Challenges Facing Contemporary Associative Approaches to Acquired Behavior

Ralph R. Miller
State University of New York at Binghamton

Despite the considerable success of contemporary associative models of learning in stimulating new behavioral research and modest success in providing direction to both neuroscience and psychotherapy, these models are confronted with at least three challenges. The first challenge is to the assumption that animals encode only one or a few summary statistics to capture what has been experienced over many training trials. This assumption is contrary to overwhelming evidence that the brain retains episodic information. The second challenge is that the learning-performance distinction has been largely ignored. Most models erroneously assume that behavior is a nearly perfect reflection of what has been encoded. The third challenge is to account for interactions between stimuli that have been presented separately (e.g., stimulus interference) as well as between stimuli that have been presented together (e.g., stimulus competition).

The purpose of this review is to assess the common denominators of most contemporary associative models of learning as a group (as opposed to specific models), with an emphasis on the major challenges facing these models. By associative, I am referring to models designed to account for Pavlovian responding, although these models have often been extended quite successfully to instrumental behavior as well. The qualifier contemporary is necessary because new associative models are likely to be proposed in the future that do not subscribe to these common denominators. Miller and Escobar (2001) discussed the dangers of contrasting whole families of models rather than specific models. Here I try to minimize this danger by limiting my remarks to families of contemporary models rather than families of models including past and future instances. There will also be some discussion of the major alternatives to contemporary associative models in order to appreciate how well they address the problems confronting associative models. The paper is organized around three basic problems: the assumption that animals encode only summary statistics about prior experiences, the assumption that behavior is a nearly veridi-

cal reflection of what is encoded, and the need to account for interactions between stimuli presented separately during training as well as between stimuli presented together during training.

Summary Statistics

Associations vs. modern associative theories

In my view, all models of learning explicitly or at least implicitly assume the existence of associations of some sort as the building blocks of memory even if they do not use the term. Examples of models that circumvent the term association include those in which timing is the central construct (e.g., Gallistel & Gibbon, 2000; Gibbon & Balsam, 1981). These models overtly deny the existence of associations; for example, Gallistel and Gibbon speak not of associations but of symbolic representations of event dyads (composed of a cue and an outcome) including their temporal relationship. However, at test, presentation of a cue activates (through inferential processes) differential expectations of the outcome based on each prior experience with that cue-outcome dyad. The implied links that hold the cue-outcome symbolic representations together are functionally very similar to what is meant by an association between the cue and outcome. Here I use cue and outcome to refer to the first and second experienced event, respectively, in a dyadic sequence, cues being equivalent to conditioned stimuli and outcomes being equivalent to unconditioned stimuli except they need not be biologically significant. Another group of models that superficially circumvents the construct of associations is based

on contingency, in which subjects encode only the frequency of occurrence of different types of events (e.g., cue and outcome present, cue alone present, outcome alone present, and neither cue or outcome present; Rescorla, 1968). However, the encoding of the cue and outcome as being simultaneously present constitutes an association. Thus, at least primitive associative constructs appear to be ubiquitous in models of learning.

Importantly, modern associative theories not only assume that prior experience is encoded in associations but that the associations are strengthened by repeated trials (i.e., recurrences of the same events). For each specific cue and outcome dyad, the mental consequence of another [repeated] pairing (i.e., a trial) takes the form of an up-dating of a single summary statistic as in the Rescorla-Wagner (1972) model (i.e., associative strength of the target cue [Vx]) or a few summary statistics as in the Pearce and Hall (1980) model (i.e., excitatory associative strength of the target cue [Vx,exc], inhibitory associative strength of the target cue [Vx,inh], and associability of the target cue [readiness to learn something new about the cue, "x]). For example, the Rescorla-Wagner model posits that subjects, after each trial with a target cue, up-date the associative status of that cue from its pretrial value according to a linear equation: associative value of the cue after the trial is equal to what it was before the trial plus a change due to what happened on that trial. Moreover, the change in associative value of the cue on that trial is a direct function of the salience of the target cue, the salience of the outcome, and the difference between experienced outcome and the outcome expected on that trial based on all cues present during that trial. Importantly, the model assumes that all a subject retains after exposure to a cue (X), which has a history of sometimes being paired with a specific outcome, is a cue X-outcome association (VX) that can be represented by a single number; there are presumably no memories of individual trials. This is analogous to prototypes theories of categorization (e.g., Reed, 1972) in which a single memory is repeatedly modified by successive training trials, and can be contrasted with instance [snapshot] theories of categorization (e.g., Logan, 1988) in which each trial creates a separate memory, with repeated trials creating very similar but still distinct memories. The assumption that only summary statistics are encoded is rarely questioned, but is contrary to considerable data. Perhaps the most compelling argument in favor of the instance view is the evidence for episodic-like memories, that is, associative memories of specific instances of events that include not only what happened, but where and when the events happened. More generally, episodic-like memory is simply an extreme example of what is sometimes referred to as source memory (memory for where and when information was obtained).

Are humans unique, with nonhumans relegated to summary statistics?

The existence of episodic memories in humans has long been accepted by memory researchers (Tulving, 1972), but has been questioned with respect to nonhuman animals (e.g., Tulving, 2002). However, in recent years, numerous researchers have concluded that at least some nonhuman species have episodic-like memory based on demonstrations that these animals appear to encode not only procedural information[1], but what, when, and where events happen. For example, Clayton and Dickinson (1998) have documented episodic-like memory in the food stashing behavior of scrub-jays (which type of food was cached where and how long ago), and Babb and Crystal (2005) and Eacott, Easton, and Zinkivskay (2005) have reported similar behavior in rats (for a review, see Zentall, 2005). Tulving (2002) questioned whether the demonstrations of episodic memory in nonhumans involves autonoetic awareness, knowledge of self, and recognition of subjective time, all of which he regards as essential components of episodic memory. Although evidence of these additional attributes of memory has yet to be obtained for nonhumans (a difficult task due to the absence of language), it seems implausible and homocentric to assume that some form of them will not be demonstrated in the future, as prior research has shown that evolution rarely results in sharp lines in basic behavior capabilities between similar species (as acknowledged by Tulving, 2002). More important for present purposes, these additional attributes are not necessary to make the point that nonhumans retain memories of specific prior events rather than merely summary statistics. Evidence strongly suggests that nonhuman subjects often store the what, when, and where of each experienced event (i.e., episodic-like memory) even if this occurs without the full features of human episodic memory.

Moreover, through a process akin to second-order conditioning, nonhuman animals appear able to integrate different temporal memories (e.g., Matzel, Held, & Miller, 1988) and different spatial memories (Blaisdell & Cook, 2004), provided the memories to be integrated share a common stimulus, thereby creating temporal and spatial relationships between stimuli that were never actually paired. That is, subjects taught separate A-B and B-C temporal or spatial relationships behave as if they have knowledge of an A-C temporal or spatial relationship. Although temporal learning surely includes subjects encoding when events occur with respect to other events within what is normally construed as a trial (i.e., temporally and spatially proximal events, see Healy, 1998; Savastano & Miller, 1998, for reviews), subjects also encode when an event occurs with respect to the arrow of time on a larger scale than within individual trials (for a review, see Crystal, in press). This temporal component of each discrete memory makes the memories of each successive event dif-

ferent even if all other external and internal stimuli are unchanged (which is unlikely). That is, even ostensibly identical training trials occur at different moments in the river of time. Thus, each trial is at least in some sense different, and consequently even contemporary associative theories would anticipate new memories being formed following each trial rather than a single memory being repeatedly updated. This is problematic for models that assume that training trials can be repeated and that summary statistics are all that animals encode. However, there are at least two ways that contemporary models of associative learning might circumvent this difficulty.

The first is to assume that subjects learn not about a single complex stimulus with many attributes (i.e., elements) processed as a single stimulus, but about each element independently along with within-compound associations that link these elemental representations. This elemental approach minimizes the problem of attributes that change from trial to trial making each trial different. That is, at least for some elements, the successive trials should not vary. However, this is not a fully adequate resolution of the problem posed by stimulus variation between successive trials because evidence of episodic memory could not be explained without assuming that each element had a different time tag for each successive trial on which it was presented. Thus, an elemental approach does not really circumvent the existence of episodic memory being inconsistent with repeated trials adding strength to existing associations.

A better defense of the association-strengthening view of contemporary associative models is provided by the position, maintained by most researchers concerned with human memory, that there are multiple memory systems (e.g., Squire, 2004). That is, although subjects have episodic memories, they may have other types of memories as well. This possibility is consistent with evidence suggesting that different types of memory are dependent upon different neuroanatomical sites and transmitter systems. Impressive double dissociation experiments support this differentiation of memory systems (for a review, see Squire & Kandel, 1999). In the study of human cognition, the conventional opposite of episodic memory (more generally, memory that includes source knowledge) is semantic memory, which lacks source knowledge (e.g., Anderson & Bower, 1973). But in simple multi-trial Pavlovian and instrumental learning tasks, the most notable of these other memory systems is procedural memory (in the broad sense), which also lacks source knowledge. Hence, procedural memory is compatible with associative models that assume successive similar trials update memories rather than create new memories.

Appealing to procedural memory as mediating associative learning, however, encounters several problems. Perhaps chief among them is the assumption that trials are repeated.

Inherent to models that depend upon summary statistics is the assumption that training trials can and sometimes do repeat themselves. Obviously multiple distinct memories are formed when the successive training trials are sufficiently dissimilar. But no two trials are ever exactly the same. Variables both external and internal to the subject are apt to change from trial to trial, not to mention the previously discussed unavoidable changes in temporal context due to the irreversible flow of time. If successive training trials differ, then even associative theories would expect distinct memories to be formed. An absence of repeated trials would make moot the central premise of contemporary associative theories, that is, summary statistics based on repeated trials. Contemporary associative theory might try to deal with this through generalization between similar memories at the time of test. But, if each trial were independently represented, this would render meaningless the basic assumption that organisms store summary statistics concerning identical trials, rather than memories of each individual trial. All contemporary associative models assume summary statistics, whereas few explicitly address stimulus generalization. Pearce (1987) is an example of a model that does formally account for generalization, but it too centrally assumes that training trials often repeat themselves with accompanying updating of summary statistics. Thus, the failure of trials to repeat themselves undermines models dependent on summary statistics.

Such problems for multiple memory systems that include procedural memory notwithstanding, one is still faced with the dissociation of behavioral tasks through lesioning of selective anatomical sites and chemical manipulation different neurotransmitter systems. However, these demonstrations do not directly speak to the representational form of the information that is encoded in these tasks. As a function of the informational nature of a specific event, memories of individual instances of prior experience, rather than mere summary statistics, may well be encoded in different neuroanatomical sites based on different neurotransmitters. The data itself is compelling, but I believe that the interpretation of the data as support for a unique procedural memory system dependent on summary statistics is less convincing than is often assumed. A direction for future research with both humans and nonhumans would be to assess procedural memory tasks to determine if there are memories of individual trials (or trial types) underlying these behaviors.

Recency-to-primacy shifts

All contemporary associative models of learning predict recency effects given conflicting phasic training. That is, more recent events are expected to result in an updating of memory that overrides previously acquired conflicting memories (sometimes called catastrophic forgetting). The

demonstrations of this are innumerable (e.g., Lopez, Shanks, Almaraz, & Fernandez, 1998). Extinction and countercon-ditioning provide two well known examples. Consider ex-tinction: Sufficient nonreinforced cue presentations (i.e., extinction treatment) following cue-outcome conditioning trials will result in a loss of most of the conditioned respond-ing acquired during the reinforced trials. Similarly, if the order of the two phases are reversed, that is, if reinforcement follows nonreinforcement, ultimately conditioned respond-ing will be observed (although latent inhibition might delay its emergence). Both of these phenomena are recency ef-fects. Prediction of recency effects is regarded as one of the great successes of contemporary associative models. But, for pragmatic reasons, most studies have used relatively short retention intervals (minutes, hours, or a few days at most) and the same context for training and testing. When appreciable retention intervals are inserted between treat-ment and testing, the effects of extinction treatment wane and conditioned responding returns (i.e., spontaneous re-covery, Pavlov, 1927; Rescorla, 1997; Stout, Amundson, & Miller, in press; Wheeler, Stout, & Miller, 2004). With re-versal of the order of the two phases of treatment, increases in the posttreatment retention interval in a latent inhibition paradigm (i.e., after the reinforced trials) often result in a decrease in conditioned responding (provided the retention interval is spent outside of the treatment context so there is no extinction of associations to the context; De la Casa & Lubow, 2000, 2002; Stout et al., in press; Wheeler et al., 2004). These shifts from a recency effect to a primacy effect are seen not only with increases in retention interval, but also with changes in the physical context between treatment and testing (i.e., AAB renewal, Bouton & Ricker, 1994). Simi-lar recency-to-primacy shifts have been observed in coun-terconditioning situations (Bouton & Peck, 1992). There is nothing surprising about these examples of recency-to-pri-macy shifts. Such shifts are seen across a much wider range of tasks with both human and nonhuman subjects (Bjork, 2001; Neath & Knoedler, 1994; Postman, Stark, & Fraser, 1968). Surely waxing primacy effects are less common than waning recency effects, but increases in primacy effects are not uncommon, and sometimes primacy effects are observed even without a long retention interval (Dennis & Ahn, 2001). Within my own laboratory, we have examined the conse-quences for conditioned responding by rats of added pre-sentations of the outcome without a preceding signal either before or after cue-outcome pairings (i.e., reinforced trials). That is, we compared [a version of] the well known effects of exposure to an outcome alone prior to Pavlovian condi-tioning trials (the US-preexposure effect) with [a version of] the little known effects of exposure to the outcome alone fol-lowing Pavlovian conditioning trials (Urushihara, Wheeler, & Miller, 2004), and observed recency effects when testing soon followed treatment and primacy effects when testing

was appreciably delayed after the completion of treatment. More concretely, we exposed rats to pairings of a click train (as a cue) with a tone (as a outcome) and later paired the tone with a footshock so the rats exhibited fear of the clicks. Ad-ditionally, some rats experienced tone-alone presentations prior to the click-tone pairings and other rats experienced the tone-alone presentations following the click-tone pair-ings. When the retention interval was short (a few days), the tone-alone presentations following the click-tone pair-ings decreased responding to the clicks more than did the tone-alone presentations preceding the tone-click pairings. But with a long retention interval (a few weeks), the tone-alone presentations preceding the tone-click pairings had the more deleterious effect on conditioned responding. This constitutes a clear demonstration of a shift with increasing retention interval from stronger effects of the most recent [relevant] training to stronger effects of the initial [relevant] training, a recency-to-primacy effect.

Central to the present assessment of contemporary asso-ciative models, recency-to-primacy shifts that occur with changes in the spatial or temporal context between the last phase of treatment and testing, such as those described above, indicate that associative accounts of recency effects that assume earlier acquired information is erased are funda-mentally in error. The reversion to behavior compatible with initial training indicates that representations of initial train-ing were retained rather than obliterated as is assumed by models that posit retention of only summary statistics which are successively updated. That is, these models assume that summary statistics are updated to reflect the last phase of training, irrevocably replacing information concerning ear-lier phases of training. Reversion to behavior indicative of initial training without further training denies the irrevocable loss of initial information. Seemingly, there is no simple fix for this failure, which is inherent to any model that assumes only summary statistics control behavior in simple learn-ing situations. Even if subjects are assumed to have both procedural and episodic-like memories, seemingly it is the episodic-like ones that influence conditioned responding be-cause procedural memory lacks information about the order of acquisition of conflicting information that appears to be necessary to account for recency-to-primacy shifts. Accep-tance of the existence of episodic memories does not itself provide a full account of recency-to-primacy shifts. But the assumption of episodic memories does at least provide a res-ervoir for the information that is revealed by such a shift in behavior.

Although the present focus is on assessing associative models, it is interesting to digress momentarily to consider the functional significance of recency-to-primacy shifts. Re-cency effects have obvious survival value in that contingen-cies change and recency effects would keep an animal in

tune with the immediately prevailing contingencies. Moreover, it is reasonable to assume that, with an increasing passage of time since the last training trial, it is less likely that the next trial will be consistent with the last trial as opposed to an average of all of the preceding trials. An animal designed to process information based on this assumption, assuming that it retains information concerning all prior trials, should switch from behavior indicative of recent events to behavior characterized by the mean value of all prior trials, but not revert to behavior reflecting initial training. Hence, the functional value of initial training being privileged (i.e., primacy effects) is unclear. Perhaps primacy effects have no functional role; that is, primacy effects may be an epiphenomenal consequence of processes that are functional in their other consequences.

The traditional process used to provide a mechanistic account of primacy effects is that initial information can be given more rehearsal time due to limited competition for rehearsal capacity by related information (e.g., Atkinson & Shiffrin, 1968). This view received support from the finding that instructed overt rehearsal enhanced primacy effects (Rundus, 1971). However, there is compelling evidence that better rehearsal is not a fully adequate explanation of primacy effects. For example, stimuli such as kaleidoscope images that seemingly defy rehearsal still yield primacy effects (e.g., Wright, Cook, Rivera, Shyan, Neiworth, & Jitsumori, 1990). Other mechanisms proposed to account for primacy effects, such as distinctiveness (e.g., Murdock, 1960), likely do contribute, but there is compelling evidence that no one of them is all encompassing (for reviews, see Hogarth & Einhorn, 1992; Wright, 1998). More contemporary accounts simply speak of a reduction in retroactive interference with increasing retention intervals (due to waning recency) allowing proactive interference to be evidenced (Wright, 1998). When there are multiple conflicting associations to the same cue, the spatiotemporal context or discrete cues present at testing, by virtue of their similarity to one or another training circumstance, likely act as occasion setters favoring the retrieval of one association as opposed to another (Bouton, 1993, Miller & Oberling, 1998). However credible this view is, it does not explain the emergence of primacy effects as opposed to equal weighting of all prior associations given long retention intervals. Perhaps we will have to accept initially received information being privileged as a primitive. Notably, all of these accounts presuppose the retention of the initially learned material, which is contradictory to the summary statistics assumption of current associative theories in the animal learning tradition.

Information capacity as an argument for summary statistics

Contemporary associative theories are conservative with respect to assumed long-term memory capacity. That is, retention of summary statistics clearly would place far fewer demands on memory capacity than would a model in which individual episodic memories are retained. But one might ask if there is really a need to be conservative. Long ago, Wooldridge (1963) performed rough calculations concerning the information capacity of the brain based on a single synapse for each bit of information, and concluded that mammals retain far more information than this assumption allows. This simply underlines how little we knew then (and know now) about how information is encoded in the brain. The one thing that is obvious is that actual storage capacity is enormous (although not without limit; see Cook, Levison, Gillett, & Blaisdell, 2005), and the capacity of vertebrate memory appears to be so high that for most purposes it is not a meaningful constraint on the different ways that acquired information might be stored. However, some models are challenged by even a very high memory capacity. For example, Gallistel and Gibbon (2000) propose that all stimulus dyads, along with each dyad's interstimulus interval, are encoded, with no limit based on the cue-outcome temporal separation. This leads to an incredible multitude of stimulus-stimulus intervals (effectively, associations with interstimulus intervals attached) which increases factorially with events encountered (roughly, time lived). This implicit assumption of almost limitless memory capacity contrasts sharply with the minimalist implications of contemporary associative models. In terms of demand on memory capacity, between these two extremes are models of learning that assume close contiguity is necessary to form an association but that organisms retain memories about different types of previously experienced events rather than merely memories of summary statistics (e.g., contingency models). In summary, it appears that the argument that the summary statistic viewpoint is supported by its being frugal with memory is not compelling. More generally, the existence of episodic memories and their clear role in conditioning tasks present a challenge to contemporary associative theories that rely exclusively on summary statistics.

Contingency theories as an alternative to associative theories

One might ask about models of learning that avoid the use of summary statistics. Contingency theories provide an approach that was popular in animal learning many years ago (Rescorla, 1968), and is still frequently used in analyzing human thought and behavior such in causal attribution (e.g., Cheng, 1997). At their heart, contingency theories assume that subjects retain memories of the different types of trials that have occurred and the frequency of each type. For example, consider a simple situation in which a tone is sometimes paired with a footshock. In contingency theories,

fear expressed to the tone is assumed to be positively correlated with the number of tone-shock pairings and also the number of trials on which no tone and no shock occur. Additionally, fear is assumed to be negatively correlated with the number of times the tone is presented without the shock and the number of times the shock is presented without the tone. Contingency theory was initially proposed to account for the response degrading consequences of unsignaled outcomes (e.g., shocks) interspersed among cue-outcome pairings (i.e., degraded contingency treatment), and also proved able to explain the response attenuating effects of outcomes presented alone prior to cue-outcome pairings (i.e., the US-preexposure effect) as well as of cues presented alone prior to cue-outcome pairings (i.e., latent inhibition) and interspersed amidst the cue-outcome pairings (i.e., partial reinforcement acquisition effects). However, associative theories soon provided an account of the degraded contingency and US-preexposure effects (e.g., Rescorla & Wagner, 1972) along with latent inhibition (e.g., Miller & Matzel, 1988; Pearce & Hall, 1980; Wagner, 1981). Contingency theory can be thought of as depending on retention of the frequency of different trial types, or, with a little elaboration, of individual trials that on each test trial are effectively counted as a function of trial type. The former, more common, version of contingency theory is seen to also use summary statistics (a number to represent the frequency of each trial type), albeit more statistics than are assumed by most contemporary associative models. Moreover, what is stored according to such contingency models does not include the temporal information that is inherent in episodic-like memories. The latter version, with each specific trial encoded, is more compatible with the observation of episodic memories. However, both versions encounter a number of problems, among them being unable to address trial order effects because frequencies of trial types are encoded without any information about the order of the different trials. Not only do these contingency models, like traditional associative models, fail to account for waning recency effects and waxing primacy, but they do not even anticipate the recency effects that are ordinarily observed with short retention intervals and no change in context prior to testing, something that traditional associative models such as Rescorla-Wagner (1972) correctly anticipate. The richer version of contingency theory with its retention of memories of individual trials, complemented with added information concerning the time at which trials occurred, is a direction that future researchers might profitably pursue (see Lopez et al., 1998, for a detailed discussion of potential modifications).

The Learning-Performance Distinction

Ever since Tolman (e.g., 1932), researchers have had some awareness of the learning-performance distinction. The primary variable upon which Tolman focused concerning the transformation of stored knowledge into behavior was motivation. Today we all acknowledge that motivation is essential for the expression of what has been learned. But associative theories in the animal learning tradition rarely go beyond motivation in differentiating encoded knowledge from behavior. Indeed, they usually only acknowledge rather than formalize the role of motivation, but that is still more attention than they give to other so-called performance variables. Unlike associative theories framed to explain human performance, there is ordinarily little concern for retrieval processes or response generation rules. Learning is an intervening variable; all we ever see is a change in behavior as a consequence of prior experience. Consistent with the misguided name learning theory and inconsistent with the actual goal of explaining acquired behavior, most modern associative theories in the animal tradition emphasize the learning (i.e., acquisition) process per se and are virtually silent concerning the transformation of acquired information into behavior. For example, Rescorla and Wagner (1972) simply say that responding is monotonically related to associative strength. This undermines the assertion that the Rescorla-Wagner model is quantitative in more than anticipating rank order of behaviors resulting from different treatments. That is, the model is quantitative only up to the point of computing associative strength, which is an intervening variable that cannot be directly measured. Some other associative models such as Wagner (1981) go a bit further in discussing the expression of behavior, but in explaining differences in behavior even these models clearly place a far greater emphasis on acquisition processes than on expression processes that are uniquely active at test. They assume that what is observed in behavior is a reliable reflection of what the subject has encoded. The failure to do more than predict rank-order differences in behavior is a major weakness of most contemporary association models in the animal tradition.

In contrast, in the study of human information processing, ever since the so-called cognitive revolution, retrieval, decision making, and response production have been given as much or more attention than acquisition. Landmark examples include Tulving and Pearlstone's (1966) distinction between available and accessible memories, with available memories being encoded but inaccessible given the immediately prevailing retrieval cues, and Tulving and Thomson's (1973) discussion of the importance for effective retrieval of common cues being similarly represented during training and testing. Spear (1973) differentiated between reversible and permanent performance failures on memory tests with nonhuman subjects; he termed the distinction lapse vs. loss. Tulving and Spear do not provide quantitative accounts of retrieval and response, but they at least discuss the variables

that seemingly influence retrieval.

A few associative models from the animal tradition, however, have given more weight to retrieval processes than acquisition. Bouton's (1993) retrieval model and Miller and Matzel's (1988) comparator hypothesis are models of this sort; and Gallistel and Gibbon's (2000) model gives greater emphasis to decision rules for responding than acquisition (although these last authors would argue that their model is not associative). Bouton, focusing on phasic reinforcement and nonreinforcement trials with a single stimulus, assumed that when there were different phases of training in which contradictory information was provided, the informational content of each phase was separately stored. Then, the presence on a test trial of occasion setting stimuli that were also present during one or another phase of training disambiguates the different meanings of the retrieval cues and determines which information set will be expressed on that test trial. This approach works particularly well for spontaneous recovery from extinction treatment (recovery as a function of the retention interval, which is a recency-to-primacy phenomenon), renewal (recovery from extinction effected by a change in context between extinction treatment and testing), and recovery from counterconditioning (i.e., recovery from retroactive outcome interference [e.g., Tone-Food, Tone-Shock, test on Tone for responding anticipatory of food] as a function of retention interval or context change). In these phenomena we see differences in responding to the target cue on a test trial that do not depend on differences in what was encoded at the time of training.

In contrast to Bouton (1993), Miller and Matzel's (1988) comparator hypothesis (for an update of this model, see Denniston, Savastano, & Miller, 2001) speaks most directly to cue interactions arising from training trials on which multiple cues are present, including cue competition (e.g., overshadowing and blocking) and conditioned inhibition. This model posits that acquisition is based on simple contiguity and that these phenomena arise from memory interactions that occur at the time of testing. In the framework of the comparator hypothesis, conditioned responding to a test cue is the result of a comparison made at the time of testing between two independently activated representations of the outcome. Responding is positively correlated with the degree to which a representation of the outcome is directly activated by the test cue (which reflects the strength of the test cue-outcome association), and negatively correlated with the degree to which a representation of the outcome is indirectly activated by the test cue conjointly through the association between the target cue and other stimuli present during training and the association between the other cues and the outcome (i.e., target cue --> other cue --> outcome). For example, according to the comparator hypothesis overshadowing is due to the association between the overshad-owed and overshadowing cues in conjunction with the association between the overshadowing cue and the outcome serving to indirectly activate a representation of the outcome (i.e., test cue--> overshadowing cue--> outcome, where the test cue is the overshadowed stimulus). This indirectly activated representation attenuates conditioned responding to the overshadowed cue which is otherwise promoted by the representation of the outcome that is activated directly by the association between the overshadowed cue and the outcome. Importantly, the association between the overshadowed cue and the outcome is assumed to be acquired unimpaired during overshadowing treatment. Additionally, within the comparator model, behavior indicative of conditioned inhibition (responding as if there is an expectation of no outcome) is due neither to a negatively valued association nor a cue-no outcome association, but to a comparison of multiple simple excitatory associations. Presumably, the target cue (i.e., the conditioned inhibitor) does not directly activate a representation of the outcome because this cue has never been paired with the outcome. But the conditioned inhibitor is able to indirectly activate a representation of the outcome as a result of the inhibitor's association to other cues that were present when it was trained and the association of these other cues to the outcome.

Notably, Bouton's (1993) retrieval model as well as Miller and Matzel's (1988) comparator hypothesis still assume the existence of summary statistics at least for simple dyadic relationships, which makes these models open to the same criticism previously leveled at more traditional associative models within the animal tradition. Obviously, any complete model must speak to what happens both at the time of training and at the time of testing. Gallistel and Gibbon (2000, as well as other proponents of timing models, see Church, 1989, for a review) go a step further than Bouton or Miller and Matzel in that they posit the encoding of individual events, along with the intervals between them and the order in which the trials occurred. However, a central weakness of Gallistel and Gibbon's model is seemingly the assumption that all dyads enter into association independent of their contiguity. To record a lifetime of memories, this assumption requires a memory store much larger than any other model. Moreover, this model appears to reject the well established principle of contiguity in that the temporal relationship of all stimulus dyads are encoded regardless of the interval between them.

The point to be made here is that, with the few exceptions mentioned above, most contemporary learning theories in the animal tradition erroneously focus on acquisition processes and fail to give adequate weight to latent information, that is, information that is stored but not immediately expressed. This point is related to the preceding argument concerning the inadequacy of summary statistics, in that the

episodic memories that animals apparently encode are often latent in the situations ordinarily examined in the laboratory owing to the procedures used. For example, in a situation in which contradictory information about a cue is provided in successive phases of training, information concerning the initial phase is latent when testing occurs soon after a later phase of training in the context used for later training.

When discussing the mechanism responsible for the absence of acquired behavior that might be expected based on prior experience, it is sometimes useful to differentiate between two forms of behavioral deficits. The first of these is behavior arising from information that was previously expressed (or was at least expressible if a test occurred, e.g., memories of initial training prior to extinction, prior to counterconditioning, and prior to backward blocking). The second of these is behavior arising from information that was never expressed as conditioned responding (e.g., memories concerning overshadowed cues, forward blocked cues, and cues subjected to latent inhibition treatment). Performance deficits based on information previously expressed that is now not being expressed arise from either a permanent loss of previously stored information or a reversible expression deficit. In contrast, performance deficits with respect to information that the subject never previously expressed presumably arises from an acquisition failure or a reversible expression deficit.

Absences of behaviors that were previously evident

Among the clearer successes of most contemporary associative theories are correct predictions concerning basic extinction, enhancement of extinction as a result of extinction treatment occurring in the presence of a second excitatory cue (e.g., Rescorla, 2000), and protection from extinction as a result of extinction treatment occurring in the presence of an inhibitory cue (e.g., Soltysik, Wolfe, Nicolas, Wilson, & Garcia-Sanchez, 1983). These phenomena can be explained equally well by either acquisition-focused (e.g., Rescorla & Wagner, 1972) or performance-focused associative models (e.g., Denniston et al., 2001). However, other phenomena related to extinction such as spontaneous recovery (a recency-to-primacy shift, Pavlov, 1927), renewal (recovery induced by a change in context from that of extinction, Bouton & Bolles, 1979), reinstatement (recovery induced by exposure to the outcome alone, Rescorla & Heth, 1975), and concurrent recovery (recovery induced by conditioning of a highly dissimilar cue to the original outcome, Weidemann & Kehoe, 2004) indicate that extinction is not erasure of stored information as is assumed by some associative models (e.g., Rescorla & Wagner, 1972). Associative models that assume extinction treatment establishes an inhibitory association (e.g., Pearce & Hall, 1980; Wagner, 1981) fare better with all of the above mentioned extinction phenomena except

concurrent recovery, provided a few assumptions are made about conditioned inhibition being more labile than conditioned excitation. Notably, models that treat extinction as the erasure of associations assume that extinguished memories are irrevocably lost, whereas models which appeal to inhibition to account for experimental extinction assume that the memory of initial reinforcement is present but silent owing to interference by the inhibitory association that is formed during extinction. Bouton (e.g., 1993) has gone furthest in developing the view that behavioral extinction and counterconditioning result from interference by inhibitory associations. In his framework, counterconditioning (e.g., Tone-Food in Phase 1 followed by Tone-Shock in Phase 2) not only results during Phase 2 in the formation of an excitatory Tone-Shock association, but also in the establishment of an Tone-Food inhibitory association. Unfortunately Bouton's retrieval model has not yet been presented as a formalized general model of learning. That is, it currently stands as a narrowly focused model used to account for interference seen when a cue is paired with different outcomes (extinction, latent inhibition, counterconditioning, and interference between outcomes), and is otherwise not invoked. Empirically, we see that the preponderance of data indicates that much if not all of these deficits reflect memories being rendered silent rather than erased, a point missed by contemporary acquisition-focused models of learning.

Other instances of behavioral differences that seemingly vanish, but are not due to an irreversible loss of information, are seen in what are called path dependent phenomena. Path independence refers to situations in which different subjects, despite prior differences in treatments and performance (i.e., different behavioral paths), exhibit common behavior and then alter their behavior in identical ways given the same new treatment. This contrasts with path dependence in which prior differential treatments and behavior cause subjects to alter their behaviors in different ways despite identical behavioral starting points and identical new treatment. Most acquisition-focused associative models of learning that rely on summary statistics assume that, if two subjects behave in the same fashion, they have stored the same summary statistics and consequently will alter their behavior in an identical manner given the same additional training, that is, path independence is anticipated. However, there are many examples of path dependence in the literature. For example, reacquisition after acquisition followed by extinction, or reacquisition after acquisition followed by counterconditioning, are examples of path dependence if extinction and counterconditioning successfully eliminated the originally acquired behavior and then the originally acquired behavior can be restored more rapidly than it took in original training, which is often the case. There are a large number of additional instances of path dependence. For ex-

ample, Brown-Su, Matzel, Gordon, and Miller (1986) found that rats asymptotically trained with a small reinforcer (tone-weak shock) and rats sub-asymptotically trained with a large reinforcer (tone-strong shock) were differentially sensitive to further training (tone-medium shock) despite starting from a common behavioral baseline (see Miller & Matzel, 1987, for a review). Path dependent phenomena in general provide numerous examples of behavior that does not accurately reflect stored information.

Absences of behaviors that might be expected but were never evident

All contemporary associative theories (acquisition- and performance-focused) can account for the commonly observed basic interactions between cues presented together (e.g., cue competition [including overshadowing and blocking], and conditioned inhibition). The ability to explain basic cue competition and anticipation-related phenomena like superconditioning (enhanced conditioned responding as a result of training a cue in the presence of a conditioned inhibitor, e.g., Wagner, 1971) and the overexpectation effect (reduced conditioned responding as a result of reinforcing a previously trained cue in the presence of a second previously trained cue, e.g., Rescorla, 1970) were impressive successes of these models. In contrast, few cognitive theories that depend on higher-order cognitive processes address cue competition beyond paraphasing the phenomena. Cheng and Novick's (1992) and Spellman's (1996) contingency theories do provide accounts of these phenomena by positing that subjects employ conditional probabilities of outcomes based on different combinations of cues present on each trial. Inference theory (DeHouwer & Beckers, 2002; Lovibond, 2003) also claims to explain these phenomena. But the predictions of inference models depend on subjects' assuming that outcomes (i.e., effects) have only one cause. Thus, these models speak to causal relationships, but not relations that are merely predictive (all causal relations are predictive as well as causal). It seems clear that blocking is stronger when the cues are causes rather than mere predictors; (DeHouwer, Beckers, & Glautier, 2002; Pineno, Denniston, Beckers, Matute, & Miller, 2005), but in these same reports there was evidence of blocking between mere predictors. That is, stimulus interactions do seem to occur between cues that are not causes.

The prediction of cue competition between predictive cues was a great initial success of contemporary associative theories. In these accounts, the target cue-outcome association was not formed and hence should not be recoverable by any treatment short of further training with the target stimulus. But since then recovery from cue competition has been effected through at least three manipulations.

The first of these manipulations is posttraining extinction of the competing stimulus (e.g., Kaufman & Bolles, 1981; Matzel, Schachtman, & Miller, 1985). If a target cue is trained in the presence of another cue, subsequent extinction (associative deflation) or reinforcement (associative inflation) of the companion cue often increases or decreases, respectively, behavioral control by the target cue. Such phenomena are collectively called retrospective revaluation. Increases in responding to the target cue as a result of posttraining deflation of a companion stimulus are relatively easy to obtain (e.g., Denniston, Savastano, Blaisdell, & Miller, 2003), but a few failures to obtain the effect have been reported (e.g., Holland, 1999) suggesting that this effect like most effects is parameter dependent (Shevill & Hall, 2004). In contrast, decreases in responding to a target cue as a result of posttraining inflation of a companion stimulus is rarely observed when the target cue signals a biologically significant outcome such as food or footshock (e.g., Grahame, Barnet, & Miller, 1992; Miller, Hallam, & Grahame, 1990), but can be observed if the procedure is embedded in a sensory preconditioning procedure so that the target cue does not have the opportunity to control behavior until after the inflation treatment is complete. An example of this is backward blocking as reported in rats by Denniston, Miller, and Matute (1996; also see Miller & Matute, 1996). They found backward blocking when a target stimulus X (e.g., a click train) in compound with a companion stimulus A (e.g., a white noise) was initially paired with an innocuous outcome B (e.g., a flashing light, that is, AX-B trials), followed by the associative inflation of A (i.e., A-B pairings in the absence of X). Finally, B was paired with a footshock. Then subjects were tested for fear of X. When the rats were tested for fear of X, the A-B pairings were seen to have reduced fear of X relative to a control group lacking the A-B pairings. Apparently organisms behave conservatively and do not readily surrender acquired behavior relevant to biologically significant outcomes. This qualifier notwithstanding, retrospective revaluation phenomena stand as evidence that blocked or overshadowed conditioned responding can be recovered without further training with the target cue. Miller and Matzel's (1988) comparator hypothesis was initially unique is explaining retrospective revaluation, claiming that the target memory was present all along but latent. That is, cue competition was viewed as a performance deficit rather than a failure to acquire the target association. However, subsequently Dickinson and Burke (1996) and Van Hamme and Wasserman (1994) proposed accounts of retrospective revaluation in which new learning about the target cue occurred on the revaluation trials despite the absence of the target cue on these trials. Accounts such as these last two view cue competition as an acquisition failure. Thus, basic retrospective revaluation phenomena fail to differentiate between acquisition-focused and performance-focused accounts of cue competition. Most recently, there have been

several reports that are problematic for each view point (e.g., Arcediano, Escobar, & Miller, 2004; Melchers, Lachnit, & Shanks, 2004; Urushihara, Stout, & Miller, 2004). Thus, it appears that neither view as they currently stand can adequately address all of the retrospective revaluation data.

In addition to retrospective revaluation, two other means of recovering responding to overshadowed and blocked cues have been identified. One of these is reminder treatments in which, prior to testing, a so-called reminder stimulus is presented (e.g., Balaz, Gutsin, Cacheiro, & Miller, 1982; Kasprow, Cacheiro, Balaz, & Miller, 1982). Reminder-induced recovery of responding is most readily accomplished by presenting the outcome (usually an unconditioned stimulus) to the subject a few times following training. Moreover, presentation of the training context alone or even the target cue itself a few times has also been found to be effective in some instances. In highly cognitive terms, one might say that the reminder stimulus acts as a potent retrieval cue, and once the target association is reactivated by this cue, the association is then re-stored in a more accessible location that allows easier retrieval on a subsequent test trial. However, such processes are not part of any contemporary formal associative model of learning. Another recovery technique is spontaneous recovery, that is, the insertion of a long retention interval before testing has been found to reduce the response deficit of cue competition (e.g., Kraemer, Lariviere, & Spear, 1988; Pineno, Urushihara, & Miller, 2005). Neither the acquisition-focused or performance-focused accounts of cue competition provide adequate explanations of why these two treatments produce recovery from cue competition. However, the observation that the absent behavior can reappear without further training appears to be more consistent with a performance-failure account of cue competition. Traditional associative models are challenged to provide complete explanations of these last two procedures for recovering responding after cue competition treatment.

Stimulus Interaction

Stimulus competition

Stimulus competition is often taken to mean cue competition. But beyond cue competition (i.e., competition between cues presented together), there is also published evidence of competition between outcomes presented together. For example, Esmoris-Arranz, Miller, and Matute (1997) and Miller and Matute (1998; also see Rescorla, 1980, for a variation on the same effect) reported that pairing a cue X (e.g., a tone) with a nontarget outcome (O1, e.g., a flashing light) followed by the tone being paired with a compound of the nontarget outcome and target outcome (O2, e.g., a white noise, that is X-[O1+O2] trials) attenuated responding to the tone after noise (O2) had been paired with a footshock,

relative to subjects that had not the X-O1 pairings. Control groups in these studies determined that simultaneous presentation of innocuous stimuli O1 and O2 did not result in one outcome distracting the subjects from the other. The O2-footshock pairings essentially rendered this preparation a form of sensory preconditioning, which served to avoid presenting two biologically significant outcomes (i.e., unconditioned stimuli) at the same time which likely would have resulted in one unconditioned stimulus distracting the subjects from the other. Contemporary acquisition-focused associative models provide no account of this effect, although it is not hard to imagine how models that posit limited attention (e.g., Sutherland & Mackintosh, 1971) might be modified to address the phenomenon in ways similar to how they account for cue competition. In contrast, Miller and Matzel's (1988) performance-focused comparator hypothesis provides a ready account of outcome competition (Miller & Escobar, 2002). Here the competing outcome (O1) is treated the same as a competing stimulus; that is, at the time of testing, the O1-O2 association (in conjunction with the X-O1 association) indirectly activates a representation of the target outcome (O2) upon the test trial presentation of X through an X_O1_O2 pathway. Then this indirectly activated representation of O2 competes with behavioral expression of the representation of O2 that is directly activated by X (i.e., the X-O2 association). Thus, this model accounts equally well for cue competition and outcome competition.

Stimulus interference

In addition to outcome competition between stimuli presented together during training, associative learning theories are challenged by an old literature, largely from the verbal learning tradition, concerning retroactive and proactive interference (see Slamecka, & Ceraso, 1960, for a review). Associative interference is commonly observed between items within a serial list (presumably represented by dyadic associations between elements) and between lists of items, when items are presented serially (as opposed to simultaneously, as in so-called competitive situations). For example, a simple instance is to train on a list of dyads including A-B (e.g., apple-chair) followed by training on another list of dyads including A-C (e.g., apple-shoe) and then test by presenting A (apple) to see if it is impaired in eliciting retrieval of B (chair), which were it to occur would be a form of retroactive outcome interference. Although the modern study of interference in Pavlovian situations began with humans (see Matute & Pineno, 1998), there are now many studies demonstrating interference effects with nonhuman subjects (e.g., Amundson, Escobar, & Miller, 2003, for proactive interference; Escobar, Matute, & Miller, 2001, for retroactive interference). Hence, interference does not require verbal abilities. Numerous studies have found that an important

requirement for associative interference is that the two associations must share a common element (e.g., Escobar et al., 2001). Let us use the notation of paired associate learning with X as the target cue, A as an alternative cue, O1 as the target outcome, and O2 as an alternative outcome (all four being tone, flashing light, white noise, click train, counterbalanced). Then retroactive cue interference is represented as X-O1 followed by A-O1 which impairs X in eliciting retrieval of O1 (and hence behavior appropriate for O1); retroactive outcome interference is represented as X-O1 followed by X-O2 which impairs X in eliciting retrieval of O1 (e.g., counterconditioning); proactive cue interference is represented A-O1 followed by X-O1 which impairs X in eliciting retrieval of O1; and proactive outcome interference is represented as X-O2 followed by X-O1 which impairs X in eliciting retrieval of O1. Schematically, interactions between stimuli can be represented in a 2 x 2 matrix (see Table 1) in which the interacting elements are either cues or outcomes and the interaction is either competition (interacting stimuli presented together) or interference (interacting stimuli presented apart). One should note that retroactive interference (based on either common cues or common outcomes) is an example of a behavioral deficit in which the target behavior was originally observed (or was at least observable). In contrast, proactive interference is an example of a behavioral deficit in which the target behavior was never evident unless there is some manipulation to counter the interference. Traditional acquisition-focused models of stimulus interaction have focused only on cell 1 of the matrix in Table 1. Miller and Escobar (2002) have shown how Miller and Matzel's (1988) comparator hypothesis address both cells 1 and 2, but not cells 3 and 4.

I previously spoke about interference effects when I was discussing shifts from recency to primacy. The point there was that these shifts were problematic for contemporary associative models. The point that I wish to make here is that the occurrence of interference between stimuli trained apart (provided that they have a common associate) is itself problematic for almost all contemporary associative models. Bouton (1993) has provided a viable model of associative outcome interference (cell 4 of Table 1). Importantly, his model is limited to outcome interference effects such as extinction, latent inhibition, and counterconditioning. It does not speak to associative cue interference. Nor does it speak to competition between stimuli (cues or outcomes) trained together any more than do the associative models that address cue competition (e.g., Pearce & Hall, 1980; Miller & Matzel, 1988; Rescorla & Wagner, 1972; Wagner, 1981) speak to interference.

Miller and Escobar (2002) suggested a generalization of Bouton's (1993) model that encompasses interference between cues (cell 3 of Table 1) as well as interference between outcomes (cell 4). Whereas Bouton spoke of interference arising when two associations to the same cue compete for retrieval when the cue is presented at test, Miller and Escobar proposed that two associations sharing one common associate compete for retrieval regardless of whether the common element is a cue (as is required by Bouton's model) or an outcome. The Miller and Escobar account of stimulus interference in conjunction with Miller and Matzel's (1988) comparator hypothesis provide a full account of both stimulus interference and stimulus competition. Moreover, this dual process approach does not posit that one process is engaged and the other disabled based on the procedure used. Rather, the interference mechanism depends on differential retrieval based on potentially interfering memories that have been acquired in different physical and/or temporal contexts. Thus, this mechanism, although always engaged, is effective primarily when the two associations have been acquired independently of one another and hence there is no within-compound association between the competing elements which would make the competition (i.e., comparator) mechanism effective. In contrast, the competition mechanism depends on the existence of a within-compound association between the two competing elements which would exist only if the two elements were trained in compound. Notably, training in compound would minimize differential retrieval of the two associations, thereby precluding interference. Thus, both processes are potentially active at the same time, but the conditions that maximize one effect are exactly those that minimize the other effect and vice versa. Thus, with any one experimental procedure, this hybrid model does not anticipate the simultaneous occurrence of both competition and interference.

Having one mechanism for interference and another mechanism for competition, however, appears unparsimonious, particularly when the two types of phenomena addressed (stimulus competition and interference) have so much in common. For example, both interference and competition appear to be reversible without further training with the target stimulus; this can be effected through massive extinction of the interfering or competing association (e.g., Amundson, et al., 2003; Kaufman & Bolles, 1981). Such similarities suggest that a single mechanism might be responsible for both interference and competition. Thus, Miller and Escobar's (2002) hybrid model is not entirely satisfying. But the central point here is not the adequacy of Miller and Escobar's model, but the failure of traditional acquisition-focused associative models of learning (e.g., Pearce & Hall, 1980, Pearce, 1987, Rescorla & Wagner, 1972, Wagner, 1981) to account for cells 2, 3, and 4 of the matrix in Table 1.

Applying Associative Theory to Complex Behavior

Although the focus of this paper has been on the chal-

Table 1: Stimulus Interactions

Procedure \ Competing stimuli	Cues	Outcomes
Stimuli trained together	Cell 1: Cue competition e.g., conventional blocking & overshadowing of cues	Cell 2: Outcome competition e.g., overshadowing & blocking of outcomes

Stimuli trained apart	Cell 3:
	Cue interference
	e.g., retroactive interference
	    X-O, A-O, test X-O
	& proactive interference
	    A-O, X-O, test X-O
		Cell 4:
		Outcome interference
		e.g., retroactive interference
		    X-O1, X-O2, test X-O1
		& proactive interference
		    X-O2, X-O1, test X-O1

lenges posed to contemporary associative theories as a result of their use of summary statistics and their overlooking both the learning-performance distinction and several types of stimulus interaction, I feel that I should briefly speak to how well do these theories account for more complex behavior. An early and continuing goal of associative models has been to provide the building blocks for models of complex behaviors such as categorization and language (e.g., Rumelhart, Hinton, & McClelland, 1986). In recent years, some of the most successful attempts to account for such complex behavior through associative principles have taken the form of connectionist models. These models, true to associative conventions, restrict themselves to variably weighted links between pairs of event representations, but deviate from traditional associative theories by assuming the existence of large numbers of interacting associations, a feature that presumably models the complexity of the brain. The successes of connectionistic models are impressive indeed (e.g., Kruschke, 1992; Love, Medin, & Gureckis, 2004), but they have their failings. One unusual failing is that the approach explains too much (Massaro, 1988). That is, these models are very flexible and can model almost any phenomena they confront, but they rarely make testable *a priori* predictions. This characteristic makes these models difficult to falsify (a feature also seen in some seemingly simple associative models). This weakness implies contemporary connectionist models tend to be under-constrained, something that may well lend itself to being corrected in the future. Researchers working with traditional (i.e., simple) associative models have generally shied away from connectionist modeling (but see Pearce, 1994). This likely reflects both some concern about these models not making testable predictions and their mathematics being a challenge to the researcher. Notably, connectionist theories, by the shear number of micro-associations that each event generates, effectively circumvent the assumption that only summary statistics are retained. Im-

portantly, connectionist models clearly demonstrate that the associative approach can in principle account for complex behavior, but the relative inability of present versions to predict new phenomena has been disappointing.

Conclusions

How well have contemporary associative theories fulfilled the purposes of models? To answer this question, we must first agree on the goals of a model of learning. Conventionally, these seem to be three-fold: 1) to direct research that leads to the discovery of new behavioral phenomena, 2) to theoretically organize behavioral phenomena including connecting with more molecular neural analysis and more molar cognitive analysis, and 3) to serve applied needs. Contemporary associative theories have been highly successful with respect to the first two goals and only modestly successful with respect to the third goal. Many new phenomena have been discovered in the course of testing different associative theories (e.g., superconditioning and overexpectation effects). Moreover, experiments focused on assessment of learning theories have provided the behavioral basis for a vast amount of illuminating neurophysiological research in recent years (for reviews, see e.g., Holland & Gallagher, 1999; Squire & Kandel, 1999; Thompson, Bao, Chen, Cipriano, Grethe, Kim, Thompson, Tracy, Weninger, & Krupa, 1997). However, this synergy here has been more between the behavioral phenomena and the neurological studies than between the associative theories per se and the neurological studies. With respect to the impact of modern associative theories on clinical psychology, the contributions have been modest. However, the more cognitive approach of newer associative theories, with their frequent references to expectations and event representations (e.g., Holland, 1990; Rescorla, 1988), has helped reconcile cognitive therapy with behavioral therapy. And at the more applied clinical level, the largest impact in recent years has seemingly come from

Bouton's (1993) model as implemented in exposure therapies (Collins & Brandon, 2002). However, although Bouton's model presumes binary associations, it does not say much about their nature or the necessary conditions for the formation or alteration of associations, thereby making it an associative theory only in the broadest sense.

What are the alternatives to associative models? Associative models have survived as long as they have, despite their several and continuing shortcomings, largely due to the lack of simple alternatives (which do not surrender degrees of freedom to a homunculus) and because their simplicity has stimulated much illuminating research. Contingency theories are often looked upon as possible alternatives to purely associative models (e.g., Cheng, 1997; White, 2003). But, as I have previously stated, contingency models implicitly use associations themselves and need revision to account for trial-order effects among other things. Other families of alternatives include inference theories (e.g., DeHouwer & Beckers, 2002; Lovibond, 2003); however, at this time these models are glaringly under specified. Another approach worthy of mention is the rate expectancy theory of Gallistel and Gibbon (2000). Even this model implicitly uses the basic associative principle of mental representations of events that are linked, but it adds a number of important novel features including episodic memories. Moreover, it is well specified. But it fails to deal with a number of phenomena such as stimulus intensity effects, and as previously mentioned makes truly heroic demands upon memory capacity and disavows the seeming importance of contiguity in acquisition.

In the broad sense of associations, there has been no plausible alternative offered to the construct of associations. Thus, the issue is not the acceptance of the construct of associations, but whether to accept some form of contemporary associative theory. Before we will be ready to do this, there are three central weaknesses of contemporary associative modeling that need to be addressed. First is the assumption that only summary statistics are retained. The fundamental assumption of all contemporary associative theories, that all that organisms encode are summary statistics, is contrary to overwhelming data. It is this assumption that makes associative models simple, and hence tractable. But it is at the cost of capturing the reality of mental life. The brain retains more than summary statistics. One viable alternative is a version of contingency theory that retains memories of instances including spatial locations as well as temporal intervals and trial-order information, and that uses higher-order decision making processes. The second weakness of contemporary associative theories is their failure to bridge the learning-performance distinction. And the third weakness is the need to account for the full family of stimulus interactions (i.e., the 2 x 2 matrix of Table 1). Future associative theories should address these phenomena.

Footnotes

[1]Here I am categorizing Pavlovian learning as a type of procedural learning, although some investigators have chosen to more narrowly differentiate between them.

References

Amundson, J.C., Escobar, M., & Miller, R.R. (2003). Proactive interference in first-order Pavlovian conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 29*, 311-322.

Anderson, J.R., & Bower, G.H. (1973). *Human associative memory.* Washington, DC: Winston.

Arcediano, F., Escobar, M., & Miller, R.R. (2004). Is stimulus competition an acquisition deficit or a performance deficit? *Psychonomic Bulletin & Review, 11*, 1105-1110.

Atkinson, R.C., & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In K.W. Spence & J.T. Spence (Eds.), *The psychology of learning and motivation, Vol. 2*. (pp. 89-195). New York: Academic Press.

Babb, S.J., & Crystal, J.D. (2005). Discrimination of what, when, and where: Implications for episodic-like memory in rats. *Learning and Motivation, 36*, 177-189.

Balaz, M.A., Gutsin, P., Cacheiro, H., & Miller, R.R. (1982). Blocking as a retrieval failure: Reactivation of associations to a blocked stimulus. *Quarterly Journal of Experimental Psychology, 34B*, 99-113.

Bjork, R.A. (2001). Recency and recovery in human memory. In H.L. Roediger III, J.S. Nairne, I. Neath, & A.M. Surprenant (Eds.), T*he nature of remembering: Essays in honor of Robert G. Crowder* (pp. 211-232). Washington, DC: American psychological Association.

Blaisdell, A.P., & Cook, R.G. (2004). Integration of spatial maps in pigeons. *Animal Cognition, 8*, 7-16.

Bouton, M.E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin, 114*, 80-99.

Bouton, M.E., & Bolles, R.C. (1979). Contextual control of the extinction of conditioned fear. *Learning and Motivation, 10*, 445-466.

Bouton, M.E., & Peck, C.A. (1992). Spontaneous recovery in cross-motivational transfer (counterconditioning). *Animal Learning & Behavior, 20*, 313-321.

Bouton, M.E., & Ricker, S.T. (1994). Renewal of extinguished responding in a second context. *Animal Learning & Behavior, 22*, 317-324.

Brown-Su, A.M., Matzel, L.D., Gordon, E.L., & Miller, R.R. (1986). Malleability of conditioned associations: Path dependence. J*ournal of Experimental Psychology: Animal Behavior Processes, 12*, 420-427.

Cheng, P.W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104*, 367-405.

Cheng, P.W., & Novick, L.R. (1992). Covariation in natural

causal induction. *Psychological Review, 99*, 365-382.

Church, R.M. (1989). Theories of timing behavior. In S.B. Klein & R.R. Mowrer (Eds.), *Contemporary learning theories: Instrumental conditioning theory and the impact of biological constraints on learning* (pp.41-71). Hillsdale, NJ: Erlbaum.

Clayton, N.S., & Dickinson, A. (1998). What, where, and when: Episodic-like memory during cache recovery by scrub jays. *Nature, 395*, 272-274.

Collins, B. N., & Brandon, T. H. (2002). Effects of extinction context and retrieval cues on alcohol cue reactivity among nonalcoholic drinkers. *Journal of Consulting & Clinical Psychology, 70*, 390-397.

Cook, R.G., Levinson, D.G., Gillett, S.R., & Blaisdell, A.P. (2005). Capacity and limits of associative memory in the pigeon. *Psychonomic Bulletin & Review, 12*, 350-358.

Crystal, J.D. (in press). Sensitivity to time: Implications for the representation of time. In E.A. Wasserman & T.R. Zentall (Eds.), *Comparative cognition: Experimental explorations of animal intelligence*. Oxford: Oxford University Press.

De Houwer, J., & Beckers, T. (2002). A review of recent developments in research and theories on human contingency learning. *Quarterly Journal of Experimental Psychology, 55B*, 289-310.

De Houwer, J., Beckers, T., & Glautier, S. (2002). Outcome and cue properties modulate blocking. *Quarterly Journal of Experimental Psychology, 55A*, 965-985.

De la Casa, L.G., & Lubow, R.E. (2000). Super-latent inhibition with delayed conditioned taste aversion testing. *Animal Learning & Behavior, 28*, 389-399.

De la Casa, L.G., & Lubow, R.E. (2002). An empirical analysis of the super-latent inhibition effect. *Animal Learning & Behavior, 30*, 112-120.

Dennis, M.J., & Ahn, W-K. (2001). Primacy in causal strength judgments: The effect of initial evidence for generative versus inhibitory relationships. *Memory & Cognition, 29*, 152-164.

Denniston, J.C., Miller, R.R., & Matute, H. (1996). Biological significance as a determinant of cue competition. *Psychological Science, 7,* 325-331.

Denniston, J.C., Savastano, H.I., Blaisdell, A.P., & Miller, R.R. (2003). Cue competition as a retrieval deficit. *Learning and Motivation, 34*, 1-31.

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). Hillsdale, NJ: Erlbaum.

Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgements. *Quarterly Journal of Experimental Psychology, 49B*, 60-80.

Eacott, M.J., Easton, A., & Zinkivskay, A. (2005). Recollection in an episodic-like memory task in the rat. *Learning and Memory, 12*, 221-223.

Escobar, M., Matute, H., & Miller, R.R. (2001). Cues trained apart compete for behavioral control in rats: Convergence with the associative interference literature. *Journal of Experimental Psychology: General, 130*, 97-115.

Esmoris-Arranz, F.J., Miller, R.R., & Matute, H. (1997). Blocking of subsequent and antecedent events. *Journal of Experimental Psychology: Animal Behavior Processes, 23*, 145-156.

Gallistel, R., & Gibbon, J. (2000). Time, rate and conditioning. *Psychological Review, 107*, 289-344.

Gibbon, J., & Balsam, P. (1981). Spreading association in time. In C.M. Locurto, H.S. Terrace, & J. Gibbon (Eds.), *Autoshaping and conditioning theory* (pp. 219-253). New York: Academic Press.

Grahame, N.J., Barnet, R.C., & Miller, R.R. (1992). Pavlovian inhibition cannot be obtained by posttraining A-US pairings: Further evidence for the empirical asymmetry of the comparator hypothesis. *Bulletin of the Psychonomic Society, 30*, 399-402.

Healy, S. (1998). *Spatial representations in animals*. Oxford, England: Oxford University Press.

Hogarth, R.M., & Einhorn, H.J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology, 24*, 1-55.

Holland, P.C. (1990). Event representation in Pavlovian conditioning: Image and action. *Cognition, 37*, 105-131.

Holland, P.C. (1999). Overshadowing and blocking as acquisition deficits: No recovery after extinction of overshadowing or blocking cues. *Quarterly Journal of Experimental Psychology, 52B*, 307-333.

Holland, P.C., & Gallagher, M. (1999). Amygdala circuitry in attentional and representational processes. *Trends in Cognitive Sciences, 3,* 65-73.

Kasprow, W.J., Cacheiro, H., Balaz, M.A., & Miller, R.R. (1982). Reminder-induced recovery of associations to an overshadowed stimulus. *Learning and Motivation, 13*, 155-166.

Kaufman, M.A., & Bolles, R.C. (1981). A nonassociative aspect of overshadowing. *Bulletin of the Psychonomic Society, 18*, 318-320.

Kraemer, P.J., Lariviere, N.A., & Spear, N.E. (1988). Expression of a taste aversion conditioned with an odor-taste compound: Overshadowing is relatively weak in weanlings and decreases over a retention interval in adults. *Animal Learning & Behavior, 16*, 164-168.

Kruschke, J.K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99*, 22-44.

Logan, G.D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*, 492-527.

Lopez, F.J., Shanks, D.R., Almaraz, J., & Fernandez, P. (1998). Effects of trial order on contingency judgments: A comparison of associative and probabilistic contrast accounts. Journal of Experimental Psychology: *Learning, Memory, and Cognition, 24*, 672-694.

Love, B.C., Medin, D.L., & Gureckis, T.M. (2004). SUSTAIN: A network model of human category learning. *Psychological Review, 111*, 309-332.

Lovibond, P.F. (2003). Causal beliefs and conditioned responses: Retrospective revaluation induced by experience and by instruction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 97-106.

Massaro, D.W. (1988). Some criticisms of connectionist models of human performance. *Journal of Memory and Language, 27*, 213-234.

Matute, H., & Pineno, O. (1998). Stimulus competition in the absence of compound conditioning. *Animal Learning & Behavior, 26*, 3-14.

Matzel, L.D., Held, F.P., & Miller, R.R. (1988). Information and expression of simultaneous and backward conditioning: Implications for contiguity theory. *Learning and Motivation, 19*, 317-344.

Matzel, L.D., Schachtman, T.R., & Miller, R.R. (1985). Recovery of an overshadowed association achieved by extinction of the overshadowing stimulus. *Learning and Motivation, 16*, 398-412.

Melchers, K.G., Lachnit, H., & Shanks, D.R. (2004). Within-compound associations in retrospective revaluation and in direct learning: A challenge for comparator theory. *Quarterly Journal of Experimental Psychology, 57B*, 25-54.

Miller, R.R., & Escobar, M. (2001). Contrasting acquisition-focused and performance-focused models of behavior change. *Current Directions in Psychological Science, 10*, 141-145.

Miller, R.R., & Escobar, M. (2002). Associative interference between cues and between outcomes presented together and presented apart: An integration. *Behavioural Processes, 57*, 163-185.

Miller, R.R., Hallam, S.C., & Grahame, N.J. (1990). Inflation of comparator stimuli following CS training. *Animal Learning & Behavior, 18*, 434-443.

Miller, R.R., & Matute, H. (1996). Biological significance in forward and backward blocking: Resolution of a discrepancy between animal conditioning and human causal judgment. *Journal of Experimental Psychology: General, 125*, 370-386.

Miller, R.R, & Matute, H. (1998). Competition between outcomes. *Psychological Science, 9*, 146-149.

Miller, R.R., & Matzel, L.D. (1987). Memory for associative history of a conditioned stimulus. *Learning and Motivation, 18*, 118-130.

Miller, R.R., & Matzel, L.D. (1988). The comparator hypothesis: A response rule for the expression of associations. In G.H. Bower (Ed.), *The psychology of learning and motivation*, Vol. 22 (pp. 51-92). San Diego, CA: Academic Press.

Miller, R.R., & Oberling, P. (1998). Analogies between occasion setting and Pavlovian conditioning. In N.A. Schmajuk & P.C. Holland (Eds.), *Occasion setting: Associative learning and cognition in animals* (pp. 3-35). Washington, DC: American Psychological Association.

Murdock, B.B., Jr. (1960). The distinctiveness of stimuli. *Psychological Review, 67*, 16-31,

Neath, I., & Knoedler, A.J. (1994). Distinctiveness and serial position effects in recognition and sentence processing. *Journal of Memory and Language, 33*, 776-795.

Pavlov, I.P. (1927). *Conditioned reflexes*. (G.V. Anrep, Ed. & Trans.) London: Oxford University Press.

Pearce, J.M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review, 94*, 61-73.

Pearce, J.M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review, 101*, 587-607.

Pearce, J.M., & Hall, G.(1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not unconditioned stimuli. *Psychological Review, 82*, 532-552.

Pineño, O., Denniston, J.C., Beckers, T., Matute, H., & Miller, R.R. (2005). Contrasting predictive and causal values of predictors and of causes. *Learning & Behavior, 33*, 184-196.

Pineño, O., Urushihara, K., & Miller, R.R. (2005). Spontaneous recovery from forward and backward blocking. *Journal of Experimental Psychology: Animal Behavior Processes, 31*, 172-183.

Postman, L., Stark, K., & Fraser, J. (1968). Temporal changes in interference. *Journal of Verbal Learning and Verbal Behavior, 7*, 672-694.

Reed, S.K. (1972). Pattern recognition and categorization. *Cognitive Psychology, 3*, 382-407.

Rescorla, R.A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology, 66*, 1-5.

Rescorla, R.A. (1970). Reduction in the effectiveness of reinforcement after prior excitatory conditioning. *Learning and Motivation, 1*, 372-381.

Rescorla, R.A. (1980). *Pavlovian second-order conditioning*. Hillsdale, NJ: Erlbaum.

Rescorla, R.A. (1988). Pavlovian conditioning: It's not what you thought it is. *American Psychologist, 43*, 151-160.

Rescorla, R.A. (1997). Spontaneous recovery after Pavlovian conditioning with multiple outcomes. *Animal Learning & Behavior, 25*, 99-107.

Rescorla, R.A. (2000). Extinction can be enhanced by a concurrent excitor. *Journal of Experimental Psychology:*

Animal Behavior Processes, 26, 251-260.

Rescorla, R.A., & Heth, C.D. (1975). Reinstatement of fear to an extinguished conditioned stimulus. *Journal of Experimental Psychology: Animal Behavior Processes, 1*, 88-96.

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

Rumelhart, D.E., Hinton, G.E., & McClelland, J.L. (1986). A general framework for parallel distributed processing. In J.L. McClellan, D.e. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition,* vol. 1: Foundations (pp. 45-76). Cambridge, MA: MIT Press.

Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology, 89*, 63-77.

Savastano, H.I., & Miller, R.R. (1998). Time as content in Pavlovian conditioning. *Behavioural Processes, 44*, 147-162.

Shevill, I., & Hall, G. (2004). Retrospective revaluation effects in the conditioned suppression procedure. *Quarterly Journal of Experimental Psychology, 57B*, 331-347.

Slamecka, N.J., & Ceraso, J. (1960). Retroactive and proactive inhibition of verbal learning. *Psychological Bulletin, 57*, 449-475.

Soltysik, S.S., Wolfe, G.E., Nicolas, T., Wilson, W.J., & Garcia-Sanchez L. (1983). Blocking of inhibitory conditioning within a serial conditioned stimulus-conditioned inhibitor compound: Maintenance of acquired behavior without an unconditioned stimulus. *Learning and Motivation, 14*, 1-29.

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory, 82*, 71-77.

Squire, L.R., & Kandel, E. R. (1999). *Memory: From mind to molecules*. W.H. Freeman & Co., New York. 1999.

Spear, N.E. (1973). Retrieval of memories in animals. *Psychological Review, 80*, 163-194.

Spellman, B. A. (1996). Acting as intuitive scientists: Contingency judgments are made while controlling for alternative potential causes. *Psychological Science, 7*, 337-342.

Stout, S.C., Amundson, J.C., & Miller, R.R. (in press). Trial order and retention interval in human predictive judgment. *Memory & Cognition.*

Sutherland, N.S., & Mackintosh, N.J. (1971). *Mechanisms of animal discrimination learning*. New York: Academic Press.

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

Thompson, R.F., Bao, S., Chen, L., Cipriano, D.D., Grethe,

J.S., Kim, J.J., Thompson, J.K., Tracy, J.A., Weninger, M.S., & Krupa, D.J. (1997). Associative learning. *International Review of Neurobiology, 41*, 151-189.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving and W. Donaldson (Eds.), *Organization of memory* (pp. 382-403). New York: Academic Press.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53*, 1-25.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessability of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5*, 381-391.

Tulving, E., & Thomson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352-73.

Urushihara, K., Wheeler, D.S., & Miller, R.R. (2004). Outcome pre- and post-exposure effects: Retention interval interacts with primacy and recency. *Journal of Experimental Psychology: Animal Behavior Processes, 30*, 283-298.

Urushihara, K., Stout, S.C., & Miller, R.R. (2004). The basic laws of conditioning differ for elemental cues and cues trained in compound. *Psychological Science, 15*, 268-271.

Van Hamme, L.J., & Wasserman, E.A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning and Motivation, 25*, 127-151.

Wagner, A.R. (1971). Elementary associations. In *H.H. Kendler & J.T. Spence (Eds.), Essays in neobehaviorism. A memorial volume to Kenneth W. Spence* (pp. 187-213). New York: Appleton-Century-Crofts.

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

Weidemann, G., & Kehoe, E.J. (2004). Recovery of the rabbit's conditioned nictitating membrane response without direct reinforcement after extinction. *Learning & Behavior, 32*, 409-426.

Wheeler, D.S., Stout, S.C., & Miller, R.R. (2004). Interaction of retention interval with CS-preexposure and extinction effects: Symmetry with respect to primacy. *Learning & Behavior, 32*, 335-347.

White, P.A. (2003). Causal judgement as the evaluation of evidence: The use of confirmatory and disconfirmatory information. *Quarterly Journal of Experimental Psychology, 56A*, 491-513.

Wooldridge, D.E. (1963). *The machinery of the brain*. New York: McGraw-Hill.

Wright, A.A. (1998). Auditory and visual serial position functions obey different laws. *Psychonomic Bulletin & Review, 5*, 564-584.

Wright, A.A., Cook, R.G., Rivera, J.J., Shyan, M.R., Nei-

worth, J.J., & Jitsumori, M. (1990). Naming, rehearsal, and interstimulus interval effects in memory processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 1043-1059.

Zentall, T.R. (2005). Animals may not be stuck in time. *Learning and Motivation, 36*, 208-225.