

COMMENTARY

Mechanism Through Methodology: No Madness to the Method

Aaron P. Blaisdell
University of California, Los Angeles

Grau and Joynes (2005) assess the current state of the field of animal learning and behavior, with particular emphasis on pedagogical and curricular issues. They suggest that the conventional framework which organizes lecture material around methodology is flawed and that an organization around mechanism should be used instead. They also advocate a shift from a purely behavioral approach to research on learning and behavior to a neural-functionalist approach more akin to contemporary behavioral neuroscience. While I support many of the suggestions for improving instruction, I disagree with their proposed shift away from purely behavioral investigations of animal behavior. Behavioral research continues to be a thriving and productive source of empirical and theoretical discoveries. The diverse array of specialized methodologies that have been developed to pursue this work are still paying dividends by illuminating the nature of behavioral mechanisms. Banishing purely behavioral approaches to learning and behavior, such as those used to study associative learning, animal cognition, and comparative psychology, would severely hamper our knowledge of behavioral mechanism.

“A complete theory of learning must speak to all of the ways in which experience can alter behavior...” (Grau & Joynes, 2005, p. 15). This statement resonates the views espoused by most contemporary researchers of learning and behavior. Grau and his students have published excellent work on the neural basis of learning in spinal rats. This research is exemplary in its use of rigorous methodology and the nuanced appreciation of the theoretical issues it raises. Thus, they are in an excellent position to comment on the state of the field of learning.

Grau and Joynes (2005) address a number of important issues that are germane to the study of learning and behavior. Most importantly, they provide a timely critique of the current state of pedagogy and curriculum. They delineate inherent problems with the manner in which theory and research are portrayed in standard textbooks for academic courses on learning. The conventional view, as Grau and Joynes see it, adopted by textbooks organizes information around the methodologies used to understand learning rather than around the mechanisms of learning. One problem with a methodological approach is that it fails to provide a coherent theme or framework in which to connect all of the disparate findings and theories. Thus, issues relevant to both Pavlovian and instrumental learning, for example, are presented in separate chapters with little integration into a larger, coherent framework. This disconnection between facts and framework probably contributes significantly to the tendency for students to perceive courses on learning as difficult or uninteresting. The conventional framework produces another negative consequence: neuroscientists who wish to study the neural mechanisms of learning often receive an antiquated (at best) or misguided (at worst) understanding of

Support was provided by NIMH Grant MH66855 (A.P. Blaisdell). Requests for reprints should be addressed to Aaron P. Blaisdell, UCLA Department of Psychology, 1285 Franz Hall, Box 951563, Los Angeles, CA, 90095-1563, U.S.A. (blaisdell@psych.ucla.edu).

learning. Imagine the deficiency in knowledge of a medical student who is taught gross anatomy and physiology, but not molecular biology or genetics. As a result, with a few notable exceptions, many neuroscientists are unaware of the progress that has been made over the past few decades in our understanding of the processes and mechanisms of learning.

Another issue that Grau and Joynes (2005) attack is the problem created by defining learning too narrowly as associative. Such a definition relegates nonassociative processes as unimportant to the study of learning. As a result, attempts to control for many alternative (to associative) causes for behavioral effects have led to the ignorance or even vilification of important processes and mechanisms of learning. For example, nonassociative processes such as habituation, sensitization, and alpha responding are typically deemed unimportant or are perceived to interfere with an investigation of mechanisms of “true” (i.e., associative) learning. However, as Grau and Joynes elegantly demonstrate in spinal rats, these nonassociative processes may play an important functional role in the plasticity of behavior. They, along with many contemporary psychologists, view each learning phenomenon studied in the laboratory as part of a behavioral system. A behavioral system consists of a set of dynamically interrelated behaviors that have been adapted to solve important biological problems. Thus rather than trying to rule out nonassociative factors as annoying experimental artifacts, they assert that nonassociative processes “can be seen as an example of biological ingenuity rather than an experimental anomaly” (p. 15). The behavioral systems approach is a continuation, or rediscovery, of comparative psychology (Papini, 2002). Others in the field of experimental animal psychology have adopted similar views. For example, my own research has emphasized the learning-performance distinction, and has been motivated by the view that performance variables are themselves important mechanisms that warrant study if we are to fully understand learning phenomena (Blaisdell, 2003; Denniston, Savastano, & Miller, 2001).

Although I concur with many of the points raised by Grau and Joynes (2005), I am concerned that they go too far in disparaging the methodological approach to learning and behavior. Methodologies have been developed to study behavioral processes with the goal of discovering mechanisms at both the psychological and biological levels of analysis. While it is true that methodologies are tools and not end products, many methodologies are needed to properly analyze and dissect the behavioral phenomena to extract converging evidence for mechanism. For example, fear and appetitive conditioning procedures each hold advantages and disadvantages for studying various aspects of Pavlovian processes, such as interval timing. On the one hand, footshock is a very precisely controllable US, allowing virtually complete control over how the subject experiences the CS-US interval. On the other hand, procedures using food reward have allowed for precise estimations of temporal expectancies using psychophysical techniques. The combination of both procedures has enhanced our understanding of processes of interval timing beyond the limits of each individual procedure. Similar arguments can be made for the use of multiple methodologies to understand many other behavioral phenomena, such as stimulus competition (also known as retrospective revaluation), mechanisms of short-term and long-term retention, the spatial and contextual control of behavior, and configural learning.

Grau and Joynes (2005) advocate a neural-functionalist approach in which more traditional studies of learning are replaced with systems-level behavioral neuroscience. However, the removal of purely behavioral studies from psychological research on learning, motivation, and cognition would severely restrict what we could understand about behavioral mechanisms. A cursory flip through any current issue of the *Journal of Experimental Psychology: Animal Behavior Processes*, *Learning & Behavior*, *Learning and Motivation*, *Quarterly Journal of Experimental Psychology, B*, *Journal for the Experimental Analysis of Behavior*, *Behavioral Processes*, *International Journal of Comparative Psychology*, or *Animal Cognition*, reveals a rich and active study of psychology at the behavioral level. For example, we are constantly advancing our understanding of Pavlovian processes through the investigation of such diverse topics as renewal of fear conditioning (Bouton & Bolles, 1979), conditioned flavor and odor aversions and preferences (Batson & Batsell, 2000; Lubow & De la Casa, 2002), appetitive conditioning (Rescorla, 2002), autoshaping (Killeen, 2003), human contingency and causal judgments (Dickinson & Burke, 1996; Matute & Miller, 1998; Van Hamme & Wasserman, 1994), and spatial and temporal relations (Blaisdell & Cook, 2005; Savastano & Miller, 1998). Although grouped by common or similar methodologies, all of these investigations tell us something about the nature of learning and performance mechanisms involved in learning about stimulus-stimulus relations. This may lead some, such as Grau and Joynes to adopt the view that “Pavlovian conditioning refers to a class of methods designed to investigate how organisms encode stimulus relationships within the environment” (p. 9). Perhaps it is more accurate to say that Pavlovian conditioning refers to a class of behavioral phenomena and the investigation of the mechanisms that support it. A set of methodologies have been developed for this purpose, and new methodologies are continually being developed. However, the focus is on the behavioral phenomena and the underlying mechanisms, at various levels of analysis. There is already a rich and thriving field of behavioral neuroscience that is very close to the neural-functionalist framework proposed by Grau and Joynes. To do away with purely behavioral approaches to learning and behavior and related disciplines, such as animal cognition and comparative psychology, endorses an exclusive rather than inclusive philosophy.

What ramifications does an inclusive approach to the study of learning mechanisms have for Grau and Joynes’ (2005) work on spinal rats? In their preparation, the connection between the spinal cord and the brain in the rat is cut. To study Pavlovian conditioning, the spinal rat is given pairings between a mild electrical stimulus (CS+) applied to one leg followed by a strong shock to the tail, interspersed with a mild electrical CS- to the other leg unpaired with strong tail shock (Grau, Salinas, Illich, & Meagher, 1990). This training establishes stronger responding to the CS+ than to the CS-. However, a novel stimulus (CS neutral) that has not undergone any conditioning elicits the same magnitude response as does the CS+ (Joynes & Grau, 1996). This suggests that spinal cord neurons do not support Pavlovian conditioning, but rather, that the response to the CS+ is due to protection from habituation. That is, the presence of the US during training prevents the *unconditioned* response to the CS+ from habituating, rather than establishing a *conditioned* response to the CS+. The empirical effect is well documented and noteworthy. Learning (habituation) occurs in the control (unpaired) group, but not in the experimental group. Grau and Joynes conclude that “Such comparisons

highlight the importance of appropriate controls (e.g., the inclusion of a novel CS) and how a mechanistically-based view of learning must remain wedded to rigorous methodology” (p. 11-12). This is a valid point. But perhaps they are a bit premature in their final assessment of spinal mechanisms of learning. They only test the CS+ relative to a novel CS within an hour after the training episode has ended. Therefore, we cannot determine from these data whether there are any long-term effects of the CS-US experience. More controls and tests are needed. For example, what would be the outcome of tests of the CS+ and the novel CS given 24 h (or longer) after the end of training? Perhaps the CS+ would exhibit a stronger response than the novel CS. That is, the rate of habituation to the CS+ might be slower than the rate of habituation to a novel CS. If so, then this would be evidence for the acquisition of a CS-US association (i.e., Pavlovian conditioning) that is stored in the spinal cord. Whatever the eventual resolution, it is important to realize that support of *either* theoretical interpretation should be of interest to students of learning. Instead of denigrating a protection-from-habituation mechanism for masquerading as Pavlovian conditioning, we should embrace it as an interesting neural solution to an important behavioral problem. As such, learning in the spinal cord, does not provide “a good example of the breakdown between methodology and mechanism” (p. 7). It serves as a good example of how methodology (along with carefully designed control manipulations) can illuminate mechanism.

A final point of contention I wish to address is with the portrayal of the Learning-and-Behavior traditionalist as a straw man. While the prevailing view used to be that associative learning was relegated as the only *true* type of learning, it is very difficult to find anyone today who holds tenaciously to this opinion. According to Grau and Joynes (2005), the traditionalist would discount the work of Kandel and his associates (Kandel & Schwartz, 1982) because learning in *Aplysia* is arguably a form of alpha conditioning. I agree that it is interesting to study the neural mechanisms of nonassociative plasticity in *Aplysia*. There are, however, invertebrate systems in which *true* Pavlovian conditioning is demonstrable. For example, repeated pairings between a light CS, which normally elicits approach in the mollusk *Hermissenda*, and a rotation US, which normally reduces the rate of locomotion, results in a conditioned response of reduced locomotion to the light CS (Rogers, Talk, & Matzel, 1994). The CR elicited by the CS is opposite to the UR prior to training, and therefore, cannot be attributed to alpha responding. Perhaps the gill-withdrawal reflex in *Aplysia* was a poor choice of behavior system in which to investigate the neurobiology of associative learning (though this does not devalue its utility for the study of other forms of behavioral plasticity). But this does not mean that this type of learning cannot be studied in simple systems. Both types of learning are important and are worthy of study.

References

- Batson, J. D., & Batsell, W. R., Jr. (2000). Augmentation, not blocking, in an A+/AX+ flavor-conditioning procedure. *Psychonomic Bulletin Review*, *7*, 466-471.
- Blaisdell, A. P. (2003). The S-R information stream: Where's the filter? *Integrative Physiological and Behavioral Science*, *38*, 146-165.
- Blaisdell, A. P., & Cook, R. G. (2005). Integration of spatial maps in pigeons. *Animal Cognition*, *8*, 7-16.

- Bouton, M. E., & Bolles, R. C. (1979). Role of conditioned contextual stimuli in reinstatement of extinguished fear. *Journal of Experimental Psychology: Animal Behavior Processes*, **5**, 368-378.
- Denniston, J. C., Savastano, H. I., & Miller, R. R. (2001). The extended comparator hypothesis: Learning by contiguity, responding by relative strength. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 65-117). Mahwah, NJ: Erlbaum.
- Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective reevaluation of causality judgements. *Quarterly Journal of Experimental Psychology*, **49B**, 60-80.
- Grau, J. W., & Joynes, R. L. (2005). A neural-functionalist approach to learning. *International Journal of Comparative Psychology*, **18**, 1-22.
- Grau, J. W., Salinas, J. A., Ilich, P. A., & Meagher, M. W. (1990). Associative learning and memory for an antinociceptive response in the spinalized rat. *Behavioral Neuroscience*, **104**, 489-494.
- Joynes, R. L., & Grau, J. W. (1996). Mechanisms of Pavlovian Conditioning: Role of protection from habituation in spinal conditioning. *Behavioral Neuroscience*, **110**, 1375-1387.
- Kandel, E. R., & Schwartz, J. H. (1982). Molecular biology of learning: modulation of transmitter release. *Science*, **218**, 433-443.
- Killeen, P. R. (2003). Complex dynamic processes in sign tracking with an omission contingency (negative automaintenance). *Journal of Experimental Psychology: Animal Behavior Processes*, **29**, 49-61.
- Lubow, R. E., & De la Casa, L. G. (2002). Superlatent inhibition and spontaneous recovery: differential effects of pre- and postconditioning CS-alone presentations after long delays in different contexts. *Animal Learning and Behavior*, **30**, 376-386.
- Matute, H., & Miller, R. R. (1998). Detecting causal relations. In W. T. O'Donohue (Ed.), *Learning and behavior therapy* (pp. 483-497). Needham Heights, MA: Allyn & Bacon, Inc.
- Papini, M. R. (2002). *Comparative psychology: Evolution and development of behavior*. Upper Saddle River, NJ: Prentice Hall.
- Rescorla, R. A. (2002). Comparison of the rates of associative change during acquisition and extinction. *Journal of Experimental Psychology: Animal Behavior Processes*, **28**, 406-415.
- Rogers, R. F., Talk, A. C., & Matzel, L. D. (1994). Trial-spacing effects in Hermissenda suggest contributions of associative and nonassociative cellular mechanisms. *Behavioral Neuroscience*, **108**, 1030-1042.
- Savastano, H. I., & Miller, R. R. (1998). Time as content in Pavlovian conditioning. *Behavioural Processes*, **44**, 147-162.
- Van Hamme, L. J., & Wasserman, E. A. (1994). Cue competition in causality judgments: The role of nonrepresentation of compound stimulus elements. *Learning and Motivation*, **25**, 127-151.

Received May 6, 2004.

Revision received June 22, 2004.

Accepted June 22, 2004.

COMMENTARY

Experimental Methods and Conceptual Confusion

Armando Machado
Universidade do Minho, Portugal

According to Grau and Joynes (2005), (1) the current classification of types of learning is based on methodology and assumes a correspondence between types of learning and distinct neural-functional mechanisms; (2) this assumption is wrong because experiments show that different mechanisms may underlie the same type of learning; consequently, (3) we should change the teaching of the psychology of learning. I argue that because Grau and Joynes misunderstood the nature of the classification of learning phenomena and cloaked their research findings with a garb of conceptual errors and infelicities, their recommendations concerning the teaching of learning should be rejected.

Grau and Joynes' (2005) overall argument in "A neural-functionalist approach to learning" may be divided into four parts. First, the current classification of types or kinds of learning (single stimulus learning, Pavlovian conditioning, and instrumental conditioning) is based on methodology. Second, this classification assumes a correspondence between each kind of learning defined methodologically and a distinct kind of neural-functional mechanism. Third, experiments show that the assumption of correspondence is incorrect because a variety of different mechanisms may underlie the same kind of learning when the latter is defined methodologically. Finally, given the empirical inadequacy of a classification based on methodology, as well as some of its additional negative consequences (e.g., a too restrictive view of learning), it should be replaced by a classification based on mechanisms. In what follows, I will argue that Grau and Joynes (1) misunderstood the nature of the classification of learning phenomena; (2) misconstrued the global significance of their own empirical findings; (3) mistook psychologists of learning for (some) neuroscientists when attributing the correspondence assumption; and, more generally, (4) cloaked their *important* empirical contributions with a garb of conceptual errors and infelicities that their recommendations concerning the teaching of learning should simply be rejected.

On the Classification of Learning Phenomena

According to Grau and Joynes, "researchers within the field of learning have traditionally divided their empirical world according to methodology, with phenomena classified as single stimulus learning, Pavlovian conditioning, or instrumental learning" (p. 1, Abstract). Thus, the "phenomena were classified together, not because they share a common underlying mechanism (the machinery that underlies learning), but rather because a similar methodology was used to infer

This work was supported by a grant from the Portuguese Science and Technology Foundation (FCT). I thank Francisco Silva, Marco Vasconcelos, Paulo Pata, and Paulo Rodrigues for helpful comments on earlier versions of this paper. Address correspondence to Armando Machado, Instituto de Educação e Psicologia, Universidade do Minho, 4710 Braga, Portugal (Armandom@iep.uminho.pt).

their presence” (p. 2). Readers may search in Grau and Joynes’ paper for a clear statement of what “methodology” means in this particular context, for the sense in which habituation/sensitization, Pavlovian and operant conditioning are *methodological* categories; they may search, but they will not find it.

Given this glaring omission, I can only conjecture: Perhaps the authors mean a classification based on the procedures implemented in the laboratory, on the kinds of events manipulated explicitly by investigators and often described in the Procedure section of experimental articles. But if this is the case, then Grau and Joynes have misunderstood the nature of the classification. For whether an instance of learning examined in the laboratory is classified as, for example, habituation does not depend exclusively on what the experimenter does (call it S), but also on how the animal reacts to it (call it R) and on the nature of the dynamic relations holding between the two (call it S-R). A rat startles when hearing a brief, loud noise. When the same noise is repeated, say, every 15 s, the startle response typically wanes and may even cease (R). If the response decrement is due to the repetition of the stimulus (S-R) then psychologists will refer to this instance of learning as habituation. All three elements—S, R, and S-R—are important to classify the phenomena and two of them, R and S-R, are not at the experimenter’s disposal; they tell us something about the animal.

Consider a second example. For a couple of hours, a male juvenile cowbird sings hundreds of songs in front a female but is consistently ignored by her. Then during one of his songs the female moves her right wing, a display called wing stroking. Subsequently, the male sings that song more often than before and that song becomes more effective at eliciting copulation postures from females than other songs (e.g., West & King, 1988). A psychologist of learning would interpret the male cowbird’s behavior as follows: Its response of singing a particular song (R) was followed by a particular female reaction (S), and for that reason it occurred more often in the future. If this interpretation is correct, then the change in the male’s singing would be classified as an instance of operant conditioning. Although the interpretation could be wrong—perhaps the male cowbird sang the song more often because the female’s wing stroking simply elicited it—the point of the example is to show that more than methodology is involved in the classification. The change in the male’s singing behavior and the causative factors responsible for it are critical to the classification.

In the same vein, one could say that operant conditioning is not about giving M&Ms to kids, presumably the methodological criterion in Grau and Joynes’ terminology, but about giving them as a consequence of some behavior (the procedure) and observing changes in that behavior (the process). Operant conditioning is a form of behavioral change defined by both procedure and process. More generally, then, operant and Pavlovian conditioning as well as habituation and sensitization describe broad kinds of interactions between an animal and its circumstances. As such they are necessarily defined by the two endpoints of the interaction (abstractly represented by S and R) and by the properties of their relation (S-R). As Keller and Schoenfeld (1950/1995) remarked, “we need a dynamic, rather than a static picture of the behavior of organisms. To describe process, not to inventory elements, is [and has been, I add] our major concern” (p. 7).

On Grau and Joynes' Neural-Functional Mechanisms

According to Grau and Joynes, the current classification of learning phenomena is associated with the “simplistic” (p. 3) assumption of an isomorphism between methodology and mechanism. “One-to-one correspondence” would be a more accurate term than isomorphism, for the assumption is that distinct neural mechanisms underlie distinct methodologically-defined kinds of learning. Be that as it may be, according to Grau and Joynes the isomorphism assumption is contrary to fact: Animals can learn about conditional and unconditional stimulus relations or about responses and their outcomes in multiple ways that engage different neural-functional mechanisms. As a consequence, the authors propose a classification of types of learning based on the variety of underlying mechanisms. This new arrangement of the field of learning provides “a better framework for linking behavior to neuroscience and cognition” (p. 4), and it also promotes a richer, less restrictive conception of learning than the old one.

Although Grau and Joynes speak repeatedly about neural mechanisms, the research they summarize does not deal with them directly. That is, the authors do not identify the specific neural structures involved during learning nor describe their interactions. Instead, they define the putative mechanisms of learning functionally; hence the name they give to their approach, neurofunctionalism. To illustrate, having established conditioned antinociception in a rat's spinal cord preparation, the authors tried “to uncover how this learning was accomplished—the *mechanism* that underlies this particular example of Pavlovian conditioning” (p. 9, italics added). They advanced three possibilities: the formation of a CS-US association, paired specific enhanced sensitization, and protection from habituation. The results of some manipulations (e.g., increase the number of CS-US pairings or the intertrial interval) led them to favor the last possibility. In summary, according to the authors, the rat's spinal cord was able to abstract the CS-US relation through the mechanism of protection from habituation. They conclude: “The methods of Pavlov provide great tools, but that is all they are. Having established the usefulness of the tools, we need to understand how systems sensitive to Pavlovian relations operate, realizing that other tools may be used to uncover the underlying machinery. Our central concern is with how the learning engine works [e.g., protection from habituation], not the tools used to take it apart [e.g., Pavlovian conditioning]” (p. 16).

I will argue that Grau and Joynes' argument suffers from a fundamental conceptual error, a sort of categorical mistake. Ironically, because of that error, the authors misconstrued their own research findings. Consider the three possibilities advanced to illustrate the mechanisms underlying Pavlovian conditioning: protection from habituation, paired specific enhanced sensitization, and the formation of an association. The authors believe these are functional descriptions and that they are somewhat better than the methodological description Pavlovian conditioning. More to the point, their key argument is that the broad (methodologically-based) category Pavlovian conditioning should be replaced by functional or mechanistically-based categories such as protection from habituation, for example. But the authors failed to notice that protection from habituation, paired specific enhanced sensitization, and the formation of an association are at best subcategories of Pavlovian conditioning; they identify in greater detail how the CS and US arrange-

ments in a particular case affected behavior; they do not show the inconsistency, irrelevance, or the misleading nature of the superordinate category Pavlovian conditioning. On the contrary, they depend on it for their own intelligibility. To remove the superordinate category Pavlovian conditioning is like removing the category *Homo* and then trying to understand the relations among *sapiens*, *erectus*, *neanderthalensis*, etc.

The conceptual error is further evinced when Grau and Joynes refer to the superordinate Pavlovian conditioning as a methodologically-defined category but refer to protection from habituation, a subordinate category, as a functional, mechanistically-defined category. I showed above that the former reference is incorrect. I will show next that the latter reference also is incorrect. If we define Pavlovian conditioning as a pairing of two stimuli (e.g., CS = tone and US = food) such that the response (e.g., a dog's salivation) elicited by the US is "transferred" to the CS, then we see that this definition (or any other similar one for that matter) involves two things: a particular arrangement of stimuli and a particular change in the animal's behavior. Consider now protection from habituation: A stimulus elicits a response that decreases when the stimulus is repeated. However, if the stimulus is paired with another stimulus (the US), then the response will not decrease when the former stimulus (call it CS) is repeated. We may say, picturesquely, that the US protected the CS from habituation. But what is the difference in the *logical* status of the two definitions? Both include an environmental manipulation (procedure) and a behavioral effect (process). In this regard, they are both defined methodologically or functionally or mechanistically (in the authors' phraseology).

Furthermore, among the three possibilities advanced to account for conditioned antinociception, the authors came to favor protection from habituation. They were influenced by a series of tests that elucidated the dynamic properties of the behavior under scrutiny (e.g., the effects of increasing the number of trials). But note that initially they had to ascertain that the conditioned antinociception itself was an instance of Pavlovian conditioning. How did they do it? Precisely in the same manner they did to ascertain that their preparation was an instance of protection from habituation: They used a variety of tests that revealed the dynamic properties of the behavior in question, properties typically associated with Pavlovian conditioning (e.g., overshadowing, blocking). The authors concluded: "Together, we believe that these observations suggest that the phenomenon is reasonably classified as an instance of Pavlovian conditioning" (p. 9). Strangely enough, following the *same* logic, in one case the authors reached a mechanistically-defined category (protection from habituation) whereas in the other case they reached a methodologically-defined category (Pavlovian conditioning). And they suggest that only the former has "biological reality" (p. 2).

Consider the foregoing issue from yet another perspective. Suppose Grau and Joynes were familiar with habituation in a specific context—when stimulus S is repeated, response R typically wanes. One day these researchers encounter a new situation—when a distinct stimulus is also present, the response to S does not habituate. They name the new phenomenon protection from habituation. However, because they are not satisfied with methodological classifications, they search for the underlying functional mechanism. Two possibilities come to their minds, one in which protection from habituation is due to the animal's state of arousal induced by the new stimulus, and another in which it is due to the pairing of S with the new

stimulus. A series of experiments with appropriate control conditions (e.g., in one the arousal state was induced by presenting the new stimulus alone; in another the interval between the two stimuli was varied considerably) convinces the two authors that their case of protection from habituation is due to stimulus pairing, not to arousal state. Grau and Joynes conclude that underlying the methodologically-defined protection from habituation is the functionally-defined Pavlovian conditioning, the opposite of their conclusion in the antinociception case! The message is clear: One should be suspicious of a way of reasoning that reaches opposite classifications from the same set of facts.

On the Assumption of Isomorphism

Concerning this assumption, I question whether any learning psychologist embraces it nowadays. (Moreover, did any learning psychologist ever believe that in protozoans and humans habituation is implemented by one and the same mechanism? Although the answer may seem obvious, before giving it one still needs a clear, working description of what “mechanism” means in this context. Grau and Joynes provide none.) Instead, I believe most remain silent on this issue, as they should, because no neural mechanism is likely to be identifiable on the basis of behavioral data exclusively (Uttal, 1998). In fact, when Grau and Joynes criticize the assumption of isomorphism they always seem to have in mind a neuroscientist, not a learning psychologist: “For *neurobiologists*, the traditional view encourages an elegantly simple linking hypothesis that couples learning about distinct environmental relations (defined by methodology) to particular biological mechanisms” (p. 2; italics added); “*neurobiologists* favor paradigms blessed as true 30 years ago” (p. 3; italics added); “*neurobiologists* assume that any preparation that demonstrates a sensitivity to a response-outcome relation can be used to elucidate the mechanisms that underlie operant learning” (p. 15; italics added). Perhaps Grau and Joynes have simply mistaken the identity of their opponent.

Conceptual Infelicities and the Teaching of Learning

I found the remaining key arguments in Grau and Joynes’ paper unconvincing because the authors were either vague or inaccurate when clarity and precision were called for. For example, it is impossible to agree or disagree with vague statements such as the following: “Seeking scientific legitimacy based on observable events, learning psychologists focused on stimuli and responses, the ‘venerable S and R’... [but] somehow the events themselves grew in stature and came to dominate our thinking about learning phenomena” (p. 2). Unfortunately, they do not tell the reader what it means for stimuli and responses to grow in stature. Similarly, it is impossible to agree or disagree with a conclusion when one does not know how key inferences were drawn: “At conferences, we and our colleagues will make light of the difficulty of maintaining student attention through a month of Pavlovian conditioning. The sad fact is that this organization, a historical artifact from our behaviorist past, leads students to a mistaken view—that the field of learning has become constricted, perhaps stagnant, and is out of tune with modern developments.” (p. 17) But how or why do students conclude that the field of learning is stagnant based on the organization of the material? How or why would

an alternative classification based on mechanism eliminate “the sad fact”? And how or why should we believe that student enthusiasm and learning would be enhanced if protection from habituation, paired specific enhanced sensitization, and similar categories replaced Pavlovian conditioning and operant conditioning? And when Grau and Joynes were more specific they often misrepresented the field. Thus when they state that “the vestiges of the traditional view have led researchers to ignore behavioral/biological mechanisms that play a pervasive role in helping an organism adjust to new environmental relations” (in reference to non-associative forms of learning) they ignore the behavioral research on habituation/sensitization and the attempts to “use” habituation to explain apparently more complex forms of learning and behavior (e.g., contrast effects in multiple schedules, McSweeney, Hinson, & Cannon, 1996; interval timing, Staddon, 2001).

As for the current teaching of the psychology of learning, the crux of the problem for Grau and Joynes seems to be the emphasis on behavior: “Researchers have learned a tremendous amount about the neurobiological mechanisms that underlie learning, but only recently has this material begun to creep into our texts. Relative to behavioral studies on learning, there are far more reports, both at conferences and in our journals, on the neurobiological mechanisms of learning. Yet, behavior has remained the central focus within the classroom” (p. 17). But in their claim for a privileged place in the classroom and textbooks for the neuroscience of learning at the cost of the psychology of learning, the authors express no scientific or pedagogical wisdom, only their peculiar biases.

In conclusion, there is much that is good and much that is novel in Grau and Joynes’ paper. But what is good is not novel and, alas, what is novel is not good. The authors seem exhausted with the teaching of learning and mistakenly conclude that the psychology of learning itself is exhausted. It is not.

References

- Grau, J. W., & Joynes, R. L. (2005). A neural-functionalist approach to learning. *International Journal of Comparative Psychology*, **18**, 1-22.
- Keller, F., & Schoenfeld, N. (1950/1995). *Principles of psychology*. Cambridge, MA: B. F. Skinner Foundation.
- McSweeney, F. K., Hinson, J. M., & Cannon, C. B. (1996). Sensitization-habituation may occur during operant conditioning. *Psychological Bulletin*, **120**, 256-271.
- Staddon, J. E. R. (2001). *Adaptive dynamics*. Cambridge, MA: MIT Press.
- West, M. J., & King, A. P. (1988). Female visual displays affect the development of male song in the cowbird. *Nature*, **334**, 244-246.
- Uttal, W. (1998). *Toward a new behaviorism: The case against perceptual reductionism*. Mahwah, NJ: Erlbaum.

Received April 5, 2004.

Revision received June 22, 2004.

Accepted June 22, 2004.

COMMENTARY

Pavlovian Conditioning Requires Ruling Out Nonassociative Factors to Claim Conditioning Occurred

Steve Reilly

University of Illinois at Chicago, U.S.A.

Todd R. Schachtman

University of Missouri, U.S.A.

In a thought provoking article based on their studies concerning the behavioral capacities of spinally-transected rats, Grau and Joynes (2005) claim that protection-from-habituation and pairing-specific enhanced sensitization are learned phenomena that occur during pairings between a conditioned stimulus and an unconditioned stimulus. Our commentary questions whether such effects: (1) should be put in the same category as associative learning; (2) necessitate or warrant a new neurofunctionalism; (3) suggest that the field should have less emphasis on the methods of Pavlov and Skinner and more focus on function and neuroscience; (4) suggest that our textbooks be revised.

Grau and Joynes (2005) propose a new approach to the study of learning, labeled "neural-functionalism," stemming from their critique of perceived problems involving the practice (i.e., methodological issues) and image management (for our students and for colleagues in other disciplines) of the field of animal learning and conditioning. We discuss these critiques and this new approach by focusing on these two problems in turn. Although some important, if primarily familiar, issues are raised, we are unconvinced by Grau and Joynes's arguments.

Methodological Issues

The views articulated by Grau and Joynes developed out of their research into the learning capacities of a reduced animal preparation, spinally-transected rats. This work, involving the analysis of behavior produced by pairing a conditioned stimulus (CS) with an unconditioned stimulus (US), appears to be amenable to explanation in terms of one of three factors: Pavlovian conditioning as it is conventionally viewed, protection from habituation (PH), or pairing-specific enhanced sensitization (PSES). Most researchers would ask whether the change in behavior observed is due to associative learning or an artifact (PH or PSES). Grau and Joynes concluded that the behavioral effects obtained in their spinal rats were not due to the former factor and, rather than accepting an artifactual interpretation, the authors took a much bolder step by declaring that PH and PSES should be considered

Preparation of this article was supported by National Institutes of Health Grants DC-004341 and DC-006456. Correspondence concerning this article and requests for reprints should be addressed to Steve Reilly, Department of Psychology, University of Illinois at Chicago, 1007 West Harrison Street, Chicago, IL 60607, U.S.A. (sreilly@uic.edu), or Todd Schachtman, Department of Psychological Sciences, University of Missouri, Columbia, MO 65211, U.S.A. (SchachtmanT@missouri.edu).

subcategories of Pavlovian conditioning. If we assume for the moment that sensitization and habituation are nonassociative effects, then Grau and Joynes are arguing that these effects are subcategories of Pavlovian conditioning because their occurrence requires the pairing of a CS with a US. This is an intriguing idea. One of the present authors contemplated this same issue when he was exploring *US reinstatement* or *reminder* effects (Schachtman et al., 1983; see Miller & Springer, 1973). A conditioning trial (say, a light-shock pairing) would be given to the subjects and poor expression of learning would result for one reason or another; then, the experimenter would administer a *reminder treatment* (e.g., a shock alone exposure) after conditioning and prior to testing for the conditioned response (CR), and this would cause learning to be expressed (i.e., a CR). If control groups (rats receiving the CS alone, US alone, or the CS and US unpaired, instead of a CS-US pairing) did not show the CR when the reminder was administered to them, the behavioral effect was assumed to depend on associative processes and was, therefore, not artifactual (i.e., not due to nonassociative factors). But one might ask, *What if, in order for the reminder treatment to produce a nonassociative effect, the subject must receive the CS and US together?* Such an effect seems unlikely at first blush, but could happen since, after all, the presentation of a light produces a certain neural response as does a shock and maybe the experience of both close together in time causes the animal to be subject to the nonassociative effects of the reminder treatment. Grau and Joynes have not only produced such effects in their reduced preparation but even have plausible nonassociative mechanisms for them (PS and PSES, assuming they are nonassociative effects). But, should these behavioral effects be explained by broadening the domain of Pavlovian learning or, alternatively, should we simply acknowledge that we do not currently have ample control conditions to rule out all possible nonassociative (i.e., artifactual) effects? Grau and Joynes argue for the former, we prefer the latter. Furthermore, their choice provides the departure point for their new conceptualization of conditioning called neural functionalism. We do not see any compelling reasons to follow in that direction.

Perception of the Field and Image Management

As we all know, when an instructor teaches an undergraduate introductory psychology class, he/she is forced to greatly simplify the theories and findings. The students are not ready to accommodate more than that. In science, there are many instances where we must draw crisp lines on a canvas when, in reality, the image is fuzzy and complicated. For instance, in the physical sciences we have courses in (and departments of) biology, physics, and chemistry when we know that these areas are conceptually overlapping and inseparable. Even in our own psychology departments with separate training areas, we have a difficult time trying to figure out where to put areas or researchers studying emotion, motivation, judgment, or even personality since such things bridge many areas within psychology. Grau and Joynes are correct that many texts/courses in conditioning are separated into Pavlovian and operant conditioning. Students may not be ready to have the text or course structured in terms of mechanism (and, by mechanism, we refer to psychological mechanism as discussed below, not biological mechanism as apparently suggested by Grau and Joynes since that would be a text or course in neurosci-

ence). Most if not all areas in psychology organize their texts and courses around procedures and behaviors rather than the mechanisms underlying them. And it is true that the *phenomena* of conditioning (e.g., a CR following Pavlovian conditioning) are operationally defined giving them the semblance of merely being a methodology. However, for most researchers investigating Pavlovian conditioning the CR reflects the formation of an association between the CS and US and the subject experiences, upon presentation of the CS, an expectancy of the US. The presentation of the CS can evoke a representation of the US in active memory. Similar processes have been explored for responses and outcomes in instrumental conditioning. As Rescorla rightfully put it in the title of his 1988 paper, "Pavlovian conditioning: It's not what you think it is." We claim that it is more than a methodology and subsumes associative learning. PH and PSES are not Pavlovian conditioning any more than pseudoconditioning is conditioning.

The functional mechanisms level of Figure 4, in the target article, mix different levels of discourse or levels of analysis. They describe two effects that are operationally defined (PH and PSES) just as Pavlovian conditioning is operationally defined. All three effects are produced by a similar environmental condition. All three phenomena should be placed on a similar level: Behavioral responses to the methodological/environmental level in Figure 4 of the paper by Grau and Joynes (i.e., pairing of a CS with a US) and then there should exist a level below that in which the underlying (psychological) mechanism of each phenomenon be determined. Associative learning is likely one mechanism of Pavlovian conditioning and this associative mechanism may very well be further explicated psychologically with respect to the expectancies, activation of event representations, retrieval, and other psychological processes potentially involved with such associations). Grau and Joynes argue for "a more mechanistically based psychology of learning" (p. 4) but we believe that the field of conditioning is already very mechanistically based. The underlying mechanisms of PH and PSES remain to be determined. In fact, conditioning theory requires that certain mechanisms that have been explored in one procedure (e.g., conditioned emotional response, conditioned eyeblink, conditioned taste aversion) be explored in other procedures (and Grau and Joynes assume, for instance, that associative basis of long-term habituation applies to spinal cases of habituation and, yet, this remains to be determined).

We struggled with some of the arguments presented in the Grau and Joynes article. In large part this was due to difficulty understanding some of the terminology and a lack of detail about some of their points. Our appreciation of this article would, for example, have benefited from defining what is meant by functional mechanism (since the authors assert that the field currently is not mechanistically based, when they seem to mean that it is not centered around neuroscience). They claim they "question restrictive views of what constitutes learning" (p. 3) and so it would have been helpful if they had defined learning itself so one can appreciate how this conceptualization restricts learning. They claim that "the architecture has changed" (p. 3) and yet do not specify what changes they are referring to. Secondly, since nonreinforced conditioning with a CS can produce conditioned inhibition, it is not clear how much the effects in Figure 2 are due to conditioned excitation to CS+ and inhibition to CS-. We also do not know if the effects obtained by Grau and Joynes are stimulus specific. We also do not understand why statements would be made that claim "Tradition...would have us ignore identified biological

mechanisms" (p. 13). Does tradition include learning theorists ignoring neuroscience? We think not. It is simply another level of discourse or analysis (in fact, a different course).

Grau and Joynes acknowledge that many of the features of neural functionalism are already fully appreciated by the researchers in the field. This is true. Nothing in the current field denies that organisms can solve an environmental puzzle in different ways (and potentially using more than one brain mechanism for each way and using different brain mechanisms for different ways). No new field of neural-functionalism is needed. The authors claim that "researchers working within the multiple memory paradigm have adopted a neurofunctionalist approach, one that focuses on the brain mechanisms that underlie information storage" (p. 5). Once gain, to us, this sounds like part of the domain of the field of neuroscience and no new label is needed.

The authors are correct that it is a challenge to make our conditioning courses palatable to students and to retain respect from our colleagues from other disciplines. Students can and do get excited when they learn about the different potential underlying psychological mechanisms of conditioning (expectancies, activation of event representations, etc.) and so we should expose these intriguing processes in our lectures. There are many things that will assist our image management with colleagues. For instance, the paper by Grau and Joynes refers to our field as the "field of learning" (p. 18) and we would guess that our human learning and cognition colleagues scoff when they see such expressions. These colleagues must think we still linger in the middle of the twentieth century when our field *was*, for the most part, the field of learning. We are the field of conditioning or the field of animal learning or, if you like, the field that explores a subset of learning processes. In the final analysis, progress in animal learning and conditioning is evolutionary in the sense that the merits of new ideas, theories, or viewpoints are determined, not by individual appeals to correctness, but by whether these new ideas have greater utility than those they seek to supplant. As it should be, time will be the judge of the merits of such a case.

References

- Grau, J. W., & Joynes, R. L. (2005). A neural-functional approach to learning. *International Journal of Comparative Psychology*, **18**, 1-22.
- Miller, R. R., & Springer, A. D. (1973). Amnesia, consolidation, and retrieval. *Psychological Review*, **80**, 69-79.
- Rescorla, R. A. (1988). Pavlovian conditioning it's not what you think it is. *American Psychologist*, **43**, 151-160.
- Schachtman, T. R., Gee, J-L, Kasprow, W. J., & Miller, R. R. (1983). Reminder-induced recovery from blocking as a function of the number of compound trials. *Learning and Motivation*, **14**, 154-164.

Received, April 16, 2004.

Revision received June 22, 2004.

Accepted June 23, 2004.

COMMENTARY

A Continuum of Learning and Memory Research

Greta Sokoloff and Joseph E. Steinmetz
Indiana University, U.S.A.

History has revealed time and time again that science is moved forward by revolutions that pit one point of view, theory, or methodology, against an opposing view. During calmer times, however, we as researchers are left to our own devices and settle into our work with little thought to the world around us. The field of learning and memory has been privy to many such revolutions in the past but has yet to form a cohesive, modern message. Grau and Joynes suggest that our strong ties to the past are to blame for a lack of progression in the field. We agree and add that the focus of the field on two extreme ends of a continuum has also held us back; suggesting that research that goes on in the middle of the continuum may be the key to leading the field out of its rut.

Grau and Joynes (2005) posit that there are two core assumptions that maintain and nourish research in the field of learning and memory. The first assumption, which is not problematic to the authors (nor to us), is that learning is essential. The second assumption, however, is where Grau and Joynes assert that the field has been led down a “problematic path” (p. 2)—that the interpretation of the assumption of generality in learning has led the field astray. Specifically, the study of phenomena with traditional methodologies that manipulate relationships between stimuli and responses, such as Pavlovian, operant, or instrumental conditioning paradigms, has resulted in the methodologies and procedures themselves taking precedence over the phenomena being studied.

Grau and Joynes have pointed out an important problem in the field of learning and memory. The way in which many in the field of learning cling to traditions established long ago is troublesome not only for the research we perform and the data we obtain but also for the dissemination of the field, and our findings, to others. This problem is especially evident in the university classroom. Few would argue the fact that when we construct a curriculum for a course on learning and memory, we perpetuate the very problem in the field that Grau and Joynes identify. Not yet has a comprehensive text been written that has even come close to tearing down “the traditional house” (p. 2) thereby reinforcing a segregation of theory with no attempt at synthesis.

Although it is quite evident that the way in which we teach learning to students is in desperate need of “modernization” it is also important for a synthesis across research programs. Many of us, however, do strive to achieve this goal. In fact, the incursion of the problem of generality in research, as emphasized by Grau and Joynes, arises from a blurring of two extreme ends of a continuum. To this extent, the authors’ criticism of using particular methodologies—setting up shop in a single room (p. 2)—although important, may be a bit overstated.

Researchers who study behavioral and biological underpinnings of learning and memory, like researchers involved in many scientific disciplines, have research interests that can be placed on a continuum. At one end of the continuum there are researchers interested in describing and uncovering the laws of behavior based primarily on a history of traditional learning theory and animal behavior. Not surprisingly, these researchers use traditional methodologies to achieve that goal. At the other end of the continuum, researchers may have little concern for learning in and of itself, instead being interested in understanding the laws of neurobiology that are associated with the behavior. These researchers also use traditional methodologies because the resultant changes in behavior that are known from previous work provide them with a means to assess neural and molecular substrates of learning. The point on the continuum that one decides to “set up shop”, therefore, depends on the specific interests, goals and training of the individual researcher.

The dichotomy that establishes the endpoints of such a continuum is not unique to the study of learning and memory. It has spurred on many important scientific revolutions; evolution, developmental biology, and psychology, and of course in learning as well—God vs. natural selection, nature vs. nurture, behaviorism vs. cognitive psychology. While these extreme positions frequently act to garner our attention and make us think more deeply about our work, the majority of us do not see the world in quite so black and white terms. It may be argued that much of the work that moves us forward and fills in the gaps arises from the gray areas, the middle of the continuum. A good example of work at the middle of the continuum is the use of fear conditioning procedures to study the role of structures like the amygdala and hippocampus in emotional learning (e.g., Fanselow, 2001; Lamprecht and LeDoux, 2004). Many laboratories have adopted variations of fear conditioning procedures to study how brain systems interact and encode this form of learning in a number of different species (including humans). Much of this work has advanced our understanding of acquired fear behaviors while at the same time advancing our understanding of the neurobiological underpinnings of the behavioral change.

It is true, as Grau and Joynes point out, that the majority of researchers utilize a rather small set of traditional methodologies (i.e., Pavlovian, operant, or instrumental conditioning). Obviously at one end of the continuum the interest is in the methodology itself, while at the other end the methodology serves as a tool for looking at other events (i.e., neural activity, protein synthesis, receptor populations). The use of traditional paradigms/methodologies as research tools, however, is not necessarily the result of “seeking a safe course” as Grau and Joynes suggest (p. 3). Many of us utilize traditional learning paradigms because they allow us to control a larger number of experimental variables so that underlying neural function can be assessed. Our own work in studying how structures like the cerebellum and hippocampus are involved in classical eyeblink conditioning provides an example of this use of a traditional paradigm (see Steinmetz, 2000, for review). Due to the elegant and comprehensive behavioral work of Isadore Gormezano, Allan Wagner, James Kehoe, Berny Schreurs, and others, we know what to expect behaviorally when manipulations such as those involving interstimulus intervals, intertrial intervals, and stimulus intensity are performed. Being able to predict what happens behaviorally when variations of this basic, traditional learning procedure

are performed has greatly aided our ability to elucidate neural processes associated with the learning. In other words, it is not so much that we are suspicious of other paradigms, rather, we use what works optimally and predictably for elucidating mechanisms underlying the phenomena we are studying.

The importance of mechanism is well-established by Grau and Joynes, and the authors suggest that mechanism should be the focus of the field as opposed to methodology. We too appreciate the importance of mechanism(s) and it is what drives our research program. Mechanism, however, is not immune from the problems suffered by methodology. Mechanism may vary due to methodology at a behavioral level (e.g., operant conditioning vs. instrumental conditioning) or at a neural level (e.g., what and how brain structures are engaged during learning). An example from the classical eyeblink conditioning literature may serve to illustrate this point. From a behavioral standpoint, evidence for the acquisition of the classically conditioned eyeblink response is the appearance of anticipatory conditioned responses (CRs), which we can call the behavioral mechanism. A number of studies have shown that a small lesions placed in the interpositus nucleus of the cerebellum can abolish this type of learning (e.g., Steinmetz, Logue, & Steinmetz, 1992). From these (and other) data it has been established that the cerebellum is critically involved in the acquisition and performance of the classically conditioned eyeblink response (see Steinmetz, 2000, for review). This is an example of a neural mechanism. Several years ago, Kelly, Zou, and Bloedel (1990) reported the appearance of eyeblink CRs in rabbits that had cerebellum and cerebral cortex removed (i.e., essentially a brainstem preparation). Do these data show that the cerebellum is not critical for classical eyeblink conditioning? Probably not. These data more likely demonstrate the uncovering of plasticity processes in the brain stem after cerebral cortical input is eliminated, a second neural mechanism that is capable of supporting the appearance of the behavioral mechanism (the conditioned eyeblink). Importantly, this second neural mechanism cannot independently support conditioned responding in the intact preparation where the cerebral cortical input is present. This example illustrates the point that a given behavioral mechanism may be accounted for by a number of different neural mechanisms that play different roles in producing the behavior depending on the components of the nervous system that are available to the organism. Indeed, many neural mechanisms may in fact be overlooked or under-appreciated because of traditional mechanisms that hold sway in the field (i.e., cortical inhibition, LTP, LTD, PKC, etc.). Again we see that the problem of generality of neural mechanisms to be just as potentially troublesome here as it is when traditional learning methodologies or phenomena are compared (e.g. Is latent inhibition mechanistically, both behaviorally and neurally, the same during appetitive and aversive forms of instrumental learning?).

Grau and Joynes should be applauded for their efforts to regroup and revitalize the field of learning and memory. Their caution of leaning too heavily on old theories, both in our research and in our classrooms, is well warranted. Their position of neural functionalism, focusing on motivational and operational behaviors and the underlying neural mechanisms, is sound. In fact, many of us would purport to already be converts or followers. In the end, our goal as researchers in the field (regardless of our location on the continuum) should encompass an understanding of learning that can explain phenomena across methodologies, in different species,

and across development. When we get to this point maybe we will have learned something.

References

Fanselow, M. S. (2001). Toward a neurobiology of functional behavior systems: Contrasting Pavlovian emotional and motor learning. In J. E. Steinmetz, M. A. Gluck, & P. R. Solomon (Eds.), *Model systems and the neurobiology of associative learning* (pp. 379-393). Hillsdale, NJ: Erlbaum.

Grau, J. W., & Joynes, R. L. (2005). A neural-functionalist approach to learning. *International Journal of Comparative Psychology*, **18**, 1-22.

Kelly, T. M., Zou, C.-C., & Bloedel, J. R. (1990). Classical conditioning of the eyeblink reflex in the decerebrate-cerebellate rabbit. *Behavioural Brain Research*, **38**, 7-18.

Lamprecht, R., & LeDoux, J. (2004). Structural plasticity and memory. *Nature Reviews Neuroscience*, **5**, 45-54.

Steinmetz, J. E. (2000). Brain substrates of classical eyeblink conditioning: A highly localized but also distributed system. *Behavioural Brain Research*, **110**, 13-24.

Steinmetz, J. E., Logue, S. F., & Steinmetz, S. S. (1992). Rabbit classically conditioned eyelid responses do not reappear after interpositus nucleus lesion and extensive post-lesion training. *Behavioural Brain Research*, **51**, 103-114.

Received April, 19, 2004.

Revision received June 22, 2004.

Accepted June 22, 2004.

COMMENTARY

Beyond Method

John Staddon

University of York, United Kingdom

Procedures are not the same as processes. But purposive modules are probably no more than a way-station to understanding learning processes and may not be as simply represented in brain neurophysiology as many seem to assume. And finally: it is impossible (and therefore unwise) to specify in advance what the ultimate theory of learning *must* explain.

It is possible to object to many specific points in this thoughtful article by Grau and Joynes (2005), but it is impossible to disagree with its overall theme. Learning and conditioning procedures *are* surely a means to an end, where the end is to understand the myriad ways in which behavior, human and animal, depends on the environment. Conditioning methods, the standard paradigms for habituation, classical and instrumental conditioning, are not themselves processes or mechanisms, nor are the phenomena they demonstrate universal building blocks for learned behavior.

When I first learned about classical conditioning, as an undergraduate, I remember being told that Western scientists had misunderstood Pavlov. "Conditioned response" for him just meant a response "conditional upon" a history of pairing between CS and US. It was simply a descriptive term, nothing more. Pavlov assumed no "conditioning process" of the sort to which Grau and Joynes now object. It is, I suppose, rather incredible that these caveats have been ignored by so many able people for so long that now the very same lesson must apparently be learned again.

So, what is to be done? The authors propose an *ism* they call neurofunctionalism: "Functionalism because the approach focuses on the identification and comparison of operational modules designed to accomplish a particular goal, be it the abstraction of environmental relations, recognition of a food source, or spatial navigation. Neural because an integral component of the approach involves the specification of the underlying neural mechanisms" (p. 5). The functional part of this idea resembles Timberlake's behavior systems approach, and the modular view now common in human evolutionary and cognitive psychology. The brain, it is argued, is not in any sense a general purpose learning system (so much for Hull, Skinner, Bitterman, and all!). Rather it is organized as a set of more or less isolated units that have specific niche-related functions such as spatial navigation or food recognition. The neural part of their idea is that there are may be more than one neural mechanism underlying a given function and that mechanisms that underlie "learning" need not exclude peripheral processes.

Well almost every science begins with classification—of species, substances, elements, whatever. Learning psychology began with a classification of procedures but, as the authors point out, that has not worked well. An appropriate analogy might be to the early chemists. Their first try was air, earth, fire, and water. A good start, but not fundamental. Physical science had to wait until Mendeleev to get a classification worked—in the sense that it led onward to a deeper understand of physics and chemistry. So, will classification based on “modules” work better than the tripartite division the authors criticize? I’m not sure, because this also is not a new idea. It is hard to see where “modules” differ much from faculty psychology, or William McDougall’s idea that we all have a number of what he called “instincts.”

If functional modules can be identified with a particular neural substrate, of course, progress will have been made. And in a few cases, this does seem to be the case. But in others, it seems pretty clear that the relevant neural structure may be interwoven with other structures serving very different functions. But these possibilities are familiar and it is probably difficult to find anyone who would disagree with this approach.

What of “mechanism,” the other half of neurofunctionalism? Mechanism is a protean term. The OED gives as a first definition: “The structure or operation of a machine or other complex system; a theory or approach relating to this.” In other words, *mechanism* can refer to structure—and I believe that Grau and Joynes intend this meaning—or to the theory of operation of the system. Pure structure is nothing but neuroanatomy. To put “mechanism” into structure implies a bottom-up approach: explaining the overall operation of the system by the properties of, and interactions among, its neural elements. I have argued elsewhere (Staddon, 2001a, Chapter 1) that explanation at this level, of a complex, evolved structure like the brain, is likely to be very difficult—much more difficult than is usually supposed. The other meaning of “mechanism” refers to a theory of operation. This, I suggest, is where both psychologists and neuroscientists should look first if they wish to solve the brain-mind problem.

Now to specifics. The authors begin their otherwise splendid article with what has become an almost obligatory swipe at behaviorism—“This trichotomy, a vestige of our behaviorist past ...” (p. 1, Abstract). But behaviorism is a philosophy rather than a specific set of scientific assertions (Staddon, 2001b). It does not rest on these distinctions. Moreover, as Roediger (2004; not noted as a behaviorist himself) recently pointed out: we are (to a degree) all behaviorists now!

The authors mention *operationalism* as a component of the framework that led to the prevailing view of associative learning. They are right to be skeptical. Premature emphasis on operational definitions has always put the cart before the horse. A mature science allows you to define your terms operationally, but the converse is not true. Coming up with an operational definition will not by itself advance science. Theoretical understanding comes first; operational definition afterwards.

In their discussion of habituation, the authors write: “learning that is thought to rely on an associative mechanism generally exhibits the opposite relation; given an equal number of pairings, spaced presentation produces a stronger CR than massed presentation...” (p. 12) This may not be true. The standard Groves

and Thompson discussion of habituation is purely qualitative, but statements like this depend on quantitative properties. The habituation property we have termed *rate sensitivity* (Staddon, 1993; Staddon & Higa, 1996) suggests that this conclusion will depend on quantitative details: When habituation is tested after training, after a long time, the long-ITI group may show more habituation than the short-ITI group. But a short time after training, these relations may be reversed. Perhaps this just makes the authors' point: You have to understand the mechanism, the theory of operation of the system, to make sense of observations like this.

The authors attribute to learning theorists the core assumption "that learning is essential" (p. 1). But this is plainly false. Many species, particularly protists, get by with nothing beyond simple habituation. Learning is of course necessary for more complex niches, but I know of no one who thinks it essential to adaptive behavior in general.

Figure 1 is problematic, because it treats the biological and functional mechanisms as separate. This cannot be true. The functional mechanism, rules of operation, of the system is implemented by the biology, it is not something separate, or separable, from it.

As someone reared in the Skinnerian tradition, I nevertheless do not see the distinction between "operant" and "instrumental" behavior that the authors discern. I think that, as historically defined, there is no defensible distinction between operant and instrumental behavior. Pecking in pigeons is "operant" by Skinner's, or anyone's definition. Yet it is clearly "constrained" in the authors' sense (Staddon & Simmelhag, 1971).

Finally, the authors speak of what "a complete theory of learning must speak to..." (p. 15). This is such a common philosophical error that it deserves a name, perhaps the *imperative fallacy*—the idea that it is possible to specify in advance of its discovery the domain of a theory. It is as if some pre-Copernican scholar argued that a complete theory of the planets *must* specify their number, size and color, as well as their orbits and duration of year. Well, yes, in a sense. But we know now that this list is apples and oranges. Newton's laws apply to only some of these things; the others are either inexplicable or attributable to other processes entirely.

But enough carping. I thank the authors for a productive attack on a difficult and important problem.

References

- Grau, J. W., & Joynes, R. L. (2005). A neural-functional approach to learning. *International Journal of Comparative Psychology*.
- Roediger, R. (2004). What happened to behaviorism? *APS Observer*, **17**, March, [pp.?).
- Staddon, J. E. R. (2001a). *Adaptive dynamics: The theoretical analysis of behavior*. Cambridge, MA: MIT/Bradford.
- Staddon, J. E. R. (2001b). *The new behaviorism: Mind, mechanism and society*. Philadelphia, PA: Psychology Press.
- Staddon, J. E. R. (1993). On rate-sensitive habituation. *Adaptive Behavior*, **1**, 421-436.
- Staddon, J. E. R., & Higa, J. J. (1996). Multiple time scales in simple habituation. *Psychological Review*, **103**, 720-733.
- Staddon, J. E. R., & Simmelhag, V. (1971). The "superstition" experiment: A reexamination of its implications for the principles of adaptive behavior. *Psychological Review*, **78**, 3-43.

Received June 10, 2004.
Revision received June 22, 2004.
Accepted June 22, 2004.

Neurofunctionalism Revisited: Learning is More Than You Think It Is

James W. Grau
Texas A&M University, U.S.A.

Robin L. Joynes
Kent State University, U.S.A.

Studies of learning in simple systems (invertebrates and spinal cord) have revealed that organisms can encode stimulus-stimulus (Pavlovian) and response-outcome (instrumental) relations in multiple ways. It is suggested that nonassociative mechanisms contribute to learning and that there is value in adopting an approach that details the neural-functional mechanisms involved. Reactions to this approach are discussed. The link between the methods of Pavlov and associative ("true") learning is deeply ingrained and, some believe, should be maintained. We suggest that there is value in dissociating the concepts and seek to clarify the implications of a neurofunctionalist approach to learning. It is argued that a neural-functional approach provides a better framework for integrating behavioral and neurobiological observations.

For close to 15 years, we and our colleagues have examined some unusual forms of learning that do not fall within the traditional categories of learning theory. Building on a foundation laid by Thompson, Steinmetz, Patterson, and others (for reviews see Grau & Joynes, 2001; Patterson, 2001), we developed model paradigms that have now been shown to have considerable clinical significance (e.g., Grau et al., 2004). They have also proven amendable to uncovering the biological substrates of learning (e.g., Joynes, Janjua, & Grau, 2004; Liu, Crown, Miranda, & Grau, 2005). Indeed, with new discoveries within the field of spinal cord nociceptive plasticity, spinal systems now rival the hippocampus, amygdala, and cerebellum as one of the most well-characterized neurobiological systems (Ji, Kohno, Moore, & Woolf, 2003). But for 15 years, reviewers have questioned whether we were examining *true* learning, providing demerits for falling outside traditional categories. Oddly, no one questioned whether the examples of plasticity were clinically significant or had widespread implications, only whether the findings were relevant to learning. Frustrated by this continued assault, we have turned the question around and asked why learning should be so narrowly defined? Why should our examples of instrumental or Pavlovian conditioning be viewed in an inferior light just because we lack a Lockian association? Why is associative learning on the pedestal and blessed as the only true form of learning, and why have we inbred the concept with our methods of studying learning to such an ex-

We would like to thank the Editor for giving us the opportunity to question some basic tenets in the field of learning and for recognizing that such an enterprise requires commentary. We also want to thank those who provided the comments. We have known some of the commentators for many years. They agreed and disagreed with us in good spirit, and we hope that they accept our reply in a similar fashion. Finally, thanks are due to Kristy Acosta, Michelle Hook, Russell Huie, Marissa Maultsby, Christine Petrich, Denise Puga, and Stephanie Washburn for their helpful comments. Preparation of this paper was supported, in part, by Grants MH 60156 and NS 41548. Correspondence concerning this article should be addressed to James W. Grau, Department of Psychology, Texas A&M University, College Station, TX 77843-4235, U.S.A. (j-grau@tamu.edu).

tent that many find it hard to dissociate the concepts? In the process of considering these questions, we have knocked associative learning off its pedestal and questioned the way that we organize and present our data. Some applauded (Blaisdell, 2005; Sokoloff & Steinmetz, 2005; Staddon, 2005), acknowledging the timely nature of our critique. Others (Machado, 2005; Reilly & Schachtman, 2005) felt that the traditional story still reads well and sought to maintain the status quo.

Three of the five commentators agreed on two basic points: (1) the way we present the field to students and our colleagues is in need of revision; and (2) the field has not done a good job of integrating new discoveries on the neurobiology of learning. They generally shared our desire to expand the domain of learning, but struggled with how this should be accomplished. Their reactions ranged from an endorsement of our perspective to a more traditional view that maintains a special pedestal for true conditioning. The other two reviewers had no such struggle—to them, there is just one type of conditioning and it is associative in nature. Conditioning in the spinal cord and *Aplysia* fails on this criterion and is best swept out of the house of learning as a nonassociative artifact. We feel that this is the wrong tact. In the sections that follow, we attempt to clarify why we came to different conclusions. We begin by discussing a number of issues that appeared to stem from common concerns, seeking to clarify our position. We then deal with specific issues raised within each of the commentaries.

The Semantic Connection: Established Associations Make it Difficult to Dissociate Association

Our aim in the target article was to challenge some long-held views. But challenging a paradigm requires that we convince the reader to temporarily suspend a well-established way of viewing the world and consider an alternative perspective. With a framework as well entrenched as the doctrines of learning theory, this is not always easy to do. The problem is that some concepts have been so strongly associated that it is difficult for us to see that a single linguistic token refers to two distinct entities. We believe that this issue arose with regard to the distinction between Pavlovian conditioning and associative learning. As Staddon reminded us, Pavlov saw a conditioned response (CR) as a response that was “conditional upon” the “history of pairing between the CS and the US. It was simply a descriptive term, nothing more” (p. 38). We use the term Pavlovian conditioning in just this sense, to refer to a kind of learning that depends on the relationship between two stimulus events. For many, though, the term Pavlovian conditioning has additional meaning, for it also implies a form of associative learning. The semantic confusion is natural. Pavlovian conditioning involves (by definition) two stimulus events (S1 and S2) that have a physical relation—the distal/proximal cues are physically associated. Add to this the long history of associative learning within philosophy and psychology, and the elegance and power of the concept, and Pavlovian conditioning becomes inexorably linked to the mechanism of associative learning.

If the only way a S1-S2 pairing could have a lasting impact was through the development of a new association, then it would make sense for Pavlovian conditioning and associative learning to be forever married. Given this, the first question we had to address is whether there are other viable mechanisms. Two

were identified: protection from habituation and pairing-specific enhanced sensitization. We also outlined several operations that could be used to distinguish protection from habituation from associative learning and provided evidence that our example of spinal conditioning relied on the former mechanism. Here was a case where learning depended on pairing two stimulus events, and the learning exhibited a variety of Pavlovian phenomena (e.g., latent inhibition, extinction, and overshadowing). Our conclusion was that Pavlovian conditioning is not necessarily associative in nature.

We then recognized that unlinking the concepts of Pavlovian conditioning and associative learning had a variety of implications that went well beyond our own data. If the concepts are disconnected, the methods of Pavlov lose some of their import. They still have value, but attention is shifted to detailing the underlying mechanisms. Further, what holds for Pavlovian conditioning would seemingly apply to other standard methods. Indeed, one could argue that such a view was already accepted in some circles. For example, researchers studying habituation have long recognized that an array of mechanisms can bring about a decrement in response magnitude and they do not appear to hold any one (not even the associative account) in greater esteem. Moreover, as noted by Staddon (2005), detailing the operational principles that constitute a functional mechanism can provide a better framework for linking behavior to neurobiological systems. Together, these considerations led us to conclude that the future of learning lies with a decreased emphasis on the methods of Pavlov and Skinner and an increased emphasis on the underlying functional mechanisms and neurobiology.

Associative Learning and the Ghost of a Straw Man

Having unpacked our rationale for a broader approach to learning, we face the issue that dominated much of the commentaries. The issue concerns the extent to which the soul of learning is tied to the concept of association. Some suggested that few still hold this view—that it is a straw man (Blaisdell, 2005). If so, we are lucky indeed because we need look no further than the commentaries for evidence that the ghost of this straw man still holds sway.

There are actually two versions of the argument that learning is necessarily associative. The most general holds that nonassociative effects do not count as learning. Staddon (2005) maintains this position when he questions whether learning is essential. Our point was that researchers who have chosen to study learning view the process as providing a key adaptive capacity. Against this, Staddon notes that some simple organisms (protists) negotiate their environment with nothing more than the capacity for habituation. The implication was that learning is unnecessary because nonassociative habituation does not count as learning. We come to a different conclusion because learning for us is not limited to associative processes—nonassociative habituation counts as learning. (For a discussion of the criteria for learning see Grau and Joynes, 2001.)

The other version of the argument is equivalent to the view we asked the reader to suspend—that Pavlovian (or instrumental) conditioning is necessarily associative in nature. We seek to counter this belief, but again, one could question whether we are battling a straw man. Blaisdell argues this (last paragraph), but a few sentences later he suggests that some invertebrate preparations are superior to

Kandel's because they exhibit "true Pavlovian conditioning" (p. 26; for a discussion of how the results in vertebrate and invertebrate models compare, see Pittenger & Kandel, 2003). It would seem that the ghost of the straw man has once again wheeled his influence. Reilly and Schachtman (2005) adopt the view in full when they ask whether we should broaden "the domain of Pavlovian learning, or alternatively, should we simply acknowledge that we do not currently have ample control conditions to rule out all possible nonassociative (i.e., artifactual) effects? Grau and Joynes argue for the former, we prefer the latter" (p. 35). It would seem that many still see Pavlovian conditioning as married to just one mechanism—associative learning.

Why do we seek to broaden the domain of Pavlovian learning? To address this question, let's consider where Reilly and Schachtman's position would lead us. Examples of learning that involve pairing-specific enhanced sensitization are relegated to artifactual status. Add to this the nagging problem that many CSs have a nasty habit of generating a CR-like response prior to conditioning, and you may find your favorite preparation placed along the nonassociative curbside. The cost of preserving the unity between Pavlovian conditioning and associative learning would be a very narrow field of study. Next, what are we to make of the equivalent use of the terms? If the words have the same meaning, why have two terms? To this, it might be argued that there are different kinds of associative learning. For example, in some cases, Pavlovian conditioning seems to reflect a S-S association while in others it appears S-R in nature (Rescorla, 1975). This move breaks the circularity and broadens the sphere of influence because now Pavlovian conditioning refers to two distinct kinds of learning, S-R versus S-S. But those instances of S-R conditioning, so cleverly used to break the circularity, have a formal similarity to pairing-specific enhanced sensitization. It would seem that the circularity was broken by recreating a hierarchical scheme similar to that illustrated in Figure 4 of the target article.

Citing Rescorla (1988a), Reilly and Schachtman suggest that Pavlovian conditioning is much more sophisticated than we suppose, involving the capacity for abstracting informational value and representing hierarchical relations. It is, of course, true that some examples of Pavlovian conditioning have an underappreciated level of complexity. Recognizing that researchers in other areas often viewed conditioning as a low-level mechanical process, Rescorla outlined a series of findings that suggested greater sophistication. However, we do not believe (as Reilly and Schachtman seem to suggest) that Rescorla intended to provide a singular view of what constitutes conditioning. If so, then any example of Pavlovian conditioning that was insensitive to complex conditional discriminations would be deemed inadequate. The field of study would be narrowed even further, perhaps to those forms of learning that are hippocampally dependent. We are certain Rescorla did not intend this. Indeed, Rescorla noted elsewhere how S-S relations can be encoded in multiple ways and cites protection from habituation as an example (Rescorla, 1988b). Recognizing the complexity of some instances of Pavlovian conditioning does not negate the importance of simple model systems that may lack the capacity to exhibit occasion setting or mediated acquisition/extinction effects (Holland, 1990).

Levels of Analysis

A number of commentaries raised issues concerning the relationship between different levels of analysis. Our intent was to show that shifting attention to the functional mechanisms that underlie learning has value. But we apparently made this push with such force that some perceived a more insidious intent—to effectively assassinate the superordinate category and call an end to detailed behavioral analysis. This was not our intent. We see behavioral studies as essential to detailing the efficient causes; to derive an accurate (hopefully, mathematical) description of the circumstances under which a phenomenon occurs and its ecological significance. Further, the delineation of new behavioral categories and their underlying relations will, of course, depend on detailed behavioral analyses. Our push was designed to encourage a shift in focus, from the usual tripartite (single stimulus learning, Pavlovian conditioning, and instrumental learning) to the underlying mechanisms, a shift we (and Staddon) believe is essential to uncovering the underlying neurobiological mechanisms. But such a shift in focus does not, in any way, eliminate the reality of the (superordinate) behavioral categories. Recognizing this, we listed Pavlovian conditioning as the superordinate category within Figure 4. In this scheme, the three functional mechanisms were depicted as subcategories of Pavlovian conditioning. Given that these relations were explicated in Figure 4, we do not understand why Machado would claim that we “failed to notice that” (p. 30) the functional mechanisms were subcategories of Pavlovian conditioning. It seems odder still that he would argue against our position by suggesting that the functional mechanisms “identify in greater detail how the CS and US arrangements in a particular case affected behavior; they do not show the inconsistency” (pp. 30-31) Machado perceives a rebuttal in language that summarizes some key features of Figure 4.

What implications does our approach have for the way we analyze instances of learning? To answer this question, it is helpful to think of the problem in terms of a two-stage process. The first issue concerns the superordinate category. Does the instance of learning represent an example of single-stimulus, Pavlovian, or instrumental learning? (It is assumed here that we have established that the behavioral effect qualifies as an instance of learning; see Grau & Joynes, 2001.) For Pavlovian conditioning, researchers have established a set of operations that can be used to demonstrate that the S-S relation matters. We will designate this set of operations as set “X” and assume that these conditions have been met. The next question concerns the nature of the underlying mechanism. Let us suppose that the example in question involves a case of protection from habituation. In Joynes and Grau (1996), a set of operations was derived (set “Y”) and used to show that this mechanism seems to underlie our example of spinal conditioning. Notice that the operations needed to classify the behavioral effect as an instance of Pavlovian conditioning (X) are not equivalent to those needed to classify the effect as a case of protection from habituation (X+Y). Given this, we are confused as to why Machado would question whether there is a “difference in the *logical* status of the two definition” (p. 31). The two sets of operations are not logically equivalent.

Does classifying a behavioral effect as a case of Pavlovian conditioning have mechanistic implications? Yes, of course it does. Within our framework it would suggest that one of three mechanisms is at work. Does Pavlovian conditioning have a biological reality? Again, of course it does. Neither here, nor in the target article, are we concerned with environmental relations in the absence of a processing organism. Nor are we concerned with stimulus events that the organism cannot sense. Pavlovian conditioning is of interest because it provides a sensible way of demonstrating that the *organism* is sensitive to a S1-S2 relation, independently of whether that sensitivity is due to associative or nonassociative mechanisms.

Views We Did Not Intend to Endorse

In a number of instances, those commenting on our target article read into it meaning that we did not intend. For example, as discussed above, we did not mean to disparage the importance of detailed behavioral analysis. Another example arose in Reilly and Schachtman's commentary when they suggested that *mechanism* for us meant *biological mechanism*. As illustrated in Figure 4 of the target article, we did not intend this narrow meaning. As indicated, mechanism was used in reference to both biological and functional systems. We believe that our use of *functional* mechanism is similar to what Reilly and Schachtman have in mind when they refer to *psychological* mechanism. We prefer functional because psychological has connotations that we would like to avoid and because the term function fixes our attention on the most pressing issue (for learning theorists)—what the mechanism is designed to do. As Staddon clarifies, mechanism here concerns a theory of operation, and as he suggests, this is often the most useful meaning of the term. Further, this use of the term can be applied at multiple levels of analysis and we agree that specifying how the components at each level operate is key to deriving their relation.

Schematically representing different levels of analysis can pose a challenge. In both Figures 1 and 4, we highlighted the difference between functional and biological descriptions by presenting each at a different level with the constructs connected by arrows. As Staddon reminds us, both levels of analysis refer to the same anatomical substrate.

Our presentation sidestepped some complexities that require clarification. In discussing behavior, function, or neurobiological systems, we can frame questions at either a local or global level. As an example of a local effect, consider a functional/biological mechanism designed to prime behavioral responding when a stimulus is reencountered (a form of sensitization). This priming effect might be linked to the secretion of a particular neural transmitter. At the level of the behavior system, the release of this transmitter could enhance food directed behavior. The global function might be described in terms of arousal and appetitive drive, and the biological system would involve a widely distributed neural circuit and a host of brain regions. We mention these possibilities because many of our examples concerned local mechanisms and simple behaviors, a focus that fits well with the basic categories of learning. It should be recognized, however, that we do not see our approach as limited to such local issues—learning can involve a local

modification within a particular component (module) of a system or a restructuring of the network that defines the system.

We also need to clarify some issues concerning the type of modularity assumed. Following Timberlake and Gallistel (Gallistel, 1980; Timberlake & Lucas, 1989), we envision learning as occurring in a type of lattice hierarchy. Suppose that we have a system that is sensitive to S-S relations and has embedded within it the capacity for protection from habituation, pairing specific enhanced sensitization, or associative learning, with the relative contribution of each varying as a function of training and other variables. How do we envision these mechanisms being distributed within the nervous system? One possibility is that each type of learning is mediated by a distinct neural component and that each mechanism handles its respective function across a range of learning phenomena. For example, the associative system might handle the linking of representations for both appetitive and aversive USs. Similarly, another neural mechanism might provide the capacity for protection from habituation and so forth. In some cases, nature may provide such simplicity, but this is likely the exception rather than the rule—that in many instances, the capacity for a given type of learning (e.g., protection from habituation) is multiply represented across the nervous system.

What about within a particular neural/behavior system? Should we expect to find each type of learning capacity residing within a distinct component of the structure? Again, nature may occasionally simplify our analysis in this fashion, but we also anticipate more complex scenarios. For example, a single anatomical structure may be capable of all three forms of learning—what may vary is the neurochemical system engaged. In this case, dissecting their contributions will require a different methodology. Rather than the traditional neural lesion approach (used in cases where distinct modules are thought to reside in distinct anatomical loci), local application of various agents may be needed to biochemically manipulate the learning processes. In terms of both operation and neurochemistry, each form of learning may exhibit a form of independence, but reside within the same set of neurons. A further complexity arises from an inherent quality of a lattice hierarchy, for the same component may subserve many masters. Such complexity undermines the plausibility of simpler views of modularity that aspire to link discrete functions to particular anatomical substrates. Rather, the function of a unit will likely depend upon the system to which it contributes. We mention these alternatives because Staddon (2005) appeared to believe that we endorsed a simpler view.

We have pushed for a focus on mechanism over methodology. Does this necessarily imply that everyone needs to change the way they attack a problem within the laboratory? No, not at all. First, as mentioned above, detailed analysis of behavior is still needed, in part because some basic questions within the field of learning have yet to be addressed, and in part because we have only just begun to flesh out the details of how the component modules are brought together to form an integrated behavior system. Do we see any problems with researchers pursuing particular model paradigms? Again, not at all. Current paradigms are built upon a rich behavioral history, and there is much to be gained from this history and the capacity to compare results across laboratories. Indeed, much of what we currently know about the principles and neurobiology of learning has been derived using well-established model systems. We only become concerned when those who have adopted a particular paradigm become dogmatic regarding its benefits and

prejudicial in their evaluation of new paradigms. But as Sokoloff & Steinmetz remind us, learning theorists are not the only ones capable of becoming dogmatic. Neuroscientists too can become single minded. For example, some might worry that the focus on NMDA-mediated plasticity has led researchers to ignore other potential mechanisms.

Specific Rebuttals

Beyond the general issues discussed above, each of the commentators raised a number of specific issues that require further attention. In the sections that follow, we respond to a selection of these issues. Due to space limits, we cannot address every comment, but instead, focus on those issues that are most central to our thesis.

Blaisdell: No Madness in Mechanism

Blaisdell (2005) agreed with many facets of our target article, and that the “disconnection between facts and framework probably contributes significantly to the tendency for students to perceive courses on learning as difficult or uninteresting” (p. 23) However, he expressed concerns regarding the perceived disparaging of behavioral approaches. We agree that behavioral analysis is required to identify and describe the operational principles that guide learning. We believe that Blaisdell would agree that the most informative behavioral data are those that yield new insights into how the system operates, and if so, the implicit focus remains on the functional mechanism.

Sokoloff and Steinmetz: An Inclusive View of Memory Research

Again, we found little to disagree with and feel that their comments helped to clarify a number of important issues. They also provided an interesting example of how multiple mechanisms can contribute to the encoding of a CS-US relation within an eyeblink paradigm and how learning within a more sophisticated system can sometimes usurp control over the process. Our only worry stems from the way such issues have been handled within the literature, where it sometimes appears that researchers hope to argue a mechanism out of existence. We expect that Sokoloff and Steinmetz (2005) would agree that a full description of the functional/neurobiological system must include all components.

Staddon: On Respondent and Operant Behavior

Staddon (2005) too helped to clarify a number of issues discussed above. He also described an interesting property of habituation (rate dependency) which, as he notes, helps to make our point. He did, though, differ on a few issues.

One difference concerns the distinction drawn between instrumental and operant conditioning. The components of our argument were first outlined in Grau, Barstow and Joynes (1998) where we recognized a potential to overstate our claims. In that paper, we examined whether spinal cord systems are sensitive to a response-outcome (R-O) relation, the distinguishing feature of instrumental learn-

ing. Building on earlier reports (Buerger & Chopin, 1976), we provided evidence that spinal cord neurons are sensitive to R-O relations and discounted a non-instrumental reactive model (a mechanical system that does not encode the R-O relation). The overall pattern suggested that neurons within the spinal cord could exhibit a form of instrumental learning. Given that many treat the terms operant and instrumental as synonyms, it was tempting to conclude (as others have done) that we also demonstrated a form of operant learning. Yet, we were nagged by a problem with this reasoning. Skinner saw behavior as falling into two categories, respondent or operant. He viewed respondent behavior in reflexive terms, as a type of elicited response. This raised a dilemma for us because we suspect that Skinner would argue that our example of spinally-mediated instrumental learning represents a case of respondent behavior (because the effective reinforcer—shock onset [Grau et al., 1998]—elicits our target response, leg flexion). One implication of this (to us) was that instrumental and operant learning do not refer to identical constructs. Another is that Skinner seemingly had additional criteria in mind when he drew the distinction between respondent and operant behavior—the key difference does not appear to depend on the R-O relation alone. Indeed, it is not clear that Skinner would necessarily deny that a R-O relation can influence a respondent. More formally, sensitivity to the R-O relation may be a necessary, but not a sufficient, condition for classifying a behavior as a Skinnerian operant. If so, the distinction between operant and respondent must depend on additional criteria. Of course, there is the well-known quality of “emitted,” but we were reluctant to build a definition on the inability to identify an effective cue. It seemed to us more profitable to ask how our examples of spinal learning differed from instances of behavior that Skinner would have likely seen as good examples of operant behavior. Two factors were identified, both of which concerned the degree to which the behavioral effect was biologically constrained. In an ideal operant situation, we could train a variety of behaviors (e.g., an increase or decrease in the response) using a variety of reinforcers (e.g., appetitive or aversive). Of course, we recognized that no learning situation is ever completely free of biological constraints. Nonetheless, it is not too difficult to find cases of human and animal behavior that seem far less constrained than spinal learning. By this analysis, demonstrating instrumental learning requires a set of operations that show that the system is sensitive to the R-O relation (set “A”). Operant learning requires evidence that the R-O relation matters plus additional criteria (e.g., that neither the behavioral change nor the reinforcer are constrained; Grau et al., 1998). If the additional criteria are defined by set “B”, a demonstration of operant behavior requires A+B.

Thus, we agree with Staddon that behavior in traditional operant paradigms is often biologically constrained to some extent. Yet, we expect that Staddon would agree that instrumental learning within the spinal cord is less flexible and depends on an elicited response. If so, it would seem our example of learning has a respondent quality. More generally, one could argue that respondent behavior is inherently more biologically constrained than operant learning. Indeed, it is tempting to posit that this continuum contributes more to the distinction between respondent and operant behavior than sensitivity to the R-O relation.

We believe that our analysis remains historically true and side steps a host of problems. If we had referred to our instrumental learning as an example of operant behavior, someone would have quickly pointed out its respondent character.

This would have then been followed by a list of criteria that seemingly distinguish traditional operants from our learning phenomena, with the conclusion being that we had failed to demonstrate operant learning. In the end, we agree that the learning differs and rather than wait for the critics charge, we chose to admit the differences. Staddon may reasonably wonder whether our attempt to ascribe additional meaning to the terms instrumental and operant has lasting merit, but we expect that he would agree that struggling with these issues is preferable to the difficulties that usually follow the casual application of behavioral terms.

Staddon concludes by questioning whether it is reasonable to suppose that we can foretell what a complete theory of learning must address. Of course we cannot, and we did not mean to be so presumptuous. Keeping with his analogy to the development of physical laws, we face a situation where many have focused on the movement of the sun and the way it is pulled across the sky. They see questions regarding the movement of other bodies as less important. We seek to broaden the class of phenomena deemed relevant. Does this mean that we may occasionally attend to an irrelevant property? Probably so, but we believe that the potential benefit of a more integrated theory is worth that risk.

Reilly and Schachtman: Learning to Include Nonassociative Factors

Reilly and Schachtman (2005) are happy with the status quo and see little reason to change the way in which they characterize and study the phenomena of learning. They chide our grandiosity, suggesting that our colleagues in human learning/cognition would scoff and accuse us of lingering in days gone by when the field of learning was the centerpiece of psychology. Yes, we have ambitious aims. The integrative field of study we envision encompasses much more than “animal learning or, if you like, the field that explores a subset of learning processes” (p. 37). By focusing on the functional mechanisms that underlie learning and memory, neurofunctionalism could provide a bridge between behavioral application (human and infrahuman) and neurobiological observations. The delineation of operational principles and linking hypotheses is not, in our mind, a subspecialty, but rather a central theme that provides an essential bridge. Of course, cognitive psychologists will recognize that we are adopting key features of the information-processing paradigm. The differences are that: (1) our models will be informed by neurobiological observations (as has become the case within cognitive neuroscience) and (2) we do not limit our attention to higher brain processes (as in traditional cognitive psychology).

Reilly and Schachtman discredit our analysis by suggesting that we mix different levels of discourse. We believe that the mixing may lie elsewhere. They suggest that Pavlovian conditioning should be at the same level as protection from habituation and pairing-specific enhanced sensitization. They recognize that all three effects are produced by a similar environmental conditioning (“pairing of a CS and a US”), and thus, should be united under an unnamed superordinate category. Presumably, this step is taken to distinguish these effects from other types of learning (e.g., instrumental conditioning). Below Pavlovian conditioning, they envision the mechanism of associative learning. Their analysis implies: (1) that protection from habituation and Pavlovian conditioning are concepts of the same type; and (2) associative learning is a mechanism, while protection from habitua-

tion is not. Yet, it would seem that a US-induced disruption in a particular process (habituation) would have mechanistic implications on par with a US-induced capacity to strengthen a connection (contrary to 2). Further, if associative learning is the subordinate mechanism to Pavlovian conditioning, what is the comparable subordinate mechanism to protection from habituation—protection from habituation? Similarly, is the mechanism underlying pairing specific enhanced sensitization, pairing specific enhanced sensitization? For both protection from habituation and pairing specific enhanced sensitization, the labels imply the functional mechanism. For Pavlovian conditioning to be a concept of the same type (for 1 to be true), it too would have to have equivalent mechanistic implications, presumably achieved through reference to the concept of associative learning. But what then, in their scheme, does the label “Pavlovian conditioning” add? It would seem that associative learning, protection from habituation, and pairing-specific enhanced sensitization are concepts of the same type. Reilly and Schachtman agree that all three are produced by similar environmental conditions. The only item missing is a name for the superordinate category, a name that refers to cases where the response observed is conditional upon “a history of pairing between the CS and US” (Staddon, 2005, p. 42). The traditional terms of Pavlovian or classical conditioning would seem appropriate, but Reilly and Schachtman cannot take this course because it requires broadening the definition of conditioning to include nonassociative mechanisms.

Contrary to our target article, Reilly and Schachtman suggest that other areas of psychology organize their text around procedures and behaviors rather than mechanisms. It is unclear to us what areas and texts they refer to because the perceptual and cognitive texts that we have rely on mechanism rather than methodology to organize the material. Our texts on perception include chapters on color vision, perceptual organization, movement, space perception, audition, and the other senses (e.g., Goldstein, 1999; Matlin & Foley, 1997; Schiffman, 2000). Popular cognitive texts include chapters on pattern recognition, attention, models of memory, imagery, expertise, reasoning, and language (e.g., Anderson, 2005; Solso, MacLin, & MacLin, 2005). In both instances, the material is being grouped according to the nature of the underlying process (mechanism).

Reilly and Schachtman attempt to discredit our focus on nonassociative mechanisms by suggesting that the “mechanisms of protection from habituation and pairing-specific enhanced sensitization remain to be determined” (p. 36). Is the counter to this that the mechanisms of associative learning have been determined? Does Kandel’s work not count as a sophisticated explanation of a mechanism that can generate pairing-specific enhanced sensitization? Do Reilly and Schachtman believe we actually know more about the functional and biological mechanisms that underlie associative learning? That, to us, would seem to be a difficult position to defend.

Regarding the nature of spinal learning, it was suggested that the weaker response observed to a CS- in a Pavlovian paradigm could reflect the development of conditioned inhibition. This is logically possible, but there is no evidence to support the proposal. Consequently, parsimony would favor maintaining an account based on simpler processes. As to their claim that we assume that the “associative basis of long-term habituation applies to spinal cases” (p. 36), we do not (Joynt & Grau, 1996), and this claim was based on prior research (Groves, Lee, &

Thompson, 1969). (The claim here is only that spinal habituation seems nonassociative in nature. As acknowledged earlier, associative learning has been shown to contribute to other examples of behavioral habituation.)

Reilly and Schachtman are correct in noting that tradition does not necessarily force us to ignore neuroscience. However, in practice, this has often been the outcome. Are we just pushing for neuroscience? No, we are not. We see a detailed description of the functional (psychological) mechanisms as central. Further, it is worth remembering that the domain of neuroscience is much broader and that much of that field works happily at the biochemical/biophysical level with little (or no) reference to the functional systems in which the entity under study might be embedded. Work that ignores function may be excellent science, and of profound long-term significance, but if it is not coupled to its function (specifying its contribution to behavior within a living organism), it is not neurofunctionalism.

Machado: Conceptual Confusion

Machado (2005) attempts to discount our approach by suggesting that we make some conceptual errors. For example, he claims that we failed to notice that Pavlovian conditioning within our framework functions as a superordinate category. But our summary figure (Figure 4) explicitly depicts this relation.

Machado then goes on to show that, if we semantically switch the meaning of our terms, confusion arises. This is hardly surprising. His example involves a CS that generates a CR-like response. Machado makes the reasonable assumption that the repeated presentation of the CS would lead to habituation and that the presentation of an extraneous stimulus counters this effect. Normally, in the absence of any evidence that the S-S relation matters, we would say that the extraneous stimulus produced a form of dishabituation. Machado suggests, instead, that we call this phenomenon protection from habituation. We would not endorse this step because, following Humphrey (1933), we require evidence that the CS-US relation matters. In the way we have used associative learning, pairing-specific enhanced sensitization, and protection from habituation, dependence upon the S-S relation is integral to their definition. Further confusion arises when Machado implies that evidence that a CS is sensitive to dishabituation, and that the S-S relation matters, is sufficient to conclude that the learning depends on protection from habituation. It is not. His example does not, in any way, discount an explanation of the S-S learning in terms of associative learning or pairing-specific enhanced sensitization. He is right to suggest that his example should arouse suspicion, but this is because it involves a semantic shell game and conclusions that do not follow.

He then suggests that our true opponents are the neuroscientists, not learning theorists. We disagree because we believe that the blame lies at our feet—with those who have spent their lives studying and characterizing learning. The neurobiologists have looked to us for guidance. If the leaders in our field argue that there is just one form of true learning, neurobiologists will focus on those examples that meet the learning theorists dictate. If our textbooks outline a perspective on learning that rang true 30 years ago, who can blame the neuroscientists for following the old course. Neuroscientists recognize that something is not working here—that the traditional description of learning does not provide a useful mapping to biological systems and fails to incorporate new findings.

Machado views our push for increased attention to the neurobiology of learning as a peculiar bias. Our claim is that a typical course on learning does not incorporate much of what has been discovered in the last 20 years and that many of these discoveries occurred within the area of neuroscience. Visits to conferences, attention to funding and hiring trends, and comparisons of relative output (in terms of number and impact) would seem to support our peculiar position.

Machado doubts that broadening the scope of a learning course to include greater emphasis on mechanisms and neuroscience will increase student enthusiasm for learning. Perhaps he is right—maybe students prefer the old. Stories of studies from the laboratories of Tolman and Sheffield routinely amuse the class. But amusement alone cannot be our criterion for inclusion. Adding structural biochemistry to a biochemistry curriculum probably does not make the course more entertaining, but it does make it more informative and up-to-date. Machado concludes that we must be exhausted with the teaching of learning and that we have mistakenly concluded the field itself is exhausted. He is right that we are disgruntled with the traditional way in which material is presented, but those who are tired do not push the field to change. Being dissatisfied is not the equivalent of tired. We have entered a new era of learning and we believe that it is time to retool (see <http://graulab.tamu.edu/j-grau/psyc606.html> for an example of how our recommendations have impacted our approach to teaching).

References

- Anderson, J. R. (2005). *Cognitive Psychology and Its Implications* (6th edition). New York: Worth Publishers.
- Blaisdell, A. P. Mechanism through methodology: No madness to the method. *International Journal of Comparative Psychology*, **18**, 23-27.
- Buerger, A. A., & Chopin, S. F. (1976). Instrumental avoidance conditioning in spinal vertebrates. *Advances in Psychobiology*, **3**, 437-461.
- Gallistel, C. R. (1980). *The organization of action: a new synthesis*. Hillsdale, N.J.: Erlbaum.
- Goldstein, E. B. (1999). *Sensation and perception* (5th edition). Pacific Grove, California: Brooks/Cole Publishing Co.
- Grau, J. W., Barstow, D. G., & Joynes, R. L. (1998). Instrumental learning within the spinal cord. I. behavioral properties. *Behavioral Neuroscience*, **112**, 1366-1386.
- Grau, J. W. & Joynes, R. L. (2001). Pavlovian and instrumental conditioning within the spinal cord: Methodological issues. In M. M. Patterson & J. W. Grau (Eds.), *Spinal cord plasticity: Alterations in reflex function*. Kluwer Academic Publishers.
- Grau, J. W., & Joynes, R. L. (2005). A neural-functionalist approach to learning. *International Journal of Comparative Psychology*, **18**, 1-22.
- Grau, J. W., Washburn, S. N., Hook, M. A., Ferguson, A. R., Crown, E. D., Garcia, G. G., Bolding, K. A., & Miranda, R. C. (2004). Uncontrollable stimulation undermines recovery after spinal cord injury. *Neurotrauma*, **21**.
- Groves, P. M., Lee, D., & Thompson, R. F. (1969). Effects of stimulus frequency and intensity on habituation and sensitization in acute spinal cat. *Physiology & Behavior*, **4**, 383-388.
- Holland, P.C. (1990). Event representation in Pavlovian conditioning: Image and action. *Cognition*, **37**, 105-131.
- Humphrey, G. (1933). *The nature of living in its relation to the living system* (pp. 165-179). New York: Harcourt, Brace & Co.
- Ji, R.-R., Kohno, T., Moore, K. A., & Woolf, C. J. (2003). Central sensitization and LTP: do pain and memory share similar mechanisms? *Trends in Neurosciences*, **26**, 696-705.
- Joynes, R. L., & Grau, J. W. (1996). Mechanisms of Pavlovian conditioning: The role of protection from habituation in spinal conditioning. *Behavioral Neuroscience*, **110**, 1375-1387.

- Joynes, R. L., Janjua, K., & Grau, J. W. (2004). Instrumental learning within the spinal cord: VI. The NMDA receptor antagonist, AP5, disrupts acquisition and maintenance of an acquired flexion response. *Behavioural Brain Research*, **154**, 431-438.
- Liu, G. T., Crown, E. D., Miranda, R. C., & Grau, J. W. (2005). Instrumental learning within the rat spinal cord: Localization of the essential neural circuit. *Behavioral Neuroscience*.
- Machado, A. (2005). Experimental methods and conceptual confusion. *International Journal of Comparative Psychology*, **18**, 28-33.
- Matlin, M. W., & Foley, H. J. (1997). *Sensation and perception* (4th edition). Boston: Allyn and Bacon.
- Patterson, M. M. (2001). Classical conditioning of spinal reflexes: The first seventy years. In J. E. Steinmetz, M. A. Gluck, & Paul R. Solomon (Eds.), *Model systems and the neurobiology of associative learning*. Mahwah, NJ: Erlbaum.
- Pittenger, C., & Kandel, R. R. (2003). In search of general mechanisms of long-lasting plasticity: Aplysia and the hippocampus. *Philosophical Transactions of the Royal Society of London [B]*, **358**, 757-763.
- Reilly, S., & Schachtman, T. R. (2005). Pavlovian conditioning requires ruling out nonassociative factors to claim conditioning occurred. *International Journal of Comparative Psychology*, **18**, 34-37.
- Rescorla, R. A. (1975). Pavlovian excitatory and inhibitory conditioning. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes: Volume 2, conditioning and behavior therapy*. Hillsdale, N.J.: Erlbaum.
- Rescorla, R. A. (1988 a). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, **43**, 151-160.
- Rescorla, R. A. (1988 b). Behavioral studies of Pavlovian conditioning. *Annual Review of Neuroscience*, **11**, 329-352.
- Schiffman, H. R. (2000). *Sensation and perception* (5th edition). New York: John Wiley & Sons.
- Sokoloff, G., & Steinmetz, J.E. (2005). A continuum of learning and memory research. *International Journal of Comparative Psychology*, **18**, 38-41.
- Solso, R. L., MacLin, M. K., & MacLin, O. H. (2005). *Cognitive Psychology* (7th edition). Boston: Allyn and Bacon.
- Staddon, J. (2005). Beyond method. *International Journal of Comparative Psychology*, **18**, 42-45.
- Timberlake, W., & Lucas, G. A. (1989). Behavior systems and learning: From misbehavior to general principles. In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Instrumental conditioning theory and the impact of biological constraints on learning* (pp. 237-275). Hillsdale, N. J.: Erlbaum.

Received October 28, 2004.

Revision received January 7, 2005.

Accepted January 7, 2005.