Causal Reasoning Versus Associative Learning: A Useful Dichotomy or a Strawman Battle in Comparative Psychology?

Daniel Hanus
Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

The debate about whether or not one could/should ascribe reasoning abilities to animals has deep historical roots and seems very up-to-date in the light of the immense body of new empirical data originating from various species and research paradigms. Associative learning (AL) seems to be a ubiquitous low-level contender for any cognitive interpretation of animal behavior, mostly because of the assumed mechanistic simplicity and phylogenetic prevalence. However, the implicit assumption that AL is simple and therefore the most parsimonious mechanism to describe seemingly complex behavior can and must be questioned on various grounds. Using recent empirical findings with chimpanzees as an example, I argue that at times inferential reasoning might be the most likely candidate to account for performance differences between experimental and control conditions. Finally, a general conclusion drawn from the current debate(s) in the field of comparative psychology could be that a dichotomist battle of 2 conceptual camps—each of which is lacking a clear and homogeneous theoretical framework—is a scientific deadlock.

Keywords: animal cognition, associative learning, inferential reasoning

The Permanent Contender

For Aristotle, human psychology consists of three hierarchical faculties: "nutrition," "perception," and "intellect," and according to him, only humans reach (or can potentially reach) the final, the "intellect" stage (Aristotle, 1999). Centuries later, Descartes famously denied animals the ability to reason and suggested that all nonhuman beings must be best described as complex "automata." According to him and others, language was supposed to be a necessary condition for mind (Descartes, 2006; Locke, 2007)—an assumption that by definition excludes all nonhuman beings from the rational circle. Countless anecdotes about seemingly "clever" behavior in animals have provoked the question of how animal behavior could be best described—if not in mental terms. The school of Behaviorism emerged as a kind of scientific counterpart to contemporary anecdotal reports that led to very rich interpretations of animal behavior (e.g., Romanes, 1970). In the early 20th century, behaviorists (e.g., Skinner, 1953; Thorndike, 1911; Watson, 1928) went beyond Descartes and entirely rejected mentalistic terminology like beliefs and desires as subjectivist fictions, even in human psychology. As a consequence, they promoted the view of psychology as the study of behavior rather than a study of the mind.[1] During the following period the famous "black box" was used as a metaphor to stress the fact that instead of postulating a homunculus-driven mind, only input (stimulus) and output (response) are the crucial components for sufficiently describing all—animal as well as human—behavior. The central idea was that associations between those components are formed automatically without any form of intentionality being involved. For Thorndike, one of the founders of behaviorism, the actual consequences of behavior did not even enter into the formation of an association, which is merely formed between the stimulus context and the individual's response (Blaisdell, 2008; Thorndike, 1898). Subsequent theories of associative learning (AL) followed this general proposal that the stimulus-response connection is *the* crucial aspect of animal learning. Since the 1970s, the theoretical focus has shifted more toward stimulus-stimulus connections (Pearce & Bouton, 2001). Crucially, the behavior described in such terms was still not thought to be goal-directed in any way but must be seen as a series of blind conditioned responses. Later theorists acknowledged the fact that certain animal actions might be driven by expectations of an outcome (Tolman, 1932) and therefore influenced by the value of that outcome (Colwill & Rescorla, 1985; Dickinson & Balleine, 1994; Rescorla, 1988).

Even though the Cognitive Revolution reopened the scientific focus toward more mentalistic interpretations of human psychology as well as animal psychology, the temptation seems to persist until today to take a learning-based explanation as basic and cognitive explanations as a kind of sophisticated titbit (Meketa, 2014)—especially when it comes to analyses of nonhuman behavior. But are there good reasons for this?

As for epistemological reasons, false positive findings (Type I errors) are taken to be more serious errors than false negatives (Type II errors), which results in a bias against attributing psychological traits to animals as common scientific practice (Mikhalevich, 2014; Sober, 2005). Given that line of reasoning, it seems somehow evident that in comparative psychology the undisputed

---

Correspondence concerning this article should be addressed to Daniel Hanus, Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, D-04103 Leipzig, Germany. E-mail: hanus@eva.mpg.de

[1] As it is not the focus of this paper I used this oversimplification that discounts the distinction between methodological and theoretical behaviourism (see Allen & Bekoff, 1997).

standard practice consists of ruling out merely associative explanations before referring to something as "cognitive" (Buckner, 2011). In other words, an associative explanation should be the null hypothesis that must be rejected before any cognitive explanation should be assumed (Dennett, 1983).[2]

However, there is still no satisfying answer to the question of "why this empirical methodology of psychology places a greater burden of proof on animal cognition research than it does on human cognition research" (Sober, 2005). Or to put it differently, why do some psychologists seemingly accept so readily the null hypothesis that different cognitive mechanisms are at work in humans and other animals? Wouldn't modern psychology be better served by removing these evidential biases and instead adopting a more balanced comparative practice in the consideration of human and nonhuman animal research? And finally, how justified is the fundamental claim that AL is the most simple and therefore most parsimonious description when it comes to animal behavior?

How Simple Is AL?

As elaborated previously, one of the main arguments why AL is or should be the permanent contender is the assumption of its simplicity. However, it is that very assumption that could be questioned on several grounds. Especially evolutionary and conceptual arguments have been formulated to support the view of AL as a continual low-level contender (Heyes, 2012; Meketa, 2014). According to the evolutionary argument, AL must be very basic as it is phylogenetically so widespread and therefore probably very old. Basic forms of conditioning can be found throughout the animal kingdom, from mammals to invertebrates (Allen, Grau, & Meagher, 2009; Heyes, 2000; Karin-D'Arcy, 2005; Shettleworth, 2010a). The central component of the conceptual argument is based on the assumption of the mechanistic simplicity of AL. Even though "the relevant metric of simplicity is rarely specified" (Buckner, 2011), there seems to be a widespread temptation to highlight the apparent similarity between the proposed automatic link-formation mechanism and the neurophysiological hardware in which it is implemented. Such supposed similarity gains plausibility, as there are some reports that conditional stimulus (CS) and unconditional stimulus (US) might indeed resemble the activation of two connected neurons and aspects of associative models might under certain circumstances reflect brain activity (Schultz & Dickinson, 2000). Nevertheless, this supposed analogical relationship between theoretical and physiological entities is far from being conclusive. For instance, links between stimuli are not implemented by dendrites, and mental representations are after all unobservable theoretical constructs without any plausible physiological equivalence (Mitchell, De Houwer, & Lovibond, 2009).

More fundamentally, there seems to be a common view that Morgan's Canon can be perceived as a special case of the general principle of parsimony when creating scientific explanations. According to the famous canon, one should not interpret an action as the outcome of a higher—and yet more complex—psychical faculty, if it can be interpreted as the outcome of a lower—and yet simpler—psychical faculty (Morgan, 1894).

Sober (2005, 2009) pointed out sagaciously that Morgan's Canon is internally inconsistent as an assumed "cladistics parsimony" allows very different interpretations. To identify the most parsimonious hypothesis, it is necessary to compare how many changes in character state have to be assumed during evolution. If one takes the original wording of Morgan and substitutes "lower" with "ancestral" and "higher" with "derived" in a phylogenetic tree, the specific implications are not as straightforward as widely assumed (Fitzpatrick, 2008; Sober, 2005, 2009). Remarkably, the parsimony principle could for example favor anthropomorphism over anthropodenial, because it is more parsimonious to view a derived behavior that is exhibited by humans and another animal as homological compared with the view that this behavior evolved independently in the two lineages. Furthermore, the assumption that the behavior exhibited is based on the same proximate mechanism is more parsimonious than to assume that humans and other animals deploy different proximate mechanisms (Sober, 2005). According to Morgan, the default assumption in comparative psychology should be that the proximate mechanism shaping our psychology differs between humans and other animals. He says "we are logically bound not to assume the existence of these higher faculties [unless we have sufficient evidence]." If we do that, we avoid the bias of anthropomorphism. However, as Sober (2005) alluded, "not assuming the existence of something is a different matter from assuming its non-existence." Moreover, there is no logical reason why that principle imposes an asymmetry between higher and lower; as it is, it seems equally problematic to assume the existence of a lower faculty without sufficient evidence.

Coming from another perspective, Gallistel (2008) pointed out that the alleged mechanistic simplicity of AL often requires enormous computational capacities. He uses navigation in honey bees (*Apis mellifera*) to argue that sometimes associative mechanisms are more demanding than alternative cognitive systems in terms of neuronal capacity and energy investment. A representational system of a mental map—and therefore a cognitive model—seems in this case to be a much more parsimonious description compared with the idea that all information is processed through associations alone (Gallistel, 2008). Along similar lines albeit with reference to social cognition, Tomasello and Call (2006) argued that a specific behaviorally based learning rule might serve as a low-level (post hoc) explanation in a particular setting but fails to explain analogous performances in other perceptual modalities and experimental paradigms. As a consequence, the low-level interpretation turns out to be less parsimonious than the cognitive one as only the latter allows for transfer benefits between perceptually different situations without assuming different learning rules for each new context.

Another source of skepticism concerning the simplicity of AL is related to its scientific history. From their onset and despite their enormous success in explaining seemingly complex behavior, AL theories faced the challenge of incompatible empirical findings. Phenomena like "backward blocking" or "recovery from overshadowing" are examples of retrospective revaluation effects that imposed problems for contemporary theories. If two cues (e.g., a light and a tone) are presented together and followed by a reward (e.g., food), both cues establish their own associative strength and each would individually elicit a conditioned response. However, if the

---

[2] Interestingly enough, Dennett himself proposed to take up the *intentional stance* to non-human beings, implying that in order to make useful predictions and to make sense of their behaviour we should view animals as intentional systems (Dennett, 1987).

training started with one cue alone (e.g., a tone = > food) and the second cue was introduced later (e.g., tone + light = > food), no association would be established between the second additional cue and the US. This phenomenon is known as *forward blocking.* *Backward blocking* follows the same logic in principle—namely that training with one cue lowers the associative strength of an additional cue—although this time the blocking occurs even though the (blocked) stimulus was not even present during the critical training trials. This is the case if training starts with two cues simultaneously (e.g., a tone + a light = > food) followed by a period in which only one of the cues leads to the same outcome (e.g., tone = > food). As a result, the absent cue (light) loses its associative power and is no longer sufficient to elicit a conditioned response. Likewise, if one cue (light) alone is not paired with the outcome (food), subjects will increase their judgments of the causal strength of the second cue (tone) during its absence. One of the most influential learning models, introduced by Rescorla and Wagner (1972), was generally able to implement "cue competition effects" (forward blocking) but failed to explain how cues can lose their associative potential while they are absent (as in backward blocking). As a reaction to this incompatibility, some authors suggested modifying the original model by implementing positive and negative learning parameters. Such theoretical maneuver seemed necessary to explain the influence of presented as well as nonpresented cues (Van Hamme & Wasserman, 1994). While up to this point stimulus—response associations were the crucial building blocks, stimulus-stimulus associations became more and more important (Chapman, 1991; Dickinson & Burke, 1996). However, even this conceptual extension did not seem sufficient, as recent findings in humans suggest that the strength of a conditioned response depends on the relationship between stimuli that were only indirectly linked through other intermediary associations (De Houwer & Beckers 2002a, 2002b). Because the relevant cues actually never occur together here, even modern associative models have problems dealing with such experimental findings (Wasserman & Castro, 2005). Because neither stimulus-response nor stimulus-stimulus association seemed sufficient anymore, the *extended comparator hypothesis* was introduced (Blaisdell, Bristol, Gunther, & Miller, 1998). The central idea here is that cues do not compete during training/learning but during performance/retrieval (Denniston, Savastano, & Miller, 2001).

Despite numerous modifications and adjustments, contemporary associative models still assume that in the case of retrospective revaluation the driving aspect that changes associative strength is not diagnostic in a sense that past events are considered and evaluated by the system. Instead learning seemed to be determined by future expectations of currently absent cues (Hagmayer & Waldmann, 2006). However, recent experimental findings do indeed suggest the existence of "true" diagnostic retrospection in humans (De Houwer, 2002; Waldmann, 2000) and rats (Waldmann, Schmid, Wong, & Blaisdell, 2012). Here, human as well as nonhuman subjects clearly distinguished between the obvious absence of a cue and the mere lack of evidence for its presence. Another aspect worth mentioning is that for classic associative accounts contiguity is the crucial parameter determining associative power. However, Rescorla (1968) himself demonstrated that contiguity alone is not sufficient to explain AL, because additional contingency is required as well. The implication for causal learning is simply that not the mere number of cause-effect pairings is

crucial to establish associative power, but also the number of cases in which the effects are observed without the causes being present (Hagmayer & Waldmann, 2006).

This very sketchy and simplified historical digest shows many modifications and adjustments of original theories over the decades provoked by incompatible empirical findings. Adapting and modifying an existing theory in the light of new empirical findings is in itself not reprehensible, and could be viewed as an example of good scientific practice. However, in addition to the development and improvement of existing AL theories, an immense diversification of models has taken place, resulting in a highly technical vocabulary (Carruthers, 2009; de Wit & Dickinson, 2009; Pearce, 2008; Shettleworth 2010a). Obviously, additional assumptions and constant theoretical adjustments are necessary to make individual associative models compatible with the pile of new experimental data. Such dynamic development leaves it almost impossible to follow the literature and theoretical diversifications appropriately from "outside." Ironically, the whole endeavor not only loses its initial functional simplicity (Sober, 2009)—the major trump of early theories of conditioning—but also its cohesive character—the declared goal of most associative theorist. It seems that today, nearly any empirical finding could potentially be simulated by an associative model (Mitchell et al., 2009) which—to put it negatively—reminds one of the overambitious behavioristic agenda to explain "everything."

Why Is Learning (Sometimes) So Difficult?

According to some authors, humans are the only species that can mentally go beyond surface features and are able to cognize the world around them in terms of complex, unobserved predicates and relations (like causality, support, force). With the so-called *reinterpretation hypothesis*, it has been argued that even though animals might have concepts of varying complexity, these concepts are always restricted to the perceptual level (Penn et al., 2008; Penn & Povinelli, 2007; Vonk & Povinelli, 2006).

A growing body of empirical data—mainly from ape and bird studies—has challenged this claim and attributed to subjects simple forms of causal reasoning in problem solving contexts (e.g., Call, 2004, 2006b; Hanus & Call, 2008, 2011; Seed & Call, 2010; Taylor, 2014). The overall conclusion drawn by the authors of those studies is that animal subjects are sensitive to the logicocausal relations of the problem they were confronted with and in that sense go "beyond the [perceptual] information given" (Waldmann, Hagmayer, & Blaisdell, 2006). In the study by Hanus and Call (2008), for example, it has been argued that chimpanzees were able to infer the location of a food reward based on the effect of its weight. Subjects were presented with two opaque cups mounted on opposite sides of a balancing beam. After the experimenter showed a food piece to the subject it disappeared in one of the cups—out of the subject's view. As a result the beam tilted, with the baited side moving downward and the empty side upward. Chimpanzee subjects preferentially inspected the cup on the lower side. However, they were indifferent in their choice when the cups were mounted on a static inclined beam without any titling (static control) or when the beam was externally pushed down to its final slanted orientation instead of letting it move by itself (external cause control). The performance differences between the causally coherent condition and the noncausal controls suggest that besides

perceptual similarities, structural differences between the conditions facilitated the observed differences in choice behavior.

Obviously such a conclusion, namely that nonhuman subjects draw weight-based causal inferences, must stand in opposition to the general claim that animals are unable to draw any kind of inference and especially to the claim that among all animals only humans understand unobservable object features—like weight (Penn, Holyoak, & Povinelli, 2008; Povinelli & Vonk, 2003, 2004; Vonk & Povinelli, 2006). As a consequence, Povinelli, who recently assigned all nonhuman beings a "World without weight," regularly takes issue with the alleged cognitive interpretations of such kinds of experiments (Povinelli, 2011). It is informative to take a closer look at the particular critique that has been brought up to promote a low(er)–level and apparently more parsimonious explanation of the study just described. Referring to Hanus and Call (2008) Povinelli argued that "No one would be surprised by the claim that chimpanzees (and many other species) know that things move down when they are released—this is simply the ability to represent and track observable objects and events." The only thing the subject had to do was to "track the highly salient visual cue of a downward movement of the cup and then select it"—a phenomenon that is known as *gravity bias* and is well described in several primate species (Hood, Hauser, Anderson, & Santos, 1999; Tomonaga, Imura, Mizuno, & Tanaka, 2007). Chimpanzee simply "track this cue regardless of whether they conceive of weight as a causal mechanism or not" (Povinelli, 2011). This criticism exemplifies a series of alternative low-level interpretations and reveals a general methodological misunderstanding, consequently missing the crucial point of this study (and similar ones). It is essential to understand that in all of the conditions (causal and noncausal) the subject did not see to which side (i.e., into which hand) the food went. Subjects just knew that a food item must be in either of the two hands and it had to *infer* into which of the cups it had been released based on how the beam then moved. The whole logic of the task was based on that diagnostic inference— seeing an effect (movement) and inferring a potential preceding cause (weight of the food). Chimpanzees that simply follow the trajectory of a falling object would be a rather trivial finding, and this was clearly not the punchline of this study. Nonetheless, even if the critique were more applicable by ascribing a general preference for downward compared with upward movements of a balance beam, it remains questionable whether such general preference is likely to be the result of a learned association.

The central question arising from those (and other) findings is: Why do apes have such a hard time learning what cognitively minded researchers call "arbitrary" or "noncausal stimulus relations (Call, 2006a; Premack & Premack, 1994)? With respect to the Hanus and Call (2008) study just described, the question would be "Even if subjects do not do so right away, why can't they just *learn* that in the noncausal control conditions the lower side of the apparatus is always the side that contains food?"—given that contingency as well as contiguity parameters indicate a fully reliable relationship (food was always in the lower cup). To be precise, apes probably can learn such arbitrary object relations, but interestingly enough that takes so much more time than learning an event sequence with an inherent causal structure (see also Schrauf & Call, 2009 for similar results). Here are two other examples of this asymmetry between arbitrary and causal relations: It seems extremely demanding for chimpanzees to learn the seemingly

simple connection between a certain color cue and the presence of food (e.g., white cups are always baited and black cups are always empty), while it takes much less time to perform perfectly when a weight cue indicates the presence of food (e.g., heavy cups are always baited and light cups are always empty; Hanus & Call, 2011). Also, the performance differences between a causal auditory cue (e.g., rattling sound elicited by solid food item inside a shaken cup) and an arbitrary auditory cue (e.g., tapping sound elicited on a stationary cup) are pointing in the same direction (Bräuer, Kaminski, Riedel, Call, & Tomasello, 2006; Call, 2004).

All of those experiments followed the same principle logic, namely to contrast two classes of event sequences. From a human perspective, one class consists of events that are connected through a causal relation while the other consists of events that are only spatial-temporally related. Whereas in the first case a physical necessity binds two events together, in the latter the connection is of purely arbitrary nature. It is crucial to accentuate that from an associationist's viewpoint no difference in performance or learning speed should be expected as long as the perceptual salience, contiguity and the previous history of exposure are identical between these two kinds of event structures. So obviously, inferential and associative approaches generate different predictions and, as recommended by many scholars, the preferred way to search for the best explanation should be empiric (e.g., Heyes, 2012; Sober, 2009). As it turned out, results from several recent studies using different experimental paradigms suggest that nonhuman subjects perform very differently in causally structured compared with arbitrarily structured event sequences (Seed, Hanus, & Call, 2011). These performance differences can already appear during the very first exposure to the problem (e.g., Hanus & Call, 2008) or become manifest through a very different and abridged learning curve (e.g., Hanus & Call, 2011).

So what exactly differentiates those two classes of events, if not contiguity and contingency (which are kept identical) or personal experience—because there are no plausible reasons to believe that the two conditions differed in that respect? It has been argued that the two classes of event or object relation differ with respect to their *salience* so that subjects' attention is unequally captured by the two scenarios. Even though nonhuman primates do not understand "anything about the underlying causal structure involved, [they are] predisposed to perceive certain clusters of features as more salient than others" (Penn & Povinelli, 2007). But the important question is: What exactly makes certain object features more salient than others? One answer could be that repeated experiences lead to the formation of associations between certain stimuli and a positive outcome (e.g., food). As mentioned previously, however, in most of those experiments a (unintended and undocumented) learning history that might lead to a conditioned preference for certain stimulus constellations is not utterly impossible but at the very least highly implausible. Apparatuses and stimuli used in such experimental setups do usually not share perceptual features with every day objects and situations the animals are confronted with. That leaves us with the question, if not the repeated exposure followed by a positive (or negative) outcome, what is the basis on which certain object or event features gain salience for the ape subjects?

Along with others, I believe that causal relations between events or objects are a special class of relation in the world, and that the ability to detect such relations is highly beneficial, especially for

mobile organisms living in rather inconstant environments. It is beneficial because it allows the individual to take a cognitive shortcut that enables them not only to predict future events and analyze past ones but also to causally intervene without taking the long and effortful path of trial and error learning. In that sense, every day experience is indeed important—how could it be otherwise—not to form blind and automatic associations, but rather to build-up causal representations that allow for cognitive transfers between structurally similar (yet perceptually dissimilar) situations. That is to say that noninferential learning processes might contribute to behavioral outcomes under certain conditions, but there are clearly circumstances in the life of nonhuman animals where that cannot be the whole story.

A False Dichotomy After All?

In the final section, I aim to delineate future perspectives by revitalizing the question of whether associative and cognitive explanations are mutually exclusive approaches to describe behavior and speculate what could be a promising avenue to enhance the field of comparative psychology. While some authors have repeatedly argued against the view that cognitive and associative descriptions are incompatible (Blaisdell, 2008; Buckner, 2011; Danks, 2009; Lagnado, Waldmann, Hagmayer, & Sloman, 2007), daily scientific practice these days seems still tacitly dominated by the "battle of the camps." However, ubiquitous as the battle is, it is difficult to precisely identify the adversaries in the arena. Alongside with the theoretical disagreements it is first and foremost the polemic tone in the battle, which complicates the issue unnecessarily.

One theoretical complication results from the fact that contemporary AL theories reveal a hybrid character, meaning they include classical cognitivist credentials and instrumental learning components at the same time (Heyes, 2012). In contrast to the strict avoidance of mentalistic terminology of the original behavioristic school, contemporary researchers in the field of AL readily implement cognitive constructs like memory, attention and representation in their terminology. Another nonbehavioristic aspect of contemporary AL is that nobody would seriously deny that AL processes happen in the mind of an individual (Heyes, 2012), which consequently lets a growing number of AL researchers acknowledge the fact that learning does not necessarily require a concurrent change in behavior but could be sufficiently described as "behaviorally silent" (Dickinson, 1980). While in classical behaviorism changes in behavior were the one and only starting point to initiate learning, modern AL models easily implement behavior as outcome without the necessity of a behavioral input (Pearce & Hall, 1980; Wagner, 1981).

Others have alluded to the more general fact that it is a theoretical misconception to view learning and cognition as conceptual opponents. As Tomasello and Call (2006) pointed out, the most complex forms of human cognition are based on language or algebra—both being learned skills. So the crucial question is not whether certain behavior is learned, but rather whether this learning is based on blind associations or on some causal or intentional inference of the situation (Tomasello & Call, 2006). It is crucial to emphasize over and over again that such "inference" does not necessarily need to be of an explicit, conscious or even verbalized nature. Furthermore, even if alleged sophisticated cognitive explanations can be challenged by low(er)-level alternative explanations of behavior, those alternatives do not need to be associative in nature. A growing body of evidence suggests that learning in humans and animals is caused by a multitude of different mechanisms, making it outdated to insist on a unique, general-purpose learning system that does the job (Gallistel, 2000; Matute & Vadillo, 2009; Meketa, 2014; Tooby & Cosmides, 2005; Waldmann, Cheng, Hagmayer, & Blaisdell, 2008).

So where does this false dichotomy and misconception of a cognition/association exclusiveness come from and why is it so tenacious? One of the main reasons for the existing disarray seems to be terminological imprecision—a regular solecism that most of us are guilty of. As Buckner (2011) notes, "associationism has been variously conflated with empiricism, behaviorism, and connectionism; and cognitivism has been conflated with rationalism, vitalism, introspectionism, the computational theory of mind, and, when applied to non-human animals, anthropomorphism." Other frequent (false) dichotomies are top-down propositional versus bottom-up stimulus-driven, abstract versus perceptually based, controlled versus automatic, symbolic versus rule-based, conscious versus unconscious, explicit versus implicit, to name but a few. One of the most promising ways to avoid such terminological confusion might be a more vivid theoretical exchange between these different schools of thought. That however seems easier postulated than done—on the one hand because of the mere theoretical diversification which makes it very hard to keep track of every conceptual development in each field, and on the other because of the regional 'scientific enculturation.' Certain departments located in certain institutions are just dominated by certain research ideologies and it is very difficult (if not impossible) for students and scholars to remain open to alternative theoretical approaches.

Furthermore, the conceptual relationship between the cognitive-associative discourse in comparative research and recent accounts of problem-solving and decision-making psychology seems blurry—at least for many of us. For example, key assumptions and implications from rational probabilistic models like Causal Bayes nets (Gopnik et al., 2004) are just not concerned with issues of high- versus low-level, animal versus human, conscious versus unconscious, or resource-dependent versus automatic (Gopnik, 2009). Another example might be contemporary connectionist models that differ from automatic link formation mechanisms, as they are able to implement top-down influences like attentional control and are also able to represent semantic information (Shanks, 2009).

So is it just a question of theoretical perspective and scientific interest after all? Obviously, different disciplines ask different questions and pursue different research agendas. Let's consider the example of a traffic light given by Tomasello (2014): One may be interested in questions like "How efficiently does this light regulate the traffic?" or "How does public transport benefit from certain traffic light systems?" In contrast, someone else may be interested in "How the electric circuits within the light work together in order to enable the emission of certain light patterns." The first type of question is concerned with a phenomenological understanding of "traffic" and its functional aspects, whereas the other type of question deals with electrical wiring and technical implementations. Crucially, even though these two types of analyses might (or might not) be seen as complementary, one approach

cannot contribute anything useful to satisfactorily answer questions asked by the other and vice versa. Electrical engineers have nothing to say about (and are probably not interested in) traffic, and traffic planners are not concerned with (and are probably not interested in) electronic circuits and light bulbs. Nevertheless, it is crucial to note that in the traffic example, the main issue seems to be the different levels of description. Theories and models within those different levels cannot contradict each other and consequently would not create different empirical predictions. Is such a solution conceivable in the case of the cognition-association debate? Is it possible, for example, that subjects might have causal representations of the world but nevertheless use associative processes for prediction and parameter learning, as some researchers have suggested (Lagnado, 2009)? While the associative system simply links stimulus A and B, a propositional causal model represents how A and B are related to each other—for example, as preceding cause and following effect (Pearl & Russell, 2001).

Despite conceptual concerns there is also a very practical reason why a false cognition-association dichotomy is problematic. False dichotomies create false battles, which suffer from the fact that nobody seems to know what exactly the two scientific camps or schools of thought actually are these days. As a result, researchers who consider themselves either more cognitively or more associatively minded target the other "camp" in a way that both intellectual opponents do not recognize themselves anymore within the other argumentation. Nobody would consider himself or herself as "association blind" or "simple minded" (Heyes, 2012) but rather feel that they have a theoretical affinity toward one or the other side. A consequence of such simplification is a polemic battle of schools or scientific camps instead of an open-ended search for the best explanatory approaches to empirical phenomena. And as the camps diverge more and more ideologically, the willingness to acknowledge the potential explanatory value of a different theoretical perspective diminishes. The scientific discourse seemed (too) long dominated by the "question of whether all thought was best conceived of as associative or propositional in nature" (Buckner, 2011) instead of asking what the circumstances are that allow or trigger associative and cognitive capacities in (human and nonhuman) animals.

In conclusion, instead of pursuing or even heating up the old battle it seems about time for new perspectives. The idea is far from new—some researchers have already advocated more integrative accounts of various kinds (e.g., Bekoff, Allen & Burghardt, 2002; Buckner, 2011; Danks, 2009). Lagnado (2009) suggested endorsing a "unified framework that takes the interaction between learning and reasoning seriously, but allows for variation in the complexity of representations and inferential processes." Even more (maybe too) ambitious is the radical proposal to rethink the foundations of cognitive science by cleaning up our terminology, first and foremost to begin with the ambiguous term of "mental representation" (Morsella, Riddle, & Bargh, 2009).

The hope for the field of comparative psychology would be that different empirical predictions that are generated by different theoretical approaches are spelled out explicitly and beforehand (Heyes, 2012). Generating hypothetical post hoc explanations might be scientifically valid but nevertheless leaves a kind of destructive aftertaste. A taste that Mitchell et al. (2009) apparently had in mind when they attested that "from experience we have learned that it is difficult to produce a pattern of data that cannot be explained by one or the other variant of these associative models."

So, instead of an "us against them" approach in which the "us" is as blurry and poorly defined as "them," we need new research paradigms that are diagnostic and—in the ideal case—generate different hypotheses that can be tested through experiment and observation. Those paradigms and testable ideas might most likely arise from intrepid collaborations of different research groups with different theoretical backgrounds (see the contribution of Dwyer & Waldmann, in this issue, for a potentially promising example).

References

Allen, C., & Bekoff, M. (1997). *Species of mind: The philosophy and biology of cognitive ethology*. Cambridge, MA: MIT Press.
Allen, C., Grau, J., & Meagher, M. (2009). The lower bounds of cognition: What do spinal cords reveal? In J. Bickle (Ed.), *The Oxford handbook of philosophy of neuroscience* (pp. 129–142). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780195304787.003.0006
Aristotle. (1999). *The metaphysics*. New York, NY: Penguin Classics.
Bekoff, M., Allen, C., & Burghardt, G. (Eds.). (2002). *The cognitive animal: Empirical and theoretical perspectives on animal cognition*. Cambridge, MA: MIT Press.
Blaisdell, A. P. (2008). Cognitive dimension of operant learning. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (Vol. 1, pp. 173–195). Oxford, UK: Academic Press. http://dx.doi.org/10.1016/B978-012370509-9.00184-4
Blaisdell, A. P., Bristol, A. S., Gunther, L. M., & Miller, R. R. (1998). Overshadowing and latent inhibition counteract each other: Support for the comparator hypothesis. *Journal of Experimental Psychology: Animal Behavior Processes, 24,* 335–351. http://dx.doi.org/10.1037/0097-7403.24.3.335
Bräuer, J., Kaminski, J., Riedel, J., Call, J., & Tomasello, M. (2006). Making inferences about the location of hidden food: Social dog, causal ape. *Journal of Comparative Psychology, 120,* 38–47. http://dx.doi.org/10.1037/0735-7036.120.1.38
Buckner, C. (2011). Two approaches to the distinction between cognition and "mere association." *International Journal of Comparative Psychology, 24,* 314–348.
Call, J. (2004). Inferences about the location of food in the great apes (*Pan paniscus, Pan troglodytes, Gorilla gorilla,* and *Pongo pygmaeus*). *Journal of Comparative Psychology, 118,* 232–241. http://dx.doi.org/10.1037/0735-7036.118.2.232
Call, J. (2006a). Descartes' two errors: Reason and reflection in the Great Apes. In S. Hurley, & M. Nudds (Eds.), *Rational animals?* (pp. 219–234). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780198528272.003.0010
Call, J. (2006b). Inferences by exclusion in the great apes: The effect of age and species. *Animal Cognition, 9,* 393–403. http://dx.doi.org/10.1007/s10071-006-0037-4
Carruthers, P. (2009). Invertebrate concepts confront the generality constraint. In R. Lurz (Ed.), *The philosophy of animal minds* (pp. 89–107). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511819001.006
Chapman, G. B. (1991). Trial order affects cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 837–854. http://dx.doi.org/10.1037/0278-7393.17.5.837
Colwill, R., & Rescorla, R. (1985). Postconditioning devaluation of a reinforce affects instrumental responding. *Journal of Experimental Psychology: Animal Behavior Processes, 11,* 120–132. http://dx.doi.org/10.1037/0097-7403.11.1.120
Danks, D. (2009). The psychology of causal perception and reasoning. In

H. Beebee, C. Hitchcock, & P. Menzies (Eds.), *The Oxford handbook of causation* (pp. 447–470). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780199279739.003.0022

De Houwer, J. (2002). Forward blocking depends on retrospective inferences about the presence of the blocked cue during the elemental phase. *Memory & Cognition, 30,* 24–33. http://dx.doi.org/10.3758/BF03195262

De Houwer, J., & Beckers, T. (2002a). Higher-order retrospective revaluation in human causal learning. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology, 55,* 137–151. http://dx.doi.org/10.1080/02724990143000216

De Houwer, J., & Beckers, T. (2002b). Second-order backward blocking and unovershadowing in human causal learning. *Experimental Psychology, 49,* 27–33. http://dx.doi.org/10.1027/1618-3169.49.1.27

Dennett, D. (1983). Intentional systems and cognitive ethology: The "Panglossian paradigm" defended. *Behavioral and Brain Sciences, 6,* 343–390. http://dx.doi.org/10.1017/S0140525X00016393

Dennett, D. (1987). *The intentional stance*. Cambridge, MA: MIT Press, Bradford Books.

Denniston, J., Savastano, H., & Miller, R. (2001). The extended comparator hypothesis: Learning by contiguity, responding by relative strength. In K. Mowrer (Ed.), *Handbook of contemporary learning theories* (pp. 65–117). Mahwah, NJ: Erlbaum.

Descartes, R. (2006). *A discourse on the method*. New York, NY: Oxford University Press.

de Wit, S., & Dickinson, A. (2009). Associative theories of goal-directed behaviour: A case for animal-human translational models. *Psychological Research, 73,* 463–476. http://dx.doi.org/10.1007/s00426-009-0230-6

Dickinson, A. (1980). *Contemporary animal learning theory*. New York, NY: Cambridge University Press.

Dickinson, A., & Balleine, B. W. (1994). Motivational control of goal-directed action. *Animal Learning & Behavior, 22,* 1–18. http://dx.doi.org/10.3758/BF03199951

Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgements. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology, 49,* 60–80. http://dx.doi.org/10.1080/713932614

Fitzpatrick, S. (2008). Doing away with Morgan's Canon. *Mind & Language, 23,* 224–246. http://dx.doi.org/10.1111/j.1468-0017.2007.00338.x

Gallistel, C. R. (2000). The replacement of general-purpose learning models with adaptively specialized learning modules. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 1179–1191). Cambridge, MA: MIT Press.

Gallistel, R. (2008). Learning and representation. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (Vol. 1, pp. 227–242). Oxford, United Kingdom: Elsevier. http://dx.doi.org/10.1016/B978-012370509-9.00082-6

Gopnik, A. (2009). Rational constructivism: A new way to bridge rationalism and empiricism [Peer commentary on "The propositional nature of human associative learning" by C. Mitchell, J. De Houwer & P. Lovibond]. *Behavioral and Brain Sciences, 32,* 208–209. http://dx.doi.org/10.1017/S0140525X0900096X

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review, 111,* 3–32. http://dx.doi.org/10.1037/0033-295X.111.1.3

Hagmayer, Y., & Waldmann, M. (2006). Kausales denken [Causal thinking]. In J. Funke (Ed.), *Enzyklopädie der Psychologie "Denken und Problemlösen", Band C/II/8* [Encyclopedia of Psychology "Thinking and Problem Solving"] (pp. 87–166). Göttingen, Germany: Hogrefe Verlag.

Hanus, D., & Call, J. (2008). Chimpanzees infer the location of a reward on the basis of the effect of its weight. *Current Biology, 18,* R370–R372. http://dx.doi.org/10.1016/j.cub.2008.02.039

Hanus, D., & Call, J. (2011). Chimpanzee problem-solving: Contrasting the use of causal and arbitrary cues. *Animal Cognition, 14,* 871–878. http://dx.doi.org/10.1007/s10071-011-0421-6

Heyes, C. (2000). Evolutionary psychology in the round. In C. M. Heyes & L. Huber (Eds.), *The evolution of cognition* (pp. 3–23). Cambridge, MA: MIT Press.

Heyes, C. (2012). Simple minds: A qualified defence of associative learning. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 367,* 2695–2703. http://dx.doi.org/10.1098/rstb.2012.0217

Hood, B., Hauser, M., Anderson, L., & Santos, L. (1999). Gravity biases in a non-human primate? *Developmental Science, 2,* 35–41. http://dx.doi.org/10.1111/1467-7687.00051

Karin-D'Arcy, M. R. (2005). The modern role of Morgan's canon in comparative psychology. *International Journal of Comparative Psychology, 18,* 179–201.

Lagnado. (2009). A causal framework for integrating learning and reasoning [Peer commentary on "The propositional nature of human associative learning" by C. Mitchell, J. De Houwer & P. Lovibond]. *Behavioral and Brain Sciences, 32,* 211–212.

Lagnado, D. A., Waldmann, M. R., Hagmayer, Y., & Sloman, S. A. (2007). Beyond covariation: Cues to causal structure. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 154–172). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195176803.003.0011

Locke, J. (2007). *An essay concerning human understanding*. Pomona, Italy: Pomona Press.

Matute, H., & Vadillo, M. A. (2009). The Proust effect and the evolution of a dual learning system. [Peer commentary on "The propositional nature of human associative learning" by C. Mitchell, J. De Houwer & P. Lovibond]. *Behavioral and Brain Sciences, 32,* 215–216. http://dx.doi.org/10.1017/S0140525X09001046

Meketa, I.. (2014). A critique of the principle for cognitive simplicity in comparative cognitzin. *Biological Philosophy, 29,* 731–745.

Mikhalevich, I. (2014). Experiment and animal minds: Why statistical choices matter. In: Philosophy of Science Association 24th Biennial Meeting, Chicago, Illinois.

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences, 32,* 183–198. http://dx.doi.org/10.1017/S0140525X09000855

Morgan, C. L. (1894). *An introduction to comparative psychology*. London, England: Walter Scott. http://dx.doi.org/10.1037/11344-000

Morsella, E., Riddle, T. A., & Bargh, J. A. (2009). Undermining the foundations: Questioning the basic notions of associationism and mental representation [Peer commentary on "The propositional nature of human associative learning" by C. Mitchell, J. De Houwer & P. Lovibond]. *Behavioral and Brain Sciences, 32,* 218–219. http://dx.doi.org/10.1017/S0140525X09001071

Pearce, J. (2008). *Animal learning and cognition: An introduction* (3rd ed.). Hove, United Kingdom: Psychology Press.

Pearce, J. M., & Bouton, M. E. (2001). Theories of associative learning in animals. *Annual Review of Psychology, 52,* 111–139. http://dx.doi.org/10.1146/annurev.psych.52.1.111

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review, 87,* 532–552. http://dx.doi.org/10.1037/0033-295X.87.6.532

Pearl, J., & Russell, S. (2001). Bayesian networks. In M. Arbib (Ed.), *Handbook of brain theory and neural networks*. Cambridge, MA: MIT Press.

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences, 31,* 109–130. http://dx.doi.org/10.1017/S0140525X08003543

Penn, D. C., & Povinelli, D. J. (2007). Causal cognition in human and nonhuman animals: A comparative, critical review. *Annual Review of Psychology, 58,* 97–118. http://dx.doi.org/10.1146/annurev.psych.58.110405.085555

Povinelli, D. (2011). *World without weight: Perspectives on an alien mind.* New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780198570967.001.0001

Povinelli, D. J., & Vonk, J. (2003). Chimpanzee minds: Suspiciously human? *Trends in Cognitive Sciences, 7,* 157–160. http://dx.doi.org/10.1016/S1364-6613(03)00053-6

Povinelli, D. J., & Vonk, J. (2004). We don't need a microscope to explore the chimpanzee's mind. *Mind & Language, 19,* 1–28.

Premack, D., & Premack, A. J. (1994). Levels of causal understanding in chimpanzees and children. *Cognition, 50,* 347–362. http://dx.doi.org/10.1016/0010-0277(94)90035-3

Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology, 66,* 1–5. http://dx.doi.org/10.1037/h0025984

Rescorla, R. A. (1988). Pavlovian conditioning. It's not what you think it is. *Journal of Experimental Psychology: Animal Behavior Processes, 43,* 151–160. http://dx.doi.org/10.1037/0003-066X.43.3.151

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York, NY: Appleton-Century-Crofts.

Romanes, G. (1970). *Animal intelligence* (Reprint of 1882 editorial published by Kegan Paul, Trench & Co., London ed.). Farnborough, England: Gregg International.

Schrauf, C., & Call, J. (2009). Great apes' performance in discriminating weight and achromatic color. *Animal Cognition, 12,* 567–574. http://dx.doi.org/10.1007/s10071-009-0216-1

Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience, 23,* 473–500. http://dx.doi.org/10.1146/annurev.neuro.23.1.473

Seed, A., & Call, J. (2010). Physical problem solving in tool-using and non-tool-using animals. In J. Moore & M. Breed (Eds.), *Encyclopaedia of Animal Behaviour* (Vol. 2, pp. 778–785). Oxford, United Kingdom: Elsevier Academic Press. http://dx.doi.org/10.1016/B978-0-08-045337-8.00090-5

Seed, A., Hanus, D., & Call, J. (2011). Causal knowledge in apes and children: More than meets the eye? In T. McCormack, C. Hoerl, & S. Butterfill (Eds.), *Tool use and causal cognition* (pp. 89–111). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199571154.003.0005

Shanks. (2009). The associative nature of human associative learning [Peer commentary on "The propositional nature of human associative learning" by C. Mitchell, J. De Houwer & P. Lovibond]. *Behavioral and Brain Sciences, 32,* 225–226.

Shettleworth, S. J. (2010a). *Cognition, evolution, and behavior* (2nd ed.). New York, NY: Oxford University Press.

Skinner, B. F. (1953). *Science and human behavior.* New York, NY: Macmillan.

Sober, E. (2005). Comparative psychology meets evolutionary biology: Morgan's canon and cladistic parsimony. In L. Datson & G. Mitman (Eds.), *Thinking with animals: New perspectives on anthropomorphism* (pp. 85–99). New York, NY: Columbia University Press.

Sober, E. (2009). Parsimony and models of animal minds. In R. Lurz (Ed.), *Philosophy of animal minds* (pp. 237–257). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511819001.014

Taylor, A. H. (2014). Corvid cognition. *Wiley Interdisciplinary Reviews: Cognitive Science, 5,* 361–372. http://dx.doi.org/10.1002/wcs.1286

Thorndike, E. (1911). *Animal intelligence.* New York, NY: Macmillan.

Thorndike, E. (1898). Animal intelligence: An experimental study of the association processes in animals. *Psychological Review: Monograph Supplements, 2,* i–109. http://dx.doi.org/10.1037/h0092987

Tolman, E. C. (1932). *Purposive behavior in animals and men.* New York, NY: Century.

Tomasello, M. (2014). *A natural history of human thinking.* Cambridge, MA: Harvard University Press. http://dx.doi.org/10.4159/9780674726369

Tomasello, M., & Call, J. (2006). Do chimpanzees know what others see—or only what they are looking at? In S. Hurley & M. Nudds (Eds.), *Rational animals?* (pp. 371–384). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780198528272.003.0017

Tomonaga, M., Imura, T., Mizuno, Y., & Tanaka, M. (2007). Gravity bias in young and adult chimpanzees (*Pan troglodytes*): Tests with a modified opaque-tubes task. *Developmental Science, 10,* 411–421. http://dx.doi.org/10.1111/j.1467-7687.2007.00594.x

Tooby, J., & Cosmides, L. (2005). Conceptual foundations of evolutionary psychology. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 5–67). Hoboken, NY: Wiley.

Van Hamme, L., & Wasserman, E. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning and Motivation, 25,* 127–151. http://dx.doi.org/10.1006/lmot.1994.1008

Vonk, J., & Povinelli, D. (2006). Similarity and difference in the conceptual systems of primates: The unobservability hypothesis. In E. M. Wassermann & T. R. Zentall (Eds.), *Comparative cognition: Experimental explorations of animal intelligence* (pp. 363–387). New York, NY: Oxford University Press.

Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.) *Information processing in animals: Memory mechanisms* (pp. 5–47). Hillsdale, NJ: Lawrence Erlbaum Associates.

Waldmann, M. R. (2000). Competition among causes but not effects in predictive and diagnostic learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 53–76. http://dx.doi.org/10.1037/0278-7393.26.1.53

Waldmann, M. R., Cheng, P. W., Hagmayer, Y., & Blaisdell, A. P. (2008). Causal learning in rats and humans: A minimal rational model. In N. Chater & M. Oaksford (Eds.), *The probabilistic mind: Prospects for Bayesian Cognitive Science* (pp. 453–484). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199216093.003.0020

Waldmann, M., Hagmayer, Y., & Blaisdell, A. (2006). Beyond the information given: Causal models in learning and reasoning. *Current Directions in Psychological Science, 15,* 307–311. http://dx.doi.org/10.1111/j.1467-8721.2006.00458.x

Waldmann, M. R., Schmid, M., Wong, J., & Blaisdell, A. P. (2012). Rats distinguish between absence of events and lack of evidence in contingency learning. *Animal Cognition, 15,* 979–990. http://dx.doi.org/10.1007/s10071-012-0524-8

Wasserman, E. A., & Castro, L. (2005). Surprise and change: Variations in the strength of present and absent cues in causal learning. *Learning & Behavior, 33,* 131–146. http://dx.doi.org/10.3758/BF03196058

Watson, J. (1928). *The ways of behaviorism.* New York, NY: Harper.