+ phase of blocking, OUT-X and C-X associations are formed. During the subsequent additivity post-training, OUT-X and C-X associations predict X, but X is not there. In the additive case, the stronger OUT extinguishes its OUT-X association faster than the weaker non-additive OUT does. During additivity post-training, presentation of the novel stimuli G and H, as well as the absence of stimuli A and X, increases novelty. Thus, because the representation of X is weaker in the additive case, attention to X increases less, and blocking is stronger than in the subtractive case (see Fig. 1, Right Panels).

Conclusion. In agreement with Mitchell et al.’s position, the SLG model suggests that associative learning can be a conscious, controlled process related to higher-order cognition. Furthermore, in addition to the above experiments, computer simulations show that the SLG model describes (a) the facilitatory effect of additivity training before backward blocking (Beckers et al. 2005), (b) maximality effects (Beckers et al. 2005, Experiment 1), (c) the facilitatory effect of subtractivity pre-training results on backward blocking (Mitchell et al. 2005), and (d) higher-order retrospective revaluation (De Houwer & Beckers 2002). Interestingly, whereas the propositional approach predicts no blocking following subtractive pre- and post-training (see Beckers et al. 2005, pp. 241, 246), the SLG model can account for those results. Furthermore, these finely graded results are also present in the model description of latent inhibition, in which weaker conditioned responding is observed.