WHAT’S ELEMENTARY ABOUT ASSOCIATIVE LEARNING?

Edward A. Wasserman
Department of Psychology, University of Iowa, Iowa City, Iowa 52242-1407

Ralph R. Miller
Department of Psychology, Binghamton University, Binghamton, New York 13902-6000

KEY WORDS: Pavlovian conditioning, Thorndikean conditioning, association formation, learned behavior, conditioned responding

ABSTRACT

The scientific study of associative learning began nearly 100 years ago with the pioneering studies of Thorndike and Pavlov, and it continues today as an active area of research and theory. Associative learning should be the foundation for our understanding of other forms of behavior and cognition in human and nonhuman animals. The laws of associative learning are complex, and many modern theorists posit the involvement of attention, memory, and information processing in such basic conditioning phenomena as overshadowing and blocking, and the effects of stimulus preexposure on later conditioning. An unresolved problem for learning theory is distinguishing the formation of associations from their behavioral expression. This and other problems will occupy future generations of behavioral scientists interested in the experimental investigation of associative learning. Neuroscientists and cognitive scientists will both contribute to and benefit from that effort in the next 100 years of inquiry.

CONTENTS

INTRODUCTION AND BROAD OVERVIEW...................................................... 574
What Is the Domain of Elementary Associative Learning? ......................................... 574
What Is the History of Associative Learning?.......................................................... 575
What Is the Current State of the Field? ................................................................ 575
What Is the Promise of the Field? ......................................................................... 575
INTRODUCTION AND BROAD OVERVIEW

Nearly every student who has taken an introductory psychology course has heard of Pavlov and Thorndike and learned a bit about their pioneering studies of classical and instrumental conditioning. However, students’ textbooks seldom inform them that associative learning is key to understanding both human and animal behavior or that the field is an active and controversial one. Unresolved issues have been revisited and new phenomena discovered for experimental study and theoretical explication. In addition, contributors from neuroscience and cognitive science are making the study of associative learning interdisciplinary.

We hope to convey the challenge and vitality of the field by reviewing the literature of the past dozen years. The last two Annual Review of Psychology articles on associative learning appeared in 1978 (Dickinson & Mackintosh) and 1982 (Rescorla & Holland). Because these articles were narrow in focus, we also cover some earlier research.

What Is the Domain of Elementary Associative Learning?

As the final notes fade at the close of the symphony’s tranquil third movement, a listener’s body tenses as she eagerly anticipates the boisterous opening of the final movement. When he nears the breakfast table, a small boy pushes aside the chair that stands between his seat and the television, thereby permitting him an unobstructed view of his favorite cartoon show. These examples illustrate the operation of associative learning, in which one event comes to be linked with another because of an individual’s past experiences.

The first example, in which the associated events do not include the individual’s behavior illustrates Pavlovian, classical, or respondent conditioning, whereas the second, in which the individual’s behavior is one of the associated events, is one of Thorndikean, instrumental, or operant conditioning. Although the events in Pavlovian and Thorndikean conditioning are decidedly different, the rules that govern these cases of learning are strikingly similar (Dickinson 1980). Largely for convenience, in this review we focus more heavily on Pavlovian conditioning.
What is the History of Associative Learning?

Associative learning has been appreciated since the time of the ancient Greeks and has long been the subject of philosophical speculation (Warren 1921). Yet it was not until the end of the nineteenth century that a true experimental science of associative learning was born. In the United States, Thorndike (1911) arranged a response-outcome contingency that permitted an animal to receive a reward when it performed a prescribed or instrumental behavior. That behavior increased in likelihood when the animal was again placed in the experimental setting. In Russia, Pavlov (1927, 1928) paired two stimuli with each other independently of the organism’s behavior. After repeated pairings, conditioned stimuli (CSs) came to elicit conditioned responses (CRs) similar to the unconditioned responses (URs) that had initially been elicited only by unconditioned stimuli (USs).

Since those discoveries, scientists around the world have been assiduously studying the principles of instrumental and classical conditioning, primarily in nonhuman animals. This work peaked in the 1930s, 1940s, and 1950s, when the research and theories of Hull, Skinner, Spence, and Tolman dominated experimental psychology (Bower & Hilgard 1981). Since the 1960s, the study of associative learning has declined in popularity, replaced by research in other aspects of cognition such as attention, memory, and information processing in human beings and animals (Wasserman 1993).

What Is the Current State of the Field?

Although no longer dominant in experimental psychology, associative learning remains a highly active area of research and theory. Work here still focuses on the behavior of animals, but the study of associative learning in human beings is growing in interest and importance (Shanks 1994). Most noteworthy is recent research in causal perception and the success that has been achieved there by associative theories (Allan 1993, Young 1995). Researchers in neuroscience and cognitive science have also become interested in associative learning.

What Is the Promise of the Field?

This objective approach to the study of associative learning has ambitious goals. To Pavlov, “along this path will be found the final triumph of the human mind over its uttermost and supreme problem—the knowledge of the mechanism and laws of human nature” (1928, p. 41). Thorndike further hoped to understand the origin and development of the human mental faculty because “out of…associative processes have arisen human consciousnesses with their sciences and arts and religions” (1911, p. 22). Far from being merely the laboratory study of learning in animals, the field still hopes to elucidate human and animal cognition by detailing the laws of associative learning.
Since the work of Thorndike and Pavlov, many research phenomena have been discovered and investigated. Those phenomena are central to the empirical corpus of the field and have inspired many theories of associative learning. Here we review some of those key phenomena.

**Simple Acquisition with a Single Cue or Response**

**CONTIGUITY AND SIMILARITY** Temporal contiguity between events has been one of the most important and well-researched factors in associative learning. In Pavlovian conditioning, temporal contiguity is defined by the CS-US interval; in Thorndikean conditioning, temporal contiguity is defined by the response-outcome delay.

Most authors who have surveyed the effects of the interstimulus interval (ISI) in Pavlovian conditioning have concluded that simultaneity of CS and US is not optimal for promoting CRs; rather, there is an advantage to the CS preceding the US by a relatively short time, with that advantage being a nonmonotonic function of the CS-US interval (Mackintosh 1983, Rescorla 1988; cf Matzel et al 1988c). The best ISI for promoting CRs is not a fixed point in time; it differs with the response, the intensity and nature of CS and US, and the temporal distribution of CS-US pairings (Atkins et al 1994, Berk & Miller 1978, Lennartz & Weinberger 1992, Schneiderman 1972). Moreover, lengthening the intertrial interval (ITI) raises the optimal ISI (Gibbon et al 1977, Levinthal et al 1985), which suggests that the key variable is the ISI/ITI ratio.

The customary interpretation of ISI effects is that the CS-US interval directly affects the acquisition of a CS-US association. However, other variables may participate. For example, limited prior training at ineffective ISIs nonetheless facilitates later CRs at an effective ISI; indeed, this facilitation is as great after pretraining at many ineffective ISIs as it is after pretraining at effective ones (Ross & Scavio 1983). In addition, reducing the number of trials per session increases the efficacy of longer ISIs (Kehoe et al 1991). These and other results are best interpreted by admitting performance variables into the behavioral equation (DeCola & Fanselow 1995, Yin et al 1994).

The low level of CRs that is often seen after simultaneous presentations of CS and US is usually attributed to a failure of acquisition. However, Rescorla (1980) argued that simultaneous events actually yield maximal learning, unless the US distracts the subject from the CS. He showed that robust acquisition occurs when the simultaneously paired events are not biologically significant and hence do not distract from one another. Suggestive of an alternative explanation, Matzel et al (1988c) and Barnet et al (1991) found that the CR
deficit is at least in part due to a failure of the organism to express an association that it had actually acquired.

In Thorndikean conditioning, increasing the delay between response and outcome reduces the efficacy of a reinforcer (for a review, see Tarpy & Sawabini 1974; for data, see Dickinson et al 1992, Wilkenfield et al 1992) or a punisher (for a review, see Azrin & Holz 1966; for data, see Baron et al 1969, Cohen 1968) in controlling learned behavior.

Spatial contiguity and interevent similarity also affect associative learning and performance. Locating CSs and USs spatially near to one another aids Pavlovian conditioning (Ellins & von Kluge 1990, Testa 1975; see also Marshall et al 1979, Rescorla & Cunningham 1979). So, too, does using CSs and USs with similar attributes (Testa 1975; see also Rescorla & Furrow 1977, Rescorla & Gillan 1980). Thorndikean discrimination learning is aided by the instrumental response spatially coinciding with the discriminative stimulus (Harrison et al 1977, Ramey & Goulet 1971; cf Rumbaugh et al 1989) and by the discriminative and reinforcing stimuli sharing similar patterns of stimulation (for a review, see JD Miller & Bowe 1982; for data, see Bowe et al 1987; cf Neill & Harrison 1987).

**NECESSITY AND SUFFICIENCY** Within the Pavlovian paradigm, sufficiency refers to the likelihood of the US occurring given the presence of the CS and can be expressed in probabilistic notation as P(US|CS), which depends on the frequencies with which the CS and US occur together and the CS occurs alone. Consistent reinforcement constitutes a sufficiency of 1.0, and extinction treatment constitutes a sufficiency of 0.0. Trials on which both the CS and US occur enhance CRs to the CS, whereas trials on which only the CS occurs impair CRs (Van Hamme & Wasserman 1994, Young 1995). Necessity refers to the US not occurring in the absence of the CS and can be expressed as 1 – P(US|No CS), which depends on the frequencies with which the US occurs alone and times when neither the US nor the CS occurs. Occasions on which only the US occurs weaken CRs, whereas occasions lacking both the CS and the US strengthen CRs, as when the ITI is lengthened (Young 1995). Within the Thorndikean paradigm, similar relations hold (Wasserman et al 1993). The concepts of sufficiency and necessity are jointly captured by the notion of contingency (discussed below).

**Complex Phenomena with a Single Cue or Response**

**EXPERIMENTAL EXTINCTION** Operationally, experimental extinction refers to the occurrence of the antecedent event (CS or instrumental response) without the reinforcer that followed it during acquisition. Empirically, extinction is the decrease in CRs that occurs because of this treatment. To explain extinction, Pavlov (1927) posited an inhibitory CS-US association that counteracted the excitatory association that was formed during acquisition; extinction was thus
a form of further learning rather than an eradication of earlier learning. Subsequently, many other explanations of extinction were proposed (for a review, see Kimble 1961), which were similar only in that they (a) rarely subscribed to unlearning and (b) could explain why inconsistent reinforcement during training generally yields greater resistance to extinction than does consistent reinforcement.

More recently, the widely cited model of Rescorla & Wagner (1972) has treated extinction as unlearning. Rescorla & Wagner made this proposal to retain the simplicity of their model, but the earlier reasons for rejecting unlearning are still valid. For instance, presentation of a strong novel stimulus before presentation of an extinguished CS temporarily restores CRs to the extinguished CS ("external inhibition"), and extinguished CRs frequently undergo partial "spontaneous" recovery with increasing retention intervals (Pavlov 1927). Contemporary research suggests that an extinguished CS has all the properties of a conditioned inhibitor (Calton et al 1996, Hart et al 1995), which brings us back almost to Pavlov’s view of extinction.

A notable addition to our knowledge about extinction is that it is context specific. Bouton and his colleagues have shown that when training occurs in one context and extinction occurs in another, CRs are still obtained if testing occurs in the first context (Bouton & Bolles 1979, Bouton & King 1983). If testing occurs in a third context, then CRs are also observed. These results suggest that, rather than responding being specific to the acquisition context, nonresponding is specific to the extinction context. Bouton (1993) and Brooks & Bouton (1994) interpreted these results in terms of the context serving as a salient retrieval cue, with the context used for acquisition training having a broader stimulus generalization gradient than the context used for extinction training. Bouton has supported this account by showing that even when acquisition and extinction occur in the same context, so that there are no CRs in that context at the end of extinction, CRs are again observed when testing occurs in another context (Bouton & Ricker 1994). This modulatory role of the context underscores the increasing attention paid to the contextual control of acquired behavior (Balsam & Tomie 1985, Swartzentruber 1995).

US-PREEXPOSURE EFFECT Repeated exposure to a reinforcer may retard the acquisition of behavioral control during later training (Randich & LoLordo 1979). Although nonassociative habituation may reduce the efficacy of the US, the prime factor responsible for the US-preexposure effect appears to be the excitatory status of the training context. During US preexposure, a context-US association may be formed that blocks the CS-US association during later training. Consistent with this interpretation, either extinction of the context or a change in context between US-preexposure and CS-US training attenuates the effect (Randich 1981). Matzel et al (1988b) further found that a conditioned
release of endorphins elicited by the context at the time of CS-US training is necessary to obtain the US-preexposure effect with aversive USs.

Although the US-preexposure effect is commonly thought to reflect a failure for the subject to acquire a CS-US association, Matzel et al (1987) found that extinguishing the training context after CS-US pairings attenuated the effect. Moreover, JS Miller et al (1993) have observed spontaneous recovery from the deficit. Thus, the US-preexposure effect is not entirely due to a deficit in acquisition. If blocking by the context is responsible for the effect, then it is blocking of expression rather than acquisition of the CS-US association.

CS-PREEXPOSURE EFFECT When subjects are extensively exposed to a CS before CS-US pairings, acquisition of behavioral control by the CS is retarded (Lubow & Moore 1959). This effect is called “latent inhibition”; however, this is a misnomer because a preexposed CS does not pass the summation test for conditioned inhibition (Reiss & Wagner 1972, Rescorla 1971). The most popular account of the CS-preexposure effect is that, during preexposure, subjects reduce their perceptual attention to the CS (Lubow et al 1976). Implicit here is the assumption that attention is necessary for acquisition. However, despite its popularity, this type of attentional theory has many problems. For example, it cannot easily explain the retardation of behavioral control found when a CS is paired with a strong US after the CS has been paired with a weaker version of the same US (Hall & Pearce 1979). Because pairings of the CS with the weak US result in behavioral control by the CS, it is unlikely that subjects have decreased their perceptual attention to the CS and hence that this decrease is the basis for the later retardation in learning when the CS is paired with the strong US. This Hall-Pearce negative transfer effect is also a problem for some attentional models that focus on associability rather than on perceptual attention (Mackintosh 1975), but not for others (Pearce & Hall 1980).

The main competitors to attentional explanations are models that focus on context-CS associations that are formed during CS preexposure. Wagner (1981) posited a context-CS association that interferes with acquisition of the CS-US association. Miller & Matzel (1988) and Grahame et al (1994) also focus on the context-CS association, but they posit that this association impairs the expression of the CS-US association. Consistent with both views, a context shift between CS preexposure and CS-US pairings (Hall & Minor 1984) or extinction of the context between CS-preexposure and CS-US pairings (Baker & Mercier 1982, Wagner 1979; cf Hall & Minor 1984) attenuates the effect. Distinguishing between these views, Grahame et al (1994) found that extinction of the context after CS-US pairings restores CRs to the CS, Kraemer et al (1991) found spontaneous recovery from the CS-preexposure deficit, and Kasprów et al (1984) found that a posttraining “reminder” can restore CRs. If acquisition were impaired, as Wagner’s account suggests, then no posttraining
treatment other than retraining should restore CRs to the CS. Thus, the CS-preexposure effect appears to be largely a failure to express, not to acquire information. If one adheres to a perceptual-attentional view, then this evidence implies that attention during training may not be necessary for acquisition, although it may well facilitate it; rather, attention during training may facilitate later expression of the acquired information, perhaps by creating more effective links to modulate information retrieval.

Operationally, CS preexposure and habituation are identical. Both effects involve the CS losing its potential to control behavior. Thus, one might suspect that they arise from a common process. However, habituation wanes rapidly with increases in the retention interval (Thompson & Spencer 1966), whereas the CS-preexposure effect does not (Hall & Schachtman 1987). In addition, habituation survives a context shift (Marlin & Miller 1981), whereas CS-preexposure does not (Hall & Honey 1989). Thus, different processes evidently underlie the CS-preexposure effect and habituation.

LEARNED irrelevance Random presentations of CSs and USs proactively interfere with associative learning. Mackintosh (1973) found that randomly presented CSs and USs produced more proactive interference than did either CSs alone or USs alone, a result he took to mean that animals can learn the irrelevance of the CS for predicting the US. Mackintosh argued that “learned irrelevance” makes it more difficult for the animal later to learn either an excitatory or an inhibitory interstimulus association (Baker 1977, Baker & Mackintosh 1977).

There is an obvious rival account of the learned irrelevance effect: namely, that it is simply the sum of the effects of CS alone and US alone preexposure. Early efforts to distinguish learned irrelevance from the summation of CS- and US-preexposure effects (Baker & Mackintosh 1979, Matzel et al. 1988d, Tomie et al. 1980) supported the existence of learned irrelevance by showing that manipulations that eliminate the CS- and US-preexposure effects do not eliminate the learned irrelevance effect. Providing more support, Bennett et al (1995) found that daily intermixtures of CSs and USs more adversely affected later conditioning than did giving the same number of CSs and USs in different blocks of sessions. However, Bonardi & Hall (1996) have shown that this effect depends on USs serving as part of the conditioning context. Thus, the reality of learned irrelevance is still uncertain.

**Complex Phenomena with Multiple Cues or Responses**

SENSORY PRECONDITIONING A common criticism of associative learning is that learning may only occur when biologically significant stimuli are the reinforcers. Yet learning often takes place without the obvious involvement of biologically significant stimuli. If this criticism were valid, then very real limits
would apply to the generality of associative principles. There is strong evidence that associative learning does occur in the absence of biologically significant stimuli. The best known case is “sensory preconditioning,” an effect first named by Brogden (1939) but first studied in 1932 by Podkopaev in Pavlov’s laboratory (Kimmel 1977). Initially, Podkopaev paired a whistle with a light (CSB-CSA). Next he paired the light with footshock (CSA-US). Finally, he tested the dog with the whistle. Despite the fact that the whistle had never been paired with footshock, it elicited leg flexion. When appropriate controls are included, this procedure provides compelling evidence of learning with biologically insignificant stimuli.

Later research has attested to the robustness of sensory preconditioning and has explored several variables that influence it (Razran 1971, Thompson 1972). Many of the variables that affect associative learning when the outcome is biologically significant affect sensory preconditioning in a similar way, which leads some authors to conclude that the same associative mechanisms are at work in both cases. Responding to CSB is an increasing function of the number of CSB-CSA pairings (Prewitt 1967), the percentage of trials on which CSB is followed by CSA (Tait et al 1971, 1972), and the intensity of CSB (Rogers 1973, Tait & Suboski 1972). Responding to CSB is reduced by presenting both CSB and CSA alone before CSB-CSA pairings (Parks 1968), by increasing the time between CSB and CSA (Lavin 1976, Spiker & Ferraro 1977), by presenting CSB alone after CSB-CSA pairings (Tait et al 1969), and by presenting CSA alone after CSA-US pairings (Rizley & Rescorla 1972). All the above research employed Pavlovian contingencies. An analogous case of “response preconditioning” with Thorndikean contingencies was reported by St. Claire-Smith & MacLaren (1983).

Sensory preconditioning has historically involved a three-step paradigm for investigating the association of biologically insignificant stimuli. Other research in Pavlov’s laboratory by Narbutovich (Kimmel 1977) entailed only a single training period. Here, a scuffling sound at one spatial site regularly preceded the spinning of a disk at another. After these pairings, the dog not only oriented toward the scuffling sound, it also turned toward the site where the disk was about to spin. The dog’s receptor orienting acts provided direct evidence of the association between the auditory and visual stimuli; no additional training was needed to prove the point. Another way to directly assay the association between neutral stimuli is to record electrical brain activity during CSB and CSA. When a visual stimulus, but not an auditory stimulus, is presented, alpha activity is blocked. After tone-light pairings, human beings show alpha blocking to the tone, which suggests the formation of a tone-light association (Smith & Putney 1979).
HIGHER-ORDER CONDITIONING  Another well-known case in which associative learning occurs in the absence of stimuli of inherent biological significance is “higher-order” conditioning, discovered by Pavlov. Here, the same phases of training are administered as in sensory preconditioning, but in the reverse order. Initially, CSA is paired with a biologically significant US (CSA-US). Then, CSB is paired with CSA (CSB-CSA). The result is that CSB now produces a CR, even though it was never paired with the US. Control groups prove that both phases of training are necessary for the resultant CR to CSB (Rescorla 1973).

Second-order conditioning has been found with numerous responses in many species, including proboscis extension in honeybees (Bitterman et al 1983), target striking in goldfish (Amiro & Bitterman 1980), sexual activity in quail (Crawford & Domjan 1995), and keypecking in chicks (Deich et al, in Wasserman 1981). Stronger responding to CSB has been reported when CSA was paired with larger USs than with smaller ones (O’Connell & Rashotte 1982) and when CSB was consistently followed by CSA than when it was inconsistently followed by CSA (Colwill & Rescorla 1985, Rashotte et al 1981). In addition, increasing the CSB-CSA interval lowers responding to CSB (Gibbs et al 1991; cf Popik et al 1979). Second-order conditioning is even demonstrable with a single CSA-US pairing followed by a single CSB-CSA pairing (Bond & DiGuisto 1976).

Much research in second-order conditioning sprang from Rizley & Rescorla’s (1972) finding that performance of the second-order CR might be independent of the ability of the first-order CS to elicit a CR. So, after CSA-US and CSB-CSA training, CSA alone presentations might weaken the CR to CSA, but not affect the CR to CSB (cf Rizley & Rescorla’s 1972 results of the same manipulation on sensory preconditioning). Insensitivity of CSA extinction on CSB responding was found by later workers (e.g. Barnet et al 1991, Robertson et al 1984), but that result has not been general. Extinction of CSA can produce reliable drops in responding to CSB (e.g. Colwill & Rescorla 1985, Hittesdorf & Richards 1982). Efforts to reconcile these conflicting findings according to the species of organism and the nature of the CSs and USs (Nairne & Rescorla 1981) were not successful. Had it been generally true that extinction of responding to CSA left responding to CSB intact, one could have concluded that second-order associations never depend on a representation of CSA. Moreover, Rescorla (1973) devalued the US representation without impairing second-order conditioning, which argues against the US representation being essential to second-order conditioning. Evidence against reinforcer-specific second-order associations also comes from Stanhope (1992), who successfully trained topographically different CRs to two different CSAs: one paired with water and the other with food. Topographically different CRs were
not formed, however, to two different CSBs, each of which was specifically paired with previously food- or water-paired CSAs.

Another glimpse into the associations involved in higher-order conditioning was provided by Stefurak et al (1990). Rats exhibit a preference for places paired with morphine, but an aversion for tastes paired with morphine. However, if a place CSB is paired with a taste CSA, then rats exhibit a place aversion, and if a taste CSB is paired with a place CSA, then rats exhibit a taste preference. Clearly, the first-order stimulus conditioned to morphine represents in memory only the single motivational effect of morphine to which it is preferentially associated.

SERIAL CONDITIONING Another situation in which associations may be formed between events lacking obvious biological significance is serial conditioning. In Pavlovian conditioning, serial conditioning might involve a succession of stimuli before reinforcement: CSB-CSA-US. In addition to the individual associations between CSB and the US and between CSA and the US, CSB and CSA may be associated with each other. Although this proposal is reasonable enough, empirical documentation has been slow in coming. One tactic has been to compare responding to CSB after serial conditioning with that obtained when the same two stimuli have been separately presented the same number of times in the same individual temporal relations to reinforcement. Should responding to CSB in the former case exceed that in the latter, it would be tempting to conclude that the superiority is due to the association between CSB and CSA. Just such a difference in responding has been reported (Schreurs et al 1993), but it has not been easy to eliminate other explanations for the result; the mere presence of CSA may help to “bridge” the temporal delay between CSB and the US or to “mark” the occurrence of CSB making it more distinctive, quite apart from any direct association between CSB and CSA (e.g. Hall & Honey 1993, Lieberman et al 1985, Rescorla 1982).

A different, more direct tactic was used by Wasserman et al (1978). In an autoshaping situation, pigeons were given two serial keylight compounds that were each followed by food: CSB1-CSA1-US and CSB2-CSA2-US. CSB1 involved two response keys lit with one color, and CSB2 involved the same two keys lit with a different color. CSA1 entailed only the left key lighted with a third stimulus, and CSA2 entailed only the right key lighted with the same third stimulus. The pigeons pecked the left key during CSB1 and the right key during CSB2, consistent with their forming highly specific associations between the CSBs and the CSAs.

Later work has used matching-to-sample procedures. The discriminative response in those Thorndikean tasks depended on the pigeon discriminating and remembering the temporal order in which two earlier key colors (C1 and C2) had occurred. Different reinforcement contingencies were in place after
C1-C2 and C2-C1 sequences. To control for the first or the second stimulus alone cueing the discriminative response, C1-C1 and C2-C2 sequences were also given. Pigeons showed clear control by the temporal order of two biologically insignificant stimuli (MacDonald 1993, Wasserman et al 1984, Weisman et al 1980). This ability to discriminate and remember the temporal order of two nonidentical stimuli implies a far richer representation of stimulus events than is usually granted to animals, and it lends support to the idea that the temporal relationship between events is encoded as part of an interevent association (Matzel et al 1988c).

COMPOUND STIMULUS CONTROL  Is the whole more than the sum of its parts? This question was central to Gestalt psychology, a school that more strongly stressed perceptual than associative principles. Still, there has been much interest in this question in the field of associative learning.

One situation that has been extensively explored involves simultaneous compound stimuli that are paired with reinforcement without the presentation of the compound’s elemental stimuli. After training, the subject is given the compound and its elements to see if the former more effectively produces behavior than the latter; if so, then there is reason to suggest that “configural” conditioning has occurred. The evidence might be even stronger if increases in the amount of compound stimulus training led to increases in compound responding and to decreases in element responding. Most theories of associative learning do not expect the compound to gain behavioral control at the expense of its elements, but this result might nonetheless occur if it takes a long time for the individual stimulus elements to “fuse” into an effective compound (Razran 1971). Greater responding to the compound than to its elements has been reported in both Pavlovian (e.g. Bellingham & Gillette 1981a, Kamin & Idrobo 1978) and Thorndikean (e.g. Bellingham & Gillette 1981b, Rescorla 1972) conditioning. However, in none of these reports were increases in compound training clearly accompanied by increases in compound responding and decreases in element responding. To observe this result seems to require the “positive patterning” procedure, which provides reinforcement on compound trials and withholds it on element trials: AB+, A−, B−. Pavlov (1927) was the first to investigate the effects of such training; he found greater responding to the compound stimulus than to its components.

Later research has confirmed the efficacy of the positive patterning procedure. TW Baker (1969) and Wickens et al (1970) not only found more responding to the compound than to its elements, they also found increases in compound responding and decreases in element responding with increases in discrimination training. Wickens et al (1970) further reported that the compound-element discrimination was a direct function of the CS-US interval, perhaps because the increasing duration of the compound stimulus and its
elements that accompanied increases in the ISI helped to fuse the elements into a configured stimulus. Consistent with this interpretation, Kehoe & Schreurs (1986) replicated the ISI result and also showed that, with the CS-US interval held constant, the compound-element discrimination was a direct function of CS duration.

Although positive patterning is often interpreted as indicating configuring, the result can also be explained by the conditioning of the individual elements and the summation of their response strengths when the elements are jointly presented. Consequently, most theorists argue that the best case for configural conditioning comes from the “negative patterning” procedure, which provides reinforcement on element trials and withholds it on compound trials: AB−, A+, B+. Robust responding on element trials but not on compound trials cannot be explained by summation. Here, too, Pavlov (1927) was the first to investigate negative patterning; he found less responding to the compound stimulus than to its components.

Later research has confirmed the efficacy of the negative patterning procedure (e.g. Redhead & Pearce 1995b, Rescorla et al 1985). Whitlow & Wagner (1972) further found that less responding to AB than to A or B cannot be attributed to some general properties of compound stimuli, like their greater intensity, numerosity, or complexity. Related results were reported by Rescorla (1972; see also Davidson et al 1993), who, along with Whitlow & Wagner, argued that the simultaneous presentation of two stimuli might involve three functional stimuli: A, B, and a unique configural cue AB. This configural cue presumably provides the critical stimulus allowing the organism to respond discriminatively in positive and negative patterning situations. The slow speed of learning positive and negative patterning discriminations suggests that the configural cue is of low salience. Hypothesizing a unique configural cue also helps us understand how biconditional discriminations such as AC+, BD+, AD−, BC− are learned (Saavedra 1975). Control by the elements would be expected to decrease because of their equivalent association with reinforcement and nonreinforcement, to be replaced by control by the unique configural cues associated with reinforcement (AC and BD) and nonreinforcement (AD and BC).

Other details of negative patterning are that discrimination learning is a direct function of the percentage of compound trials (Bellingham et al 1985, Kehoe & Graham 1988) and the CS-US interval (Kehoe & Graham 1988). In addition, prior positive patterning impairs negative patterning but not vice versa (Bellingham et al 1985). Furthermore, adding a common cue to both compound and element trials slows the speed of negative patterning (Pearce & Redhead 1993, Rescorla 1972; also see Redhead & Pearce 1995a), presumably because this addition makes the task more difficult. Finally, and controver-
sially, given the animal data attesting to the specificity of configural stimulus control (Davidson et al. 1993, Rescorla 1972, Whitlow & Wagner 1972), human beings may learn a general compound stimulus rule on both negative and positive patterning discriminations. They may respond less to the novel CD compound than to the novel elements C and D after AB−, A+, B+ training, whereas they may respond more to the novel CD compound than to the novel elements C and D after AB+, A−, B− training (Kleinschmidt & Lachnit 1993, Lachnit & Kimmel 1993).

CUE COMPETITION  
Cue competition refers to a retardation in behavioral control by a target CS due to its being trained in the presence of another CS that is more salient or a better predictor of the US. Examples include overshadowing (Pavlov 1927), blocking (Kamin 1968), and the relative validity effect (Wagner et al. 1968), each of which occurs in both Thorndikean and Pavlovian conditioning (Pearce & Hall 1978, Williams 1975). Overshadowing (Mackintosh & Reese 1979) and blocking (Balaz et al. 1982b) can be obtained with a single compound trial, but each effect is more robust with several compound trials. However, overshadowing (Billingham & Gillette 1981a) and blocking (Azorlosa & Cicala 1986) can be eliminated by a sufficiently large number of compound trials. The latter findings are problematic for many contemporary theories of learning, as are the results of several posttraining manipulations that yield a partial or complete restoration of responding to target CSs after cue competition has occurred, including (a) extinction of the successfully competing CS (Cole et al. 1995, Kaufman & Bolles 1981, Matzel et al. 1985), (b) “reminder” treatments with a component of training like the CS or US (Balaz 1982a, Cole et al 1996, Kasprow et al. 1982), and (c) spontaneous recovery (Kraemer et al. 1988, JS Miller et al. 1990). This reversibility without further CS-US training suggests that cue competition is due, at least in part, to a failure of the organism to express a CS-US association that it had in fact acquired.

Although overshadowing is quite general, the opposite outcome is sometimes found; the presence of a second CS may actually enhance CRs to the less salient CS (Clarke et al. 1979, Palmerino et al. 1980). Such potentiation is usually (but not exclusively) observed when the more salient CS is a gustatory cue. Potentiation has most commonly been observed when the less salient CS is olfactory, but it has also been obtained with CSs in other modalities (Best et al. 1985). Durlach & Rescorla (1980) argued that potentiation may be due to within-compound associations, with the indirect (second-order) association between the target CS and the US that is mediated by the CS of greater salience more than compensating for the reduction in CRs due to overshadowing of the direct (first-order) association between the target CS and the US.
One widely held view of blocking is that “surprise” may be necessary for learning about X (Kamin 1968), and that forward blocking is due to the absence of surprise when the US occurs in Phase 2 (AX-US). Presumably, Phase 1 training (A-US) establishes an expectation of the US based on A. In this view, a change in the US between Phases 1 and 2 should attenuate blocking. Consistent with this prediction, Dickinson et al (1976) and Mackintosh et al (1977) found that omission of a second US after Phase 1 training with two successive USs on each trial (A-US-US) produced “unblocking” (i.e. less blocking). However, several more recent studies have failed to obtain unblocking with notable, albeit different, changes in the US between Phases 1 and 2 (Ganesan & Pearce 1988, Williams 1994). The basis for this disparity is not yet clear (cf Holland 1988).

One variant in the conventional blocking procedure is “backward blocking,” in which AX-US trials precede A-US trials. Backward blocking has received recent attention because it is often obtained with human subjects in causal judgment tasks (Chapman 1991, Shanks 1985), but not with animals in Pavlovian tasks (RR Miller et al 1990). In addition, various models of learning make different predictions about backward blocking. Rescorla & Wagner (1972) predicted forward, but not backward, blocking, whereas Miller & Matzel (1988) predicted both forward and backward blocking. Van Hamme & Wasserman’s (1994) Revised Rescorla-Wagner model more accurately predicts backward blocking in some situations but not others. Empirically, Miller & Matute (1996) and Denniston et al (1996) have found that the observed discrepancies in backward blocking do not arise from potential differences between causal judgment tasks and Pavlovian tasks nor from differences between human and nonhuman subjects; rather, the critical factor appears to be whether the to-be-blocked stimulus has ever been able to elicit robust responding. Robust responding can be elicited by either effective USs or effective CSs. Hall et al (1977) and Miller & Matute (1996) also found that such effective cues are protected against forward blocking.

CONDITIONED INHIBITION Prior discussion considered circumstances under which a CS comes to elicit a CR. One way to view this general result is to regard the CS as activating an expectation of the US, which in turn evokes the CR. Under other training circumstances, a CS can come to reduce the incidence of the CR; such a CS is called a conditioned inhibitor, and it is often viewed as activating an expectation of the nonoccurrence of the US (Wagner & Rescorla 1972).

The earliest measure of conditioned inhibition was developed by Pavlov (1927), who proposed that a stimulus be called a conditioned inhibitor if it passed a “summation” test, which consists of presenting the target stimulus in compound with a previously established conditioned excitor to see whether the
target decreases CRs to the excitor. However, if the conditioned excitor used in
the summation test had been presented together with the target stimulus during
the putative inhibitory training, then the reduced CRs could reflect a discrimi-
nation between the excitor alone (reinforced in training) and the compound of
the excitor and the target stimulus (not reinforced in training). To eliminate
this possibility, Pavlov also did summation tests using conditioned excitors
that had no prior history with the target stimulus. If the target stimulus reduces
CRs to such a “transfer” excitor, then discrimination explanations become
implausible. Still, increased attention to the target stimulus at the cost of
attention to the transfer excitor and/or stimulus generalization decrement from
the transfer excitor alone to the compound of the transfer excitor plus the target
stimulus could account for a summation effect.

To preclude both attentional and generalization decrement accounts, Res-
corla (1969) and Hearst (1972) advocated a second test for inhibition: namely,
that the target stimulus should be paired with the US to see if CRs to the target
stimulus would require more pairings to emerge than would be the case with-
out prior inhibitory training. This “retardation” test is itself subject to an
attentional explanation: Inhibitory training could decrease attention to the
target CS, so that less is learned about it during its pairings with the US.
However, Rescorla and Hearst argued that inhibition training could not both
increase and decrease attention. Therefore, if a target stimulus passes both
summation and retardation tests, then one might reject attentional accounts
(see also Papini & Bitterman 1993, Williams et al 1992). This conclusion
hinges on the possibly incorrect assumption that the same type of attention is
involved in the two tests.

A third assay of inhibition is the approach/withdrawal test of Wasserman et
al (1974), who proposed that subjects will approach a CS with an excitatory
relationship to an appetitive US and will withdraw from a CS with an inhibi-
tory relationship to an appetitive US. These predictions are reversed for aver-
sive USs. The approach/withdrawal test avoids the attentional, configurational,
and generalization decrement accounts of summation and retardation tests.
This test assumes that excitation and inhibition are mutually exclusive, but
some data argue against this assumption (Droungas & LoLordo 1994, Matzel
et al 1988a). Agreement over the assessment of conditioned inhibition lies in
the future.

Generally, a stimulus becomes an inhibitor when it is presented in com-
 pound with another stimulus that predicts the occurrence of a US, but the US
fails to occur. Moreover, Pavlov (1927) found that the inhibitory value of a
stimulus increased with the magnitude of the omitted US. The first method
reported to produce conditioned inhibition used reinforced presentations of an
excitatory CS (A-US) interspersed with nonreinforced presentations of a com-
pound stimulus comprising the reinforced CS and the target cue (AX-No US) (Pavlov 1927). A second technique (the method of negative contingency) for producing conditioned inhibition consists of presenting the target stimulus and the US in an explicitly unpaired fashion (Rescorla 1968). However, if one regards the context as becoming a conditioned excitor because of the unsignaled USs administered in this procedure, then this method is really the same as Pavlov’s, with the context playing the role of the excitor. There are other procedures for producing conditioned inhibition, but they are often less effective than the above two methods (for a review, see LoLordo & Fairless 1985).

OCCASION SETTING When a stimulus is presented with a partially reinforced target CS and reinforcement depends on the presence or absence of that stimulus, we call the stimulus a positive feature in the former case and a negative feature in the latter case. When the feature is simultaneous with the target CS, subjects often learn the feature-positive (-negative) problem by treating the feature like an excitatory (inhibitory) CS. However, when it precedes the CS, the feature often acts by modulating the associative potential of the target CS; in this capacity, the feature is called a positive or negative occasion setter (Holland 1983, 1992).

Some factors that favor a stimulus becoming an occasion setter rather than a simple CS include: (a) inserting a temporal gap (5–15 s) between feature offset and CS onset, (b) having the CS (rather than the feature) occur at the optimal CS-US interval, (c) using a feature from a different modality than the CS, and (d) using a feature that is less salient than the CS (Holland 1985, 1986). Notably, the occasion-setting properties of a stimulus and its properties as a Pavlovian CS are largely independent. Thus, training a negative occasion setter to be a Pavlovian excitor does not degrade its occasion-setting potential, nor does experimental extinction degrade the occasion-setting potential of a stimulus (Holland 1989). A stimulus can even be an occasion setter and a Pavlovian CS simultaneously without appreciable interaction (Holland 1989).

The notion of occasion setting was introduced by Skinner (1938), who proposed that discriminative stimuli set the occasion for operant behavior to be reinforced. Recently, the concept has been applied primarily to Pavlovian conditioning, where it is now attracting considerable attention because occasion setting is seen as one way in which associative theory might explain complex patterns of behaviors, such as biconditional discriminations, that previously appeared to be refractory to simple associative analysis. One problem that could undermine the interpretive value of occasion setting is that substantial training is required before animals master serial feature discriminations. However, research has shown that entire contexts can serve as occasion setters and that contextual stimuli more rapidly acquire occasion-setting properties than discrete stimuli (Bouton & Swartzentruber 1986). Contextual con-
control of associative behavior has long been recognized and occasion setting is one likely means by which contextual control occurs.

THEORETICAL PERSPECTIVES AND CONTROVERSIES

To explain the rich empirical literature of associative learning, some of which we have just reviewed, many theories have been advanced. In this section, we highlight and evaluate some of the most prominent of the recent theories. We also examine the interrelation between learning and expressed behavior.

Contiguity Models

Pavlov (1927) believed that close temporal contiguity was critical for associative learning; the CS and US had to occur near in time for the CS to come to evoke a CR. Long delays between CS and US did not support conditioning, although they could maintain CRs that had been established with more contiguous stimulation (Lucas et al 1981, Ross et al 1979). Thorndike also emphasized temporal contiguity. When a response is shortly followed by a consequence, a reward makes it more likely and a punisher makes it less likely that the situation will again elicit that response. Later work has led theorists to reconsider these ideas. The notion of absolute temporal contiguity has yielded to a relativistic alternative.

Gibbon & Balsam (1981) theorized that the CS-US interval contributes to the strength of CRs relative to the time between USs. In the autoshaping situation “the association between the signal and food depends on an appreciation of the improvement in the average delay to reinforcement in the signal compared with the average delay overall or in the background” (p. 248). Jenkins et al (1981) reached a similar conclusion, that “the effectiveness of a reference trial depends on the waiting time in the reference trial relative to the overall waiting time between feedings within the experimental setting” (p. 276). These two theories are very similar, but they do have distinguishable behavioral implications (Lucas & Wasserman 1982). Although these theories describe acquisition well, they have limitations; for example, they cannot explain US intensity effects (Lucas & Wasserman 1982), and they do not address cue competition.

The idea of relative contiguity can be extended to Thorndikean conditioning. Wasserman & Neunaber (1986) have restated relative contiguity in operant conditioning as “the extent to which the delay of reinforcement after occurrences of the…operant response differs from that after [any of the organism’s other] responses” (p. 31) (see also Fantino & Abarca 1985). Through this temporal discrimination, organisms may be able to detect cases in which one of their responses hastens or delays the occurrence of outcomes. They may also be able to discriminate cases in which their responses have no effect at all...
on the occurrence of outcomes. The notion of relative contiguity is closely aligned with contingency formulations of learning (Peterson et al. 1993).

**Contingency and Rule-Based Models**

A different analysis of Pavlovian and Thorndikean experiments is traceable to Tolman (1932), who argued that learning was basically a matter of discovering “what leads to what.” The concept of “contingency” provides one means of tightening and operationalizing Tolman’s informal approach. In Pavlovian conditioning, the contingency between CS and US can be formalized as the difference between the probability of the US given the occurrence of the CS and the probability of the US given the nonoccurrence of the CS: \( \Delta P = P(US|CS) - P(US|\text{No CS}) \). Emphasis on CS-US contingency broadens the scope of the Pavlovian procedure because it emphasizes that events may be either positively or negatively correlated with one another. Positive CS-US contingencies involve signaled increases in US likelihood, whereas negative CS-US contingencies involve signaled decreases in US likelihood (as in conditioned inhibition). Thanks largely to the efforts of Rescorla (1968), we now know that animals learn about negative CS-US contingencies as well as about positive ones. Thus, nonpairings of CS and US influence behavior as do pairings of CS and US. Tolman’s suggestion that animals learn “what leads to what” is also applicable to Thorndikean conditioning. Here, contingency can be defined by the difference between the probability of an outcome given the occurrence of a response and the probability of an outcome given the nonoccurrence of the response: \( \Delta P = P(O|R) - P(O|\text{No R}) \).

Modern interest in the role of contingency in learning was rekindled by observations that presenting the US without the CS during training decreased CRs to the CS (Rescorla 1968). This finding challenged the view that acquired behavior depended only on CS-US contiguity (cf. Papini & Bitterman 1990). Early contingency theories were intended to replace contiguity as the explanation of acquisition. However, the observation that extinction of the training context can enhance CRs to the CS (Hallam et al. 1992) suggests that contingency may affect the expression, rather than the acquisition of acquired behavior. Thus, Hallam et al. proposed that acquisition of information is governed by contiguity, whereas behavioral expression of that information is governed by contingency. Consistent with this view, Miller & Matzel (1988) have noted that the calculation of contingency requires the subject to know the frequency with which the CS and US have occurred together; but registering a CS-US pairing requires learning an association between CS and US. Hence, learning must occur before contingency can be calculated, and contingency cannot be the basis of learning.

Many important research results in recent years were stimulated by contingency theories. One of these results is that the four types of interevent combi-
nations (e.g. CS and US present) do not contribute equally to acquired behavior. Despite having equal normative weights, they have different psychological weights. Wasserman et al (1990) and Kao & Wasserman (1993) have found that CS-US trials have more impact on behavior than CS only trials and US only trials, both of which in turn have more impact than No CS–No US trials. Various associative (Van Hamme & Wasserman 1994) and rule-based (Buschmeier 1991) models have been designed to incorporate these interevent biases.

A second important result is that USs that are not signaled by the target CS (US only trials) apparently do not enter into a subject’s estimation of contingency if the US is signaled by another “cover” CS (Cooper 1991, Durlach 1983; cf Jenkins & Lambos 1983). The protective effect of cover stimuli was problematic for early forms of contingency theory, but it can be readily reconciled with more recent versions (Cheng 1993).

An important application of contingency theory was provided by Seligman et al (1971), who argued that a history of uncontrollable aversive events may render organisms “helpless.” When later given control over those events, organisms with such prior training may be retarded in learning escape or avoidance responses. “Learned helplessness” has not only had great influence on laboratory research but on clinical psychology as well (Peterson et al 1993). A prior history of uncontrollable appetitive events may also render organisms “lazy” when they are later given control over those events (Engberg et al 1972; cf Beatty & Maki 1979, Schwartz et al 1974).

**Associative Models**

Over the past 25 years, associative models designed primarily to explain cue competition have dominated the field of elementary learning. We briefly review five of them here. We selected these models because each proposes something different from the others, but all of them posit associative links between representations of events.

**RESCORLA-WAGNER** The Rescorla-Wagner (1972) model is a direct descendant of the Bush-Mosteller (1951) linear operator model, in which the strength of an association between stimulus $X$ and a US is constrained by limiting the increase in associative strength to a multiplicative factor of $\lambda - V_X$, where $\lambda$ is the maximum amount of associative strength that can be supported by the US on that trial and $V_X$ is the current associative strength of the stimulus. This model predicts no interaction between the associative status of one stimulus and the change in associative status of another simultaneously presented stimulus; thus, it cannot explain cue competition. The Rescorla-Wagner model surmounts this difficulty by limiting the increase in associative strength to a multiplicative factor of $\lambda - V_\Sigma$, where $V_\Sigma$ is the total associative strength (with respect to the US) of all CSs present on that trial. The difference $\lambda - V_\Sigma$ represents the disparity
between what the subject expected on that trial and what actually occurred, in accord with Kamin’s (1968) contention that learning depends on the “surprisingness” of the US.

More specifically, the Rescorla-Wagner model states that conditioned responding is based on the current associative strength of all stimuli present on a given trial, and that after the trial, the associative value (V) of each stimulus (X) is adjusted according to the formula: \( V_X(\text{new}) = V_X(\text{old}) + \Delta V_X \), where \( \Delta V_X \) (the change in associative strength because of the last trial) = \( \alpha \beta (\lambda - V) \), with \( \alpha \) and \( \beta \) being the associabilities of CS and US, respectively. Extinction is treated like acquisition, except that \( \lambda \) for US absent is zero and \( \beta \) for US absent is nonzero, but smaller than for US present. Conditioned excitation corresponds to positive values of \( V_X \), whereas conditioned inhibition corresponds to negative values of \( V_X \).

The Rescorla-Wagner model explains most cases of cue competition including forward blocking, overshadowing, and the relative validity effect. It also accounts for many conditioning phenomena like acquisition, extinction, discrimination, conditioned inhibition, contingency effects, and the US-preexposure effect. Still, the model suffers from important failures. It does not predict the CS-preexposure effect or one-trial overshadowing (James & Wagner 1980, Mackintosh & Reese 1979); it treats extinction as unlearning, which makes it unable to explain external inhibition, spontaneous recovery, or reminder-induced recovery from extinction (Rescorla & Heth 1975); it wrongly predicts that nonreinforced presentations of a conditioned inhibitor will reduce inhibition (Zimmer-Hart & Rescorla 1974); and it attributes cue competition and the US-preexposure effect to acquisition failure, whereas recovery from these deficits occurs without further training (Balaz et al 1982a, Cole et al 1995, Kaufman & Bolles 1981, Kraemer et al 1988, Matzel et al 1987). Miller et al (1995) reviewed the many successes and failures of the model.

The Rescorla-Wagner model has been the dominant theory of associative learning in the last quarter of the twentieth century. Its success is to be seen more in the research that it stimulated than in the accuracy of its predictions. Our empirical knowledge of learning has been greatly increased because of the many tests of the model’s predictions. Moreover, considerable attention has been given to the model because it is a form of the Widrow-Hoff (1960) delta rule, which is frequently used to correct the weights of connections between units in connectionist models (Quinlan 1991).

The Rescorla-Wagner model assumes that a CS that is absent on a given trial will undergo no change in associative status because \( \alpha = 0 \). This zero value of \( \alpha \) contrasts with the nonzero value of \( \beta \) when a US is absent. Van Hamme & Wasserman (1994) proposed that \( \alpha \) may take a nonzero value on trials in which the CS is absent, if that trial is in some way related to CS
training. Their Revised Rescorla-Wagner model treats CS and US more symmetrically (although at the cost of adding a parameter), and it is able to predict several phenomena that were highly problematic for the original model such as backward blocking, which can be observed with both human (Chapman 1991, Shanks 1985) and nonhuman subjects (Denniston et al 1996, Miller & Matute 1996).

**Mackintosh** Unlike the Rescorla-Wagner model, which assumes that CS processing does not change ($\alpha$ is constant), Mackintosh’s (1975) model posits changes in CS processing as a result of experience ($\alpha$ is variable). It states that on each trial, the associative strength of a stimulus changes according to the rule: $\Delta V_X = \alpha \theta (\lambda - V_X)$, where $\theta$ is a growth parameter. After this change in $V_X$, the associability ($\alpha$) of the cue that best predicted the outcome on that trial will increase and the associabilities of all other cues present on that trial will decrease. The change in $\alpha$ does not affect the present trial, but it will affect learning on the next trial on which the cue occurs. Moreover, the magnitude of changes in $\alpha$ will be inversely related to the accuracy of a cue’s predictions about the US.

To appreciate the model, consider its account of overshadowing (AX-US). On Trial 1, A and X each gain associative value as if the other were not present. Assuming that A is more salient than X ($\alpha_A > \alpha_X$), $\Delta V_A$ will be larger than $\Delta V_X$. On Trial 2, A and X again gain associative value as if the other were not present; but, at the beginning of Trial 2, there is now a difference in their predictive values for the US (A is a better predictor than X). As a result, after Trial 2, $\alpha_A$ increases and $\alpha_X$ decreases. So, on Trial 3, learning about X will be impaired relative to what would have happened if A had not accompanied X on Trials 1 and 2. Mackintosh’s model can explain multitrial but not one-trial overshadowing (James & Wagner 1980, Mackintosh & Reese 1979). It can also explain the relative validity effect and multitrial blocking but not one-trial blocking (Balaz et al 1982b). In addition, unlike the Rescorla-Wagner model, Mackintosh’s model does not predict reciprocal overshadowing between A and X; the data on this issue are mixed (Mackintosh 1971, 1976). Although the model is often described as predicting the CS-preexposure effect, it does so only by making a special supposition explicitly to accommodate the result. The model has stimulated rather little research.

**Pearce-Hall** The Pearce-Hall (1980) model assumes that growth in associability is limited entirely by changes in CS processing. It states that after a trial, $V_X(\text{new}) = V_X(\text{old}) + \Delta V_X$, and $\Delta V_X = S_X \alpha_X \lambda$, with $\alpha_X$ representing the associability of X, $S_X$ the salience of X (a parameter that is largely a function of CS intensity), and $\lambda$ the maximum amount of associative strength that the US can support. The rule for modifying CS associability is $\alpha_X = |\lambda - V_X|$, where $V_X$
is the sum of the strengths of all cues presented on the last trial on which X occurred and $\lambda$ is based on the US presented on that trial. In essence, the model says that the associability of a CS reflects the degree to which the US on the last presentation of the CS was surprising.

The model has enjoyed success in explaining select phenomena that are unexplained by other theories. For example, both the CS-preexposure effect (Lubow & Moore 1959) and the Hall-Pearce (1979) negative transfer effect can be understood in terms of the consistent outcomes of Phase 1 training decreasing the associability of the CS. The model also successfully predicts that the Hall-Pearce effect will be attenuated by a single nonreinforced trial between Phases 1 and 2 because the absence of a US on that trial is surprising and it consequently increases the associability of the CS (Hall & Pearce 1982).

Given the model’s reliance on CS associability to explain learning, an independent measure of associability would be desirable. Toward this end, data have been reported that suggest that the orienting response to a stimulus is a valid measure of its associability (Kaye & Pearce 1984).

PEARCE In Rescorla & Wagner’s model, associative strength is largely allocated to the high-salience elements of compound stimuli and little to the presumably low-salience configural cues involving some or all of the presented elements. This model has had moderate success in accounting for research in compound stimulus control.

Pearce (1987, 1994) has proposed an alternative account of the acquisition of control by compound stimuli that even more explicitly and heavily weights configural cues. According to Pearce, any compound conditioning trial results in a single configural representation of all the stimuli that are present; that representation then enters into association with the outcome of the trial. Pearce’s theory predicts that AB+ training will result in the AB compound receiving all the direct reinforcement available, with responding to A and B alone occurring only insofar as they resemble the AB compound. It further predicts that AB+, A−, B− training will produce positive patterning and that AB−, A+, B+ training will produce negative patterning because all three stimulus conditions in these two tasks differ discriminably from one another. Pearce’s theory also correctly predicts that adding a common element to both positive and negative CSs of the negative patterning problem will slow learning (Pearce & Redhead 1993, Rescorla 1972; see also Redhead & Pearce 1995a), a prediction not made by the Rescorla-Wagner model. Finally, Pearce’s theory outperforms the Rescorla-Wagner model in predicting the correct ordering of responding to the seven stimulus patterns of the following discrimination: (A+, B+, C+) > (AB+, AC+, BC+) > (ABC−) (Redhead & Pearce 1995a). Pearce’s model is an interesting and important theory that merits further experimental examination.
WAGNER Wagner’s (1981) SOP (sometimes opponent process) model is a real-time theory of learning that borrows much of its architecture from Atkinson & Shiffrin (1968) and embellishes it with features of Wagner’s (1976, 1978) priming theory. The model is too complex to detail here. Suffice it to say that the model avoids directly hypothesizing a role for the surprise value of the US, but it achieves the same end through rules about how event representations are distributed among three different memory states and how learning and the expression of memory differentially depend on the contents of these states. The model explains the attenuated UR observed to a US that is given immediately after an initial US or an established CS for that US (Terry 1976). Albert et al (1993) and Ayres & Albert (1990) have also used SOP to explain some of the effects of different ISIs with CSs of different lengths, but the model failed on other counts.

Wagner & Brandon (1989) proposed AESOP (affective extension of SOP) to explain (a) divergences among multiple behavioral indices of a single CS-US association, (b) occasion-setting phenomena, and (c) CRs sometimes resembling URs and sometimes appearing diametrically opposed to URs. AESOP differs from SOP by hypothesizing a dual representation of each CS-US association: one for the emotive qualities of the US (which support general preparatory behavior) and the other for the remaining sensory qualities of the US [which support more focused consummatory behavior (Konorski 1967)]. These representations have different parameters in the equations describing their processing; sensory information is subject to more rapid decay. This difference results in emotive conditioning (e.g. conditioned suppression) being acquired at longer ISIs than sensory conditioning (e.g. eyelid conditioning). AESOP explains occasion setting by the occasion setter eliciting emotive responding that modulates the CR later made to the CS (Bombace et al 1991, Brandon & Wagner 1991, Brandon et al 1991). Given the explanatory successes of SOP and AESOP, it is perhaps surprising that these models have not received more attention; however, they may be too complex to afford unambiguous predictions.

Learning Versus Expression

The very name of the field—associative learning—connotes the acquisition of information. However, the true concern of the field is *acquired behavior*, which involves not only acquisition but also retrieval from long-term memory, decision making, and response generation. The error in the field’s recent narrow focus on acquisition is becoming increasingly clear. There is a growing trend for workers to view many behavioral deficits as failures to express information that had actually been acquired (Bouton 1993, Miller et al 1986, Miller & Matzel 1988, Spear 1973). An earlier generation of theorists better
appreciated the importance of performance or expression variables (Hull 1952, Spence 1956, Tolman 1932; for a review, see Wasserman 1981).

It is important to note the asymmetry in documenting failures of expression and learning: Expression failures can be demonstrated by recovery from the deficit without further training, but acquisition failures defy definitive demonstration. Any failure to demonstrate recovery from a behavioral deficit might reflect a true acquisition failure or an inadequate recovery treatment.

From an expression-failure viewpoint, acquisition might very well be explained by a simple associative rule such as contiguity because phenomena like cue competition can now be explained by performance factors. We next describe two models, each of which focuses on factors that appear to independently modulate the behavioral expression of an association.

**COMPARATOR THEORY** Miller & Matzel’s (1988) comparator hypothesis posits that CRs result from a comparison between a US representation that is directly activated by the CS and a US representation that is indirectly activated by the CS through a “comparator” stimulus. Comparator stimuli are background or other discrete cues that were prominent during CS training. Excitatory responding to the CS is anticipated to the degree that the CS better predicts the US than does the comparator. Behavior indicating inhibition is anticipated to the degree that the comparator better predicts the US than does the CS. This view is much like $\Delta P$ in contingency theory; however, the comparator hypothesis assumes that the critical comparison between these expectations occurs during testing, rather than exclusively during training, as is assumed by contingency theory. Thus, the comparator hypothesis is a model of performance rather than acquisition.

The comparator hypothesis predicts that posttraining changes in the associative value of the comparator stimulus should modify CRs to the CS. Specifically, deflating the associative status of the comparator should attenuate behavioral inhibition and enhance behavioral excitation, and vice versa for inflating the associative status of the comparator. Consistent with this view, Grahame et al (1994) found that posttraining extinction of the experimental context attenuates the CS-preexposure effect, and Matzel et al (1987) found that the same treatment attenuates the US-preexposure effect. Cole et al (1995), Kaufman & Bolles (1981), and Matzel et al (1985) found that posttraining extinction of a competing cue attenuates different forms of cue competition. Hallam et al (1992), Kaplan (1985), and Kasprov et al (1987) found decreases in conditioned inhibition due to posttraining extinction of the excitatory CS or context that was present during inhibition training. Until recently, the comparator hypothesis had little success with its predictions about posttraining inflation of comparator stimuli (RR Miller et al 1990; cf Denniston et
al 1996, Miller & Matute 1996). The hypothesis also predicts one-trial overshadowing (Mackintosh & Reese 1979) and is unique in correctly predicting enhanced inhibition after nonreinforcement of the inhibitor (Zimmer-Hart & Rescorla 1974).

TEMPORAL CODING HYPOTHESIS The weak CRs that are often seen after simultaneous conditioning pose a serious problem for contiguity theory. Barnet et al (1991) and Matzel et al (1988c) explained this result with their temporal coding hypothesis, which states that (a) learning depends on temporal contiguity, (b) the temporal relationship between events is encoded as part of the association, and (c) the temporal relationship between events is a determinant of the nature of the CR. These researchers showed that simultaneous associations may be formed, but may not be expressed because most measures of conditioning involve anticipatory behavior, which would be inappropriate here. In line with this account, Barnet et al (1993) found that one simultaneously trained CS will block control by another simultaneously trained CS, but not by a serially trained CS, and vice versa. Barnet & Miller (1996) also found that, in conditioned inhibition training, subjects encode when an omitted US should have occurred. Holland et al (1996) have extended the idea of temporal encoding to serial occasion setting. They found that features best modulate responding when they precede the testing CS by the same interval that had separated the feature and the training CS.

CONCLUDING COMMENTS

What’s elementary about associative learning? In accord with Thorndike’s and Pavlov’s early speculations, elementary associative learning still seems able to serve as the foundation for our understanding of many complex forms of behavior and cognition. However, our review reveals a rich body of knowledge about associations that surely causes us to question the simplicity of even this basic brand of mentation. Other cognitive processes such as attention, memory, and information processing are now being invoked to help explain the facts of associative learning. The next several years of research will be exciting ones, as neuroscientists and cognitive scientists join experimental psychologists in an interdisciplinary attack on the challenging problems of associative learning and behavior change.

ACKNOWLEDGMENTS

We thank the National Institute of Mental Health for supporting the research of EAW (MH47313 and MH51562) and RRM (MH33881).
Literature Cited


Bennett CH, Maldonado A, Mackintosh NJ. 1995. Learned irrelevance is not the sum of exposure to CS and US. Q. J. Exp. Psychol. 48B:117–28


Holland PC, Hamlin PA, Parsons JP. 1996.
Kamimura et al. 1981, pp. 255–84
Lubow RE, Moore AV. 1959. Latent inhibi-
tion; the effect of nonreinforced exposure to the conditioned stimulus. *J. Comp. Physiol. Psychol.* 52:415–19


Palmerino CC, Rusiniak DW, Garcia J. 1980. Flavor-illness aversions: the peculiar roles...
of odor and taste in memory for poisons. 

Science 208:753–55

Papini MR, Bitterman ME. 1990. The role of contingency in classical conditioning. Psychol Rev. 97:396–403


Parks ER. 1968. The orienting reaction as a mediator of sensory preconditioning. Psychon Sci. 11:11–12

Pavlov IP. 1928. Experiments concerning the salivatory reflexes of dogs. New York: International


Rescorla RA, Heth CD. 1975. Reinstatement of fear to an extinguished conditioned
Shanks DR. 1985. Forward and backward blocking in human contingency judgement. Q. J. Exp. Psychol. 37B:1–21
Smith ST, Putney RT. 1979. Temporal factors in the sensitivity of the alpha rhythm to habituation and conditioning. Physiol. Psychol. 7:381–86


CONTENTS

CREATIVE HYPOTHESIS GENERATING IN PSYCHOLOGY:
Some Useful Heuristics, William J. McGuire 1

WOMEN AND CAREER DEVELOPMENT: A Decade of Research,
Susan D. Phillips, Anne R. Imhoff 31

HUMAN-COMPUTER INTERACTION: Psychology as a Science of
Design, John M. Carroll 61

NEUROBIOLOGICAL CONSTRAINTS ON BEHAVIORAL
MODELS OF MOTIVATION, Karim Nader, Antoine Bechara, Derek
van der Kooy 85

MUSIC PERFORMANCE, Caroline Palmer 115

GENDER, RACIAL, ETHNIC, SEXUAL, AND CLASS IDENTITIES,
Deborah E. S. Frable 139

DISCOURSE COMPREHENSION, Arthur C. Graesser, Keith K.
Millis, Rolf A. Zwaan 163

THE EFFECTS OF STRESSFUL LIFE EVENTS ON DEPRESSION,
Ronald C. Kessler 191

LANGUAGE ACQUISITION: The Acquisition of Linguistic Structure
in Normal and Special Populations, Janet L. McDonald 215

FAMILIES AS SYSTEMS, Martha J. Cox, Blair Paley 243

VISUAL ATTENTION: Control, Representation, and Time Course,
Howard E. Egeth, Steven Yantis 269

PERSONNEL SELECTION, Walter C. Borman, Mary Ann Hanson,
Jerry W. Hedge 299

THE USE OF ANIMAL MODELS TO STUDY THE EFFECTS OF
AGING ON COGNITION, Michela Gallagher, Peter R. Rapp 339

KEY ISSUES IN THE DEVELOPMENT OF AGGRESSION AND
VIOLENCE FROM CHILDHOOD TO EARLY ADULTHOOD, Rolf
Loeber, Dale Hay 371

HEALTH PSYCHOLOGY: What is an Unhealthy Environment and
How Does It Get Under the Skin?, Shelley E. Taylor, Rena L. Repetti,
Teresa Seeman 411

PSYCHOSOCIAL TREATMENTS FOR POSTTRAUMATIC

PSYCHOBIOLOGICAL MODELS OF HIPPOCAMPAL FUNCTION
IN LEARNING AND MEMORY, Mark A. Gluck, Catherine E. Myers 481

ORGANIZATIONAL BEHAVIOR IN THE NEW
ORGANIZATIONAL ERA, Denise M. Rousseau 515

DECLARATIVE MEMORY: Insights from Cognitive Neurobiology,
Howard Eichenbaum 547

WHAT'S ELEMENTARY ABOUT ASSOCIATIVE LEARNING,
Edward A. Wasserman, Ralph R. Miller 573

ATTITUDES AND ATTITUDE CHANGE, Richard E. Petty, Duane
T. Wegener, Leandre R. Fabrigar 609

CENTRAL CHOLINERGIC SYSTEMS AND COGNITION, Barry J.
Everitt, Trevor W. Robbins 649