



Published in final edited form as:

Neurobiol Learn Mem. 2016 May ; 131: 207–215. doi:10.1016/j.nlm.2016.02.014.

Mini-Review: Prediction errors, attention and associative learning

Peter C. Holland and Felipe L. Schifino

Johns Hopkins University

Abstract

Most modern theories of associative learning emphasize a critical role for prediction error (PE, the difference between received and expected events). One class of theories, exemplified by the Rescorla-Wagner (1972) model, asserts that PE determines the effectiveness of the reinforcer or unconditioned stimulus (US): surprising reinforcers are more effective than expected ones. A second class, represented by the Pearce-Hall (1980) model, argues that PE determines the associability of conditioned stimuli (CSs), the rate at which they may enter into new learning: the surprising delivery or omission of a reinforcer enhances subsequent processing of the CSs that were present when PE was induced. In this mini-review we describe evidence, mostly from our laboratory, for PE-induced changes in the associability of both CSs and USs, and the brain systems involved in the coding, storage and retrieval of these altered associability values. This evidence favors a number of modifications to behavioral models of how PE influences event processing, and suggests the involvement of widespread brain systems in animals' responses to PE.

Keywords

attention; prediction error; Pearce-Hall model; associability; associative learning

1. Introduction

Theories of associative learning characterize animals' sensitivity to relations among events. Although early theories emphasized the importance of contiguity between cues (conditioned stimuli, CSs) and the events they predict (unconditioned stimuli, USs), contemporary theories stress the role of reinforcer *prediction error* (PE), the difference between the expected and received value of a reinforcer. Indeed, much modern behavioral, neurobiological, and theoretical investigation of learning hinges on roles of PE. For example, Schultz and Dickinson (2000) noted:

Contact information: Peter C. Holland, Department of Psychological and Brain Sciences, 3400 North Charles Street, Baltimore, MD 21218 USA, pch@jhu.edu, Voice: 410 516-6396, FAX: 410 516-4478.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

(Dr. Schifino is now at the Intramural Research Program of the National Institute on Drug Abuse, 251 Bayview Blvd., Baltimore, MD, 21224.)

Prediction errors can be used in postsynaptic structures for the immediate selection of behavior or for synaptic changes underlying behavioral learning. The coding of prediction errors may represent a basic mode of brain function that may also contribute to the processing of sensory information and the short-term control of behavior.

Most modern learning theories accept CS-US contiguity as a critical learning variable, but redefine that contiguity as between *effective* CSs and USs. One class of models asserts that PE determines the effectiveness of USs (e.g., Rescorla & Wagner, 1972; Sutton & Barto, 1981) and another suggests that PE determines eligibility of CSs (e.g., Mackintosh, 1975; Pearce & Hall, 1980). Here, we discuss roles for PE in altering stimulus *associability*, the ease or rate with which a stimulus may enter into associations. After briefly contrasting representative “reinforcement” and “attention” models, we focus on behavioral and brain system aspects of surprise-induced enhancement of associability.

1.1. The Rescorla-Wagner model

The most well-known model that incorporates PE in learning is the Rescorla-Wagner model (RW; Rescorla & Wagner, 1972). In this “reinforcement” model, the effectiveness of a US in modifying CS-US associations is determined by the difference between the value (associative strength, V) supportable by the US (λ) and the aggregate strength of all CSs present on a learning trial (V_{agg}). Thus, if the US is unexpected (V_{agg} is low, and $\lambda - V_{agg}$ is high) the increment in conditioning (ΔV) for each CS is large, whereas if the US is already anticipated on the basis of past learning (and $\lambda - V_{agg}$ is low), new learning is minimal. Formally, $\Delta V_A = \alpha_A \beta (\lambda - V_{agg})$, where A represents a stimulus, and α and β refer to (constant) rate parameters for the CS and US, respectively. A powerful example of RW’s explanatory power is the case of blocking (Kamin, 1968). Consider two groups of rats that receive pairings of a tone+light compound stimulus with a US. For a blocking group, the tone was pretrained to asymptote with the same US, whereas for the control group it was not. Although rats in the control group acquire a substantial CR to the light, the rats in the blocking group do not. Within R-W, substantial conditioning accrues to both tone and light in the control group because the US was initially unexpected on compound trials. By contrast, in the blocking group, because the US was already well-predicted by the tone, the US is no longer effective ($\lambda - V_{agg} = 0$) in supporting new learning in the compound phase.

An important feature of RW is that its equation provides a simple basis for distinguishing between excitatory ($V > 0$) and inhibitory ($V < 0$) CSs, as well as specifying the conditions under which excitatory [$(\lambda - V_{agg}) > 0$, a positive PE] and inhibitory [$(\lambda - V_{agg}) < 0$, a negative PE] learning occur. Thus, the difference between the expected and received values of a US determines both the magnitude and direction of learning. These simple assumptions allowed the model to account for many puzzling associative learning phenomena, as well as to predict several counterintuitive outcomes. The popularity of this model soared further when Schultz (Schultz & Dickinson, 2000) showed that midbrain dopamine neurons show corresponding PE-reflecting firing patterns. That is, early in CS-US pairings, the US provoked firing bursts (large positive PE), but as training continued, the increasingly-expected US produced smaller and smaller firing-rate increases, and instead the CS came to

control those responses. Furthermore, if an expected US was omitted, these neurons showed reductions in firing rates, relative to baseline, consistent with a negative PE.

1.2. The Pearce-Hall model

In the Pearce-Hall model (PH; Pearce & Hall, 1980), processing of the US remains constant, but PE modulates processing of the cues present when the PE is induced. Specifically, the *associability* (α) of a cue is proportional to the absolute value of the aggregate PE on previous trials. Thus, the change in the strength of a cue A on a learning trial n is $\Delta V_A = \alpha_{An} \lambda$, where $\alpha_{An} \approx |(\lambda - V_{agg})_{n-1}|$. Because α is assumed to begin at a high level (novel cues attract attention) and USs are initially unexpected, learning occurs rapidly on early trials. But as the US becomes better predicted, α_A declines, and learning slows. Thus, this model asserts that as the US becomes better predicted by CSs, the associabilities of those cues decrease. Considerable evidence supports this somewhat counterintuitive claim. For example, repeated pairings of a CS with one US make it more difficult to associate that CS with another (e.g., larger or smaller) US (Holland, 2005, Hall & Pearce, 1979). Similarly, if a cue is first paired with a US in compound with another CS that already predicts that US, not only does the new cue fail to acquire a CR during compound conditioning (blocking), but also it is slow to acquire associations if paired later with a US in the absence of the previously-trained CS. Furthermore, this slower learning about predictive cues does not occur in rats with lesions of the hippocampus or of its cholinergic afferentation (Baxter et al., 1997, 1999; Han et al., 1995; Holland & Fox, 2003).

By contrast, if the US is poorly-predicted (such that PE remains large), a cue's associability may be maintained at a higher level. For example, if a cue is followed by the US on only half the trials, negative PEs will ensue on nonreinforced trials and positive PEs will occur on reinforced trials, keeping the CS eligible for more rapid subsequent learning than if it had been consistently paired with the US (Holland, 2005). Likewise, rats that previously received a mix of light→tone and light-alone trials in a “serial prediction task” (§2.1) acquire light-food associations more rapidly than rats that first receive only light→tone pairings (Holland et al., 2002). Importantly, after a cue's associability has been driven down by consistent pairing with a US or other event, it can be restored simply by presenting that cue without the expected event. Thus, the surprising omission of an expected event can enhance subsequent cue learning. Interestingly, although we have demonstrated involvement of a few brain regions in associability decreases, and identified an extensive circuit critical for associability increases (§2), we have never found convincing evidence for a brain structure critical to both. Thus, mechanisms for increases and decreases in associability appear to be separable, despite being described with a common equation in PH.

1.2.1. Attention in learning and action—Although intuition and much data tell us that animals should attend more to reliable predictors of important events (Mackintosh, 1975; Anderson et al., 2011), within PH reliable predictors of important events are attended to *less* than unreliable predictors. However, the ecological demands on attention in learning and action are likely quite different. Whereas it seems clear that action decisions are optimized by a bias to attend to the most reliable predictors of the future, in learning it may be more adaptive to bias attention to cues whose consequences are not yet well-known, rather than to

preferentially attend to cues whose predictive powers are already established. Recognizing this dilemma, PH distinguished between “controlled” and “automatic” attention, leaving open the possibility that attention for learning and action could follow different rules and use different information: the α that informs learning may not direct action. Behavioral and neuroscientific evidence from our laboratory (e.g., Holland & Maddux, 2010; Maddux & Holland, 2011) showed that in many circumstances, whereas attention in learning is guided by PE, attention in action is guided by prediction. Furthermore, these aspects of attention may be expressed and altered independently, in terms of both behavior and neural systems. In tasks that demand simultaneous action decisions and new learning, Maddux et al. (2007) found that rats may simultaneously attend to one element in an array for purposes of action but to another element for purposes of new learning, and that these functions may be doubly-dissociated within medial prefrontal and parietal regions. However, in this review we focus solely on surprise-induced increases in stimulus associability.

2. Surprise-induced increases in associability

Rescorla and Holland (1982) noted that the study of any learning process involves consideration of three questions: (1) what are the conditions that produce the learning, (2) what are the contents of that learning (what information is acquired and how is it represented), and (3) how is that learning revealed in the performance of the organism? An extension of that approach to the search for neural bases of surprise-induced enhancements of cue associability distinguishes among *surprise module* systems responsible for altering associability at the time of surprise, *storage module* systems that represent altered associability values, and *expression module* systems that access those altered values to implement more rapid learning later. In this review, we describe investigations of behavioral and brain functions within these three modules.

According to PH, anytime a reinforcer is unexpectedly presented or omitted, the associability of all contiguous cues should be enhanced, encouraging future learning about those cues. However, because reinforcement models also predict enhanced learning when reinforcers are unexpected, investigations of associability changes have primarily used test procedures that pit reinforcer- and cue-processing accounts against each other. For example, in the *unblocking* procedure, animals are first trained with pairings of one cue (A) with a US, followed by pairings of a compound of A and a new cue, X, with either the same US, a lower-value US, or a higher-value US. If the US is unchanged when X is added, no PE ensues, and little is learned about X (*blocking*, §1.1). However, if the value of the US is shifted upward, the ensuing positive PE should, within RW, provide an effective reinforcer for excitatory learning about X, and, within PH, enhance the associability of X, speeding its association with the US. Thus, the occurrence of upshift unblocking does not distinguish between changes in cue or reinforcer processing. By contrast, if the value of the US is shifted downward when X is added to A, according to RW, the resultant negative PE should produce inhibitory learning about X, whereas PH again anticipates enhanced associability of X, permitting its excitatory association with the new, lower-valued US under some circumstances. Thus, the observation of excitatory learning to X in downshift unblocking is unambiguously attributable to PH-related processes, and brain manipulations that disrupt that learning would inform systems responsible for those processes.

2.1. The Serial Prediction Task (SPT)

Our studies of brain mechanisms for associability changes have used several different procedures to induce negative PEs, including downshift unblocking. However, we have most often used a *serial prediction task* (SPT; Wilson et al., 1992), which provides analytic advantages over most other procedures. First, it permits assessment of the effects of brain manipulations on both enhancements and losses (§1.2) in associability. Second, by arranging the induction of negative PEs (and presumably the recalculation of cue associability) in one phase of the experiment but assessing the expression of altered cue associability as more rapid learning in a subsequent phase, it allows us to separately examine the effects of brain manipulations on the induction and expression of associability changes.

In this procedure (Figure 1A), rats first receive consistent serial light→tone pairings in an *expectancy* phase to establish the light as a highly valid predictor of the tone. Half of those serial compounds are reinforced with food delivery and half are not. Next, in a *surprise* phase, for *shift* rats the tone is omitted on the nonreinforced trials, whereas *consistent* rats continue receiving the same trials as in the expectancy phase. Finally, the associability of the light is assessed in a test phase in which the light is directly paired with food. Within PH, as the light comes to predict the tone in the expectancy phase, its associability decreases. In support of that assertion, Holland et al. (2002) found that the more consistent training is given, the slower acquisition to the light is in test. However, violation of that prediction in shift rats during the surprise phase restores or enhances that associability. Indeed, among intact rats, shift rats routinely show faster learning of the new light-food relation in the test phase, compared to consistent rats (Holland & Gallagher, 1993a; Wilson et al., 1992; Figure 1B). Notably, this shift advantage is not anticipated by most other learning theories. For example, to the extent that long-delay light-food conditioning might occur to the light in the expectancy phase, RW would expect the shift treatment to especially extinguish that conditioning because light-alone presentations are always nonreinforced and the tone is always present on reinforced trials. Likewise, most other attentional theories (e.g., Mackintosh, 1975) also predict a shift *disadvantage*, because losses in associability that occur when cues are nonreinforced would be greater when the light is presented alone in the shift group than when it is presented in compound with the tone in the consistent group (Lubow et al., 1982).

Rats with impaired function in brain regions involved in surprise-induced associability changes either fail to show the shift advantage in the SPT or show a shift *disadvantage*. For example, Holland and Gallagher (1993a) found that among rats with bilateral lesions of the amygdala central nucleus (CeA), shift rats showed *slower* learning about the light in test than consistent rats, which learned at the same rate as sham-lesioned consistent rats (Figure 1C). Observations of this shift disadvantage in lesioned rats suggests that PH associability-enhancing processes are normally overlaid on other processing mechanisms, which are revealed when PH processes are eliminated. Furthermore, the identical rates of test acquisition among lesioned and intact rats in the consistent condition indicate that CeA is not involved in associability losses. If it were, then after consistent training, lesioned rats would learn more rapidly than controls in testing.

Comparable lesion experiments identified a number of brain regions as critical to the display of a shift advantage in this task, including substantia nigra pars compacta (SNc, Lee et al., 2006), substantia innominata/nucleus basalis magnocellularis (SI; Chiba et al., 1995) and its cholinergic afferentation of the posterior parietal cortex (PPC; Bucci et al., 1998), the dorsolateral striatum (DLS; Asem et al., 2015), the lateral hypothalamus (LH; Wheeler et al., 2014), and portions of the medial prefrontal cortex (Maddux, 2008). As with CeA lesions, none of these manipulations affected losses in cue associability in the consistent condition.

2.1.1. Module-limited function in brain processing—To separate a brain region's involvement in registration of surprise and/or the recalculation of α (associability) from its function in expressing that altered value in new learning, we transiently altered function in that region pharmacologically prior to either surprise or test sessions. Holland and Gallagher (2006; Figure 1D) found that disrupting function of CeA prior to surprise sessions or of SI prior to test sessions prevented the normal shift advantage. The shift advantage was observed in both vehicle-infused control rats and in rats that experienced CeA disruptions at the time of test or SI disruptions at the time of surprise. Thus, CeA is a critical part of the surprise module but is not needed for expression of enhanced associability once α is recalculated, whereas SI function is unnecessary at the time of surprise but is essential for expression of enhanced associability in new learning. There was no evidence that either structure was involved in reductions in associability: the rats in the consistent groups all learned at the same rate. Subsequent experiments of this type suggest the SNc (Lee et al., 2008) to be within the surprise module only, and the secondary visual cortex (SVC; Schiffino, 2015; Schiffino & Holland, 2016b) and the DLS (Asem et al., 2015) as participating in the expression module only. We found only one region, the PPC, critical to both the surprise and expression modules (Schiffino et al., 2014a). We discuss the importance of this observation in §3.

2.1.2. Regional Connectivity: Disconnection lesions—A disadvantage of bilateral lesion/inactivation procedures in circuit-mapping is that they affect all functions of the targeted region. The *disconnection* procedure reduces this complication by focusing on the necessity of convergence of information processing by two different brain regions. With this procedure, a lesion or inactivation of one structure in one hemisphere is combined with a contralateral lesion/inactivation of a second structure. Because the projections of the regions we investigated are almost exclusively ipsilateral, such a contralateral lesion greatly reduces the opportunity for convergence of information processing by those two regions, but in each hemisphere spares functions that involve interactions of one target structure with all other regions. SPT performance of rats with contralateral lesions was compared with that of rats with ipsilateral lesions/inactivations of the same two structures, which control for the amount of damage to each structure, but leave the opportunity for convergence intact in one hemisphere. Conveniently, in most cases, we found that rats with ipsilateral damage were unimpaired. We found that lesions that disconnected CeA from SNc (Lee et al., 2006), SI (Han et al., 1999), or DLS (Esber et al., 2015) all abolished the shift advantage found in ipsilateral controls, without producing differences in performance in the consistent condition. By contrast, Wheeler et al. (2014) found that both contralateral and ipsilateral

lesions of CeA and LH, as well as unilateral lesions of LH alone, eliminated the shift advantage, suggesting that LH's influence on salience enhancement did not depend on any privileged convergence of information processing by CeA and LH.

2.1.3. Regional connectivity: FOS expression and anatomical tracing—The disconnection procedure does not distinguish among direct single-synapse connections between two regions, indirect multi-synapse connections, or the case in which a third region requires activation by both of the two target regions. Nor does it permit inferences about the direction of information flow between reciprocally-connected regions. To examine direct functional connections engaged by surprise, we combined the use of retrograde tracers with assessment of the expression of FOS protein, often used as a token of neural activity. Rats were exposed to training that established certain expectancies, and then tested under conditions in which those expectancies were either violated (PE) or confirmed. Comparisons of FOS expression across these test conditions permitted inferences about participation of neurons in various brain regions in processing of PE and other events. To examine functional connections among regions, prior to training we infused retrograde tracers (Flurogold or CTb) into target regions known to receive afferents from neurons in a region of interest. Observation of double-labeling for FOS and the tracer in the region of interest allowed us to determine whether neurons that project to a particular target region are important in coding PEs, and hence whether that target region could receive PE information directly from the region of interest.

In these experiments we sought only to observe brain responses to the induction of positive or negative PEs themselves, apart from changes in associability. Thus, we abandoned the SPT for a simple discriminative conditioning procedure in which one cue (CS+) was paired with a food US and another cue (CS-) was not. After learning was complete, rats were tested with one of four combinations of CS and US: CS- and food (surprising food; positive PE), CS+ and no-food (surprising omission of food; negative PE), CS+ and food (expected food; no PE), and CS- and no-food (expected no-food; no PE). These studies revealed that whereas CeA neurons that project to VTA coded only the presence of CS+, CeA neurons that project to SNc coded both positive and negative PEs (Lee et al., 2010). This observation indicates that the CeA-SNc interaction revealed by our disconnection lesion experiments (§2.1.2) involves (at least in part) direct projections from CeA to SNc. Interestingly, dopaminergic (TH+) neurons in SNc coded surprise less powerfully than TH- neurons (Holland et al., 2016). Likewise, surprise was strongly coded by TH- neurons in substantia nigra reticularis (SNr), a major output nucleus for the basal ganglia, which sends efferents to thalamic regions that project to PPC (Deniau & Chevalier, 1992). By contrast, CeA neurons that project to SI did not appear to code surprise, nor did the double-labeled subset of SNc neurons that projected to SI (Holland et al., 2016). These last observations weigh against both an early hypothesis (Gallagher & Holland, 1992) that CeA conveys surprise information directly to SI, and a later one (Holland & Gallagher, 2006) that SI receives CeA-generated PE information via SNc. However, the lack of a role for these projections at the time of surprise is consistent with the possibility that SI does not participate in a surprise module, but is only engaged later, at the time of new learning.

2.2. Beyond the serial prediction task

Although we noted analytical advantages of the SPT, any single procedure is unlikely to capture all aspects of associability change. In this section we describe experiments with a number of procedures, which support the assertions we made based on the SPT, but also indicate that different tasks can engage different brain systems, and that the conditions that favor associability change are more extensive and complex than indicated by the original PH model.

Gallagher and Holland's (1992) claim that the CeA might be involved with the modulation of attention in associative learning was initially met with skepticism. Thus, we examined the effects of CeA lesions on learning in several procedures that induce negative PEs in different ways. Rats with CeA lesions showed learning deficits specific to associability enhancements in serial negative patterning (Holland et al., 2000), temporal uncertainty (Wheeler & Holland, 2011), downshift unblocking (§2, Holland & Gallagher, 1993b), and overexpectation (Haney et al., 2010; Holland, 2016) procedures. By contrast, CeA-lesioned rats showed no deficits in learning in upshift unblocking (§2, Holland & Gallagher, 1993b; Holland, 2006), or when negative PEs were induced using feature negative (conditioned inhibition) procedures (Holland et al., 2000; Holland 2012, 2016). We consider some especially informative procedures in the remainder of §2.

2.2.1. Unblocking and stimulus associability—In the unblocking procedure (§2), animals are first trained with A-US pairings, followed by pairings of AX with either the same US, a lower-value US, or a higher-value US. According to PH, either upshifts or downshifts in US value would produce PEs, and thus enhance the associability of cues present at that time, enabling excitatory learning between X and the US under some circumstances. For example, if the initial training involves two serially-presented USs (A→US1→US2), and the second is omitted when X is added (AX→US1), excitatory X→US1 associations could be formed. Such excitatory learning is indeed often observed in these circumstances (Dickinson, et al., 1976; Holland, 1988). However, rats with CeA lesions failed to show unblocking with such downshifts but were unimpaired in unblocking with upshifts (Holland and Gallagher, 1993b), even under conditions in which it could not be attributable to RW mechanisms (§1.1, Holland, 2006). Thus, we concluded that CeA was involved only in enhancing associability induced by negative PEs. That conclusion was supported by an electrophysiological study of rats performing a value-reversal task in which reinforcer value was shifted up or down (or maintained) over successive blocks of trials. Neurons in CeA were found to increase firing rates after downshifts, but not upshifts (Calu et al., 2010)

By contrast, BLA lesions disrupt unblocking with either upshifts or downshifts in US value, including a case in which upshift effects could be attributed specifically to associability enhancements (Esber & Holland, 2014; Chang et al., 2012). Furthermore, both electrophysiological (Esber et al., 2012; Roesch et al., 2010) and IEG studies like those described previously for CeA (Holland et al., 2016) showed evidence for coding of both positive and negative PE in BLA. Interestingly, although early studies showed parallel outcomes with downshift unblocking and SPT in studies involving lesions of CeA (Holland

& Gallagher, 1993a,b) or cholinergic innervation of PPC (Bucci et al., 1998), in two experiments (Holland et al., 2001; unpublished data, 2010), we found no evidence for an effect of BLA lesions in SPT. These contrasting observations suggest that BLA is especially engaged in processing shifts in the value of biologically-significant events (USs), which do not occur in the SPT (in that task, only the predictive relation between two CSs is manipulated.)

2.2.2. Overexpectation and cue associability—The overexpectation procedure (Rescorla, 1999) produces negative PEs without changing either the value or probability of the US. After two CSs are each separately paired with the US until learning is near asymptote, the two CSs are compounded and followed by the same US. Because the aggregate predicted value is greater than the value of the delivered US, a negative PE is produced. According to RW, that PE should encourage inhibitory learning to A and B until their aggregate prediction matches the delivered US. According to PH, that PE might also enhance the associabilities of A and B, allowing rapid losses in their strengths. Consistent with that conjecture, Haney et al. (2010) and Holland (2016) found that these losses were slowed by CeA lesions/inactivations, but not by BLA lesions/inactivations. Thus, although CeA is critical to associability enhancements that occur with negative PEs in overexpectation, the SPT, and unblocking with downshifts, BLA is needed only for the last of these, consistent with our conjecture that BLA may be especially engaged for processing of shifts in reinforcer value.

2.3. Changes in reinforcer associability

Although PH describes a mechanism for alterations in the associability of CSs, evidence suggests that the effects of surprising reinforcer presentation or omission extends to reinforcers, altering the rate at which they can participate in new associations. Consider a downshift unblocking experiment, in which one cue (A) is first paired with a sequence of one US (US1) followed by another US (US2), and then AX is followed by US1 alone. Although within PH the negative PE resulting from US omission should enhance the associabilities of the CSs, it might also enhance the associability of the remaining US1. Indeed, one might argue that because US1 is closer in time than the CSs to the omitted US2 (a variable that affects associability changes; Holland, 1988), US1 might especially benefit. Several findings from unblocking experiments support this possibility.

First, if induction of a negative PE enhances cue associability, then that cue should acquire learning faster in any task. However, we found no CeA lesion deficit when we tested rats' ability to form inhibitory X-US2 associations after US2 omission in an unblocking experiment (Holland & Kenmuir, 2005), despite observing lesion deficits in forming excitatory X→US1 associations. These findings are consistent with the idea that omission of US2 makes US1, rather than X, more associable. Second, if US associability changes in unblocking, we would expect the same effects even if X was absent when the US value was shifted. Holland and Kenmuir (2005) first trained rats with A→US1→US2, and then omitted US2 (A→US1). Subsequent X→US1 learning was enhanced, but X→US2 learning was not. Importantly, this enhancement was CeA-dependent. Enhanced X→US1 but not X→US2

learning was also found in analogous upshift experiments, but these enhancements were not CeA-dependent, as we also observed in standard upshift unblocking (Holland, 2006).

2.4. Assignment of credit in associability changes. Insights from conditioned inhibition procedures

Although many procedures that induce negative PEs produce CeA-dependent associability enhancements, the feature-negative (or “conditioned inhibition”) procedure, $A \rightarrow US$, $XA \rightarrow \text{nothing}$, appeared to be a puzzling exception. According to PH, the large negative PE on early XA trials should enhance the associability of both A and X , contributing to rapid learning of inhibitory $X \rightarrow US$ (or $X \rightarrow \text{“no US”}$) associations, which underlie solution of the discrimination. However, across a range of cue modalities, saliences, temporal arrangements, and amounts of training (Holland et al., 2000; Holland, 2012; Holland & Kenmuir, 2005) we never observed an effect of CeA lesions on acquisition of these discriminations.

Holland (2012) suggested a modification of PH such that a cue’s associability benefits from PEs only if that cue participated in the generation of the PE. In the conditioned inhibition procedure, although a large negative PE is induced early on nonreinforced AX trials, X is always presented nonreinforced, and does not itself generate a prediction of the US. If associability increments are assigned only to cues that contributed to the PE, rather than distributed to all cues present, as PH assumed, we would expect $A \rightarrow US$, $AX \rightarrow \text{nothing}$ procedures to enhance A ’s associability but leave X ’s unchanged. Holland (2012) found that to be the case, and A ’s enhanced associability was eliminated by CeA lesions. Importantly, this outcome is also consistent with an earlier modification to PH (Pearce & Mackintosh, 2010), in which cue associability is determined by individual PE ($\alpha_A = |\lambda - V_A|$). An elaboration of the overexpectation experiment (§2.2.2) extended this observation (Holland, 2016). A novel X cue was introduced when the overexpectation of reinforcement was produced by compounding two previously-reinforced elements ($A \rightarrow US$, $B \rightarrow US$ followed by $ABX \rightarrow US$). According to RW, because the three cues are followed by a negative PE, they should each acquire inhibitory learning, resulting in the loss of conditioning to A and B , and the establishment of net inhibition to X , despite its consistent pairing with food. We found this pattern of results in intact rats. Within PH, this negative PE should also enhance the associability of each of these cues in intact rats, speeding these changes. Thus, these changes should occur more slowly in CeA-lesioned rats. However, although we indeed observed less loss of responding to A and B in lesioned rats than in controls (as in §2.2.2), there were no such differences in inhibitory learning to X . This pattern of results is consistent with CeA-dependent associability enhancements for A and B (which contributed to the negative PE) but not X (which did not). Notably, these results (and those described in §2.2.2) are inconsistent with Pearce and Mackintosh’s (2010) suggestion that associability changes are linked to cues’ individual error terms, because those PEs for A and B in the overexpectation experiment are initially zero.

3. Memory for altered associability

3.1. PPC as a storage site for associability memories

Within PH, the violation of outcome expectancies today alters the associability of cues tomorrow. Thus, there must be some relatively permanent memory of this altered cue associability, hence a *storage module*. The most obvious storage site candidates would be regions whose function was required both when surprise-induced changes in associability were encoded and when they were retrieved for new learning. Although many years of investigation failed to reveal such a region, Schiffino et al. (2014a) found that normal PPC function is critical to both “surprise” and “expression” modules, identifying PPC as a reasonable candidate for the storage of altered associability memories. Subsequently, we found that infusion of the protein-synthesis inhibitor anisomycin into PPC immediately after surprise sessions in SPT prevented (in a time-dependent manner) the consolidation of an altered associability memory.

Schiffino et al. (2014b) then identified cue-specific neurons in PPC that appeared to track and represent associability changes. One visual cue (CS1) was paired with 1 drop of sucrose delivered to one sucrose well, and a second visual cue (CS3, differing in flash rate) was paired with delivery of 3 drops of sucrose to another well. Rats received sucrose only if they entered the correct well. After a rat reached 85% correct responding, it received an extended block of trials in which the reward for one of the cues (counterbalanced) was shifted to the other reward, followed by a second block of trials in which the reward for the other cue was shifted. Because the correct responses remained the same for each cue, changes in neuronal firing were related to reward changes and not changes in choice responding. Of 145 units recorded in 5 rats, 71% were responsive to the cue light onsets, increasing or decreasing their firing rates relative to pre-trial baseline rates. Of the 42 units that showed rate increases, 12 selectively increased their firing rates to the cue that accompanied a reward shift. These modified responses to the cues were maintained over time, consistent with a memory representation of associability. This was the first observation of the modulation by surprise of neural firing to cues themselves. Previous research that examined amygdala neurons (e.g., Calu et al., 2010; Roesch et al., 2010) identified regions in which *reward-responsive* neurons increased their firing rates when cue-reinforcer contingencies were changed. Thus, they were sensitive to PE, representing surprise itself, whereas Schiffino et al.'s (2014b) study showed the modulation of *cue-responsive* neuron activity by surprise, a more direct correlate of associability. The representation of altered associability values in PPC is consistent with prior identification of PPC as a critical component of attention networks (Petersen & Posner, 2012; Reep & Corwin, 2009), which may supply visuospatial priority maps to direct the deployment of attention (Bisley & Goldberg, 2010; Gottlieb et al., 2014). Our data indicate that when animals seek to learn, stimuli with uncertain relationships are given priority over others in part through activity of PPC.

3.2. Systems consolidation of associability by CeA

Another set of studies (Schiffino, 2015; Schiffino & Holland, 2016a) showed that the consolidation of a surprise-enhanced associability memory in the SPT depends on CeA processing 1-4 hours after surprise sessions. Post-surprise infusions of anisomycin,

lidocaine, or muscimol prevented subsequent display of surprise-enhanced associability. Because previous studies suggested that CeA function is unnecessary for expression of associability enhancements that were induced previously when CeA function was intact (Holland & Gallagher, 2006), we interpreted these results as indicating that post-surprise activity of CeA (“surprise replay”) is necessary for the consolidation of altered associability memories elsewhere (McGaugh, 2000), such as in PPC.

4. Summary

4.1. Behavioral implications

Our use of both behavioral and neural systems approaches has extended understanding of associability changes in associative learning. First, increases and decreases in associability may be separately determined, despite being specified by a single equation in PH. As noted in §1.2, lesions that affect one do not affect the other. Furthermore, these changes often appear to develop at very different rates. For example, in SPT, associability increases reach asymptote in one or two surprise sessions, whereas associability decreases accumulate over tens of consistent sessions (Holland et al., 2002). This asymmetry follows from the necessity to gradually acquire CS-US associations in the expectancy phase of SPT to drive associability down, whereas maximum PE in the surprise phase occurs immediately. Interestingly, to reduce this asymmetry, Pearce et al. (1982) added a dampening factor to make changes slower and more symmetrical. In that version of PH, α is determined by the events on the previous trial (as in PH) plus an exponentially-weighted moving average of its preceding values, $\alpha_n = \gamma|(\lambda - \text{Vagg})_{n-1}| + (1-\gamma)\alpha_{n-1}$. The truth may lie between; although Holland et al.’s (2002) data support the original PH version, changes in neuronal responses that track prediction error are better fit with the latter function (Calu et al., 2010; Roesch et al., 2010).

Second, several experiments suggested different mechanisms for associability increases produced by positive and negative PEs. For example, whereas processing of both is represented in BLA, only associability enhancements produced by negative PE demand CeA processing (§2.2.1). Interestingly, although RW posits a simple symmetrical way of dealing with the effects of positive and negative PEs (§1.1), PH has a more elaborate way of dealing with these changes (beyond the scope of this review), which may demand extra layers of processing. Third, the *assignment of credit* experiments (§2.4) suggest that only cues that contribute to the aggregate PE receive associability enhancements. Finally, PH seems to apply to changes in the associability of USs as well as CSs: surprising presentation or omission of US2 in a US1→US2 sequence enhances the ability of US1 to enter into new associations, apart from any changes in CS associability (§2.3).

4.2. Brain circuitry

Gallagher and Holland (1992) proposed a simple circuit in which CeA-processed surprise information was propagated to the cholinergic basal forebrain, which in turn broadly modulated cortical activity. Evidence now suggests a much broader, but still-unspecified network, involving not only these elements, but also midbrain modulation of striatal and cortical circuitry. We have identified brain circuitry for surprise-enhanced associability in

three modules. A surprise module in which PEs are calculated and/or used to alter cue associability includes CeA, BLA, PPC, portions of medial pre-frontal cortex, SNc, and perhaps SNr, although the role of these structures may be task-specific. Evidence that other behavioral functions that require PE (such as conditioned inhibition) are spared by lesions of these regions suggest that they may be specialized for using PE to alter stimulus associability, rather than being critical for calculating those PEs. Early hypotheses of how recalculated associability values are conveyed to PPC for storage have not been confirmed, and remain a subject of speculation (Schiffino et al., 2014a; Schiffino, 2015). Likewise, we identified several regions that are critical to the expression of enhanced associability in faster learning but not to the original encoding of altered associability, including SI, SVC, DLS, and LH. It is important to recognize that the involvement of these regions was specific to learning after associability enhancements: we observed no role in learning in control procedures (e.g. the consistent condition in SPT). Again, we can only speculate about circuitry involved in the transmission of altered associability memory to these regions (Schiffino et al., 2014a; Schiffino, 2015). However, it seems reasonable to presume that these memories are embedded in circuits responsible for associative learning, rather than those involved in early-stage stimulus processing. If early-stage cue processing were enhanced by PE, then one might expect that all cue functions would be enhanced. Our observations that attention in action is biased towards more predictive cues, whereas attention for learning (associability) is correlated with PE, suggest otherwise.

Finally, we must reiterate that our review is largely confined to mechanisms for surprise-induced enhancements of cue associability, within the context of the PH model. We have ignored evidence that in many cases highly predictive cues can acquire high salience, not just for controlling action (as we described in §1.2.1), but also in new learning (e.g., Mackintosh, 1975; Pearce & Mackintosh, 2010). In part to deal with increasing evidence that both prediction and prediction error can affect cue salience, many new attention-based models of associative learning have been proposed (e.g., Esber & Haselgrove, 2011; LePelley, 2004; Pearce & Mackintosh, 2010). A fuller characterization of brain mechanisms for attention in associative learning in the future demands consideration of insights gleaned from those efforts.

Acknowledgments

Funding acknowledgement: Most of the research described in this article was funded by grant MH-53667 from the National Institutes of Health.

References

- Anderson BA, Laurent PA, Yantis S. Value-driven attentional capture. *Proceedings of the National Academy of Sciences (USA)*. 2011; 108:10367–10371.
- Asem JSA, Schiffino FL, Holland PC. Dorsolateral striatum is critical for the expression of surprise-induced enhancements in cue associability. *European Journal of Neuroscience*. 2015; 42:2203–2213. [PubMed: 26108257]
- Baxter MG, Gallagher M, Holland PC. Blocking can occur without losses in attention in rats with selective removal of hippocampal cholinergic input. *Behavioral Neuroscience*. 1999; 113:881–890. [PubMed: 10571472]

- Baxter M,G, Holland PC, Gallagher M. Disruption of decrements in conditioned stimulus processing by selective removal of hippocampal cholinergic input. *Journal of Neuroscience*. 1997; 17:5230–5236. [PubMed: 9185560]
- Bisley JW, Goldberg ME. Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience*. 2010; 33:1–21.
- Bucci DJ, Holland PC, Gallagher M. Removal of cholinergic input to rat posterior parietal cortex disrupts incremental processing of conditioned stimuli. *Journal of Neuroscience*. 1998; 18:8038–8046. [PubMed: 9742170]
- Calu DJ, Roesch MR, Haney RZ, Holland PC, Schoenbaum G. Neural correlates of variations in event processing during learning in central nucleus of amygdala. *Neuron*. 2010; 68:991–1001. [PubMed: 21145010]
- Chang SE, McDannald MA, Wheeler DS, Holland PC. The effects of basolateral amygdala lesions on unblocking. *Behavioral Neuroscience*. 2012; 126:279–287. [PubMed: 22448857]
- Chiba AA, Bucci DJ, Holland PC, Gallagher M. Basal forebrain cholinergic lesions disrupt increments but not decrements in conditioned stimulus processing. *Journal of Neuroscience*. 1995; 15:7315–7322. [PubMed: 7472485]
- Deniau JM, Chevalier G. The lamellar organization of the rat substantia nigra pars reticulata: Distribution of projection neurons. *Neuroscience*. 1992; 46:361–377. [PubMed: 1542412]
- Dickinson A, Hall G, Mackintosh NJ. Surprise and the attenuation of blocking. *Journal of Experimental Psychology: Animal Behavior Processes*. 1976; 2:313–322.
- Esber GR, Haselgrove M. Reconciling the influence of predictiveness and uncertainty on stimulus salience: a model of attention in associative learning. *Proceedings of the Royal Society B Biological Sciences*. 2011; 278:2553–2561. [PubMed: 21653585]
- Esber GR, Holland PC. Basolateral amygdala is necessary for negative prediction errors to enhance cue salience, but not to produce conditioned inhibition. *European Journal of Neuroscience*. 2014; 40:3328–3337. [PubMed: 25135841]
- Esber GR, Roesch MR, Bali S, Trageser J, Bissonette GB, Puche AC, Holland PC, Schoenbaum G. Attention-related Pearce-Kay-Hall signals in basolateral amygdala require the midbrain dopaminergic system. *Biological Psychiatry*. 2012; 72:1012–1019. [PubMed: 22763185]
- Esber GR, Torres-Tristani K, Holland PC. Amygdalo-striatal interaction in the enhancement of stimulus salience in associative learning. *Behavioral Neuroscience*. 2015; 129:87–95. [PubMed: 25730120]
- Gallagher, M.; Holland, PC. Understanding the function of central nucleus: Is simple conditioning enough?. In: Aggleton, J., editor. *The Amygdala: Neurobiological aspects of emotion, memory, and mental dysfunction*. Wiley-Liss, Inc.; New York: 1992. p. 307-321.
- Gottlieb J, Hayhoe M, Hikosaka O, Rangel A. Attention, reward, and information seeking. *Journal of Neuroscience*. 2014; 34:15497–15504. [PubMed: 25392517]
- Hall G, Pearce JM. Latent inhibition of a CS during CS-US pairings. *Journal of Experimental Psychology: Animal Behavior Processes*. 1979; 5:31–42. [PubMed: 528877]
- Han J-S, Gallagher M, Holland PC. Hippocampal lesions disrupt decrements but not increments in conditioned stimulus processing. *Journal of Neuroscience*. 1995; 15:7323–7329. [PubMed: 7472486]
- Han J-S, Holland PC, Gallagher M. Disconnection of the amygdala central nucleus and substantia innominata/nucleus basalis disrupts increments in conditioned stimulus processing in rats. *Behavioral Neuroscience*. 1999; 113:143–151. [PubMed: 10197914]
- Haney RZ, Calu DJ, Takahashi YK, Hughes BW, Schoenbaum G. Inactivation of the central but not the basolateral nucleus of the amygdala disrupts learning in response to overexpectation of reward. *Journal of Neuroscience*. 2010; 30:2911–2917. [PubMed: 20181588]
- Holland PC. Excitation and inhibition in unblocking. *Journal of Experimental Psychology: Animal Behavior Processes*. 1988; 14:261–279. [PubMed: 3404081]
- Holland PC. Amount of training effects in representation-mediated food aversion learning: No evidence for a role for associability changes. *Learning & Behavior*. 2005; 33:464–478. [PubMed: 16573217]

- Holland PC. Enhanced conditioning produced by surprising increases in reinforcer value are unaffected by lesions of the amygdala central nucleus. *Neurobiology of Learning and Memory*. 2006; 85:30–35. [PubMed: 16102981]
- Holland PC. Role of amygdala central nucleus in feature negative discriminations. *Behavioral Neuroscience*. 2012; 126:670–680. [PubMed: 22889308]
- Holland PC. Effects of amygdala lesions on overexpectation phenomena in food cup approach and autoshaping procedures. 2016 Submitted for publication.
- Holland PC, Bashaw M, Quinn J. Amount of training and stimulus salience affects associability changes in serial conditioning. *Behavioural Processes*. 2002; 59:169–183. [PubMed: 12270519]
- Holland PC, Fox GD. Effects of hippocampal lesions in overshadowing and blocking procedures. *Behavioral Neuroscience*. 2003; 117:650–656. [PubMed: 12802893]
- Holland PC, Gallagher M. Amygdala central nucleus lesions disrupt increments, but not decrements, in conditioned stimulus processing. *Behavioral Neuroscience*. 1993; 107:246–253. (a). [PubMed: 8484890]
- Holland PC, Gallagher M. Effects of amygdala central nucleus lesions on blocking and unblocking. *Behavioral Neuroscience*. 1993; 107:235–245. (b). [PubMed: 8484889]
- Holland PC, Gallagher M. Different roles for amygdala central nucleus and substantia innominata in the surprise-induced enhancement of learning. *Journal of Neuroscience*. 2006; 26:3791–3797. 4715. [PubMed: 16597732]
- Holland PC, Hatfield T, Gallagher M. Rats with lesions of basolateral amygdala show normal increases in conditioned stimulus processing but reduced conditioned potentiation of eating. *Behavioral Neuroscience*. 2001; 115:945–950. [PubMed: 11508734]
- Holland PC, Kenmuir C. Variations in unconditioned stimulus processing in unblocking. *Journal of Experimental Psychology: Animal Behavior Processes*. 2005; 31:155–171. [PubMed: 15839773]
- Holland, PC.; Maddux, J-M. Brain systems of attention in associative learning. In: Mitchell, CJ.; LePelley, ME., editors. *Attention and learning*. Oxford University Press; Oxford: 2010. p. 305-349.
- Holland PC, Park JH, Kang G, Wheeler DS. Interactions among amygdala, midbrain, and basal forebrain with surprising reward and nonreward. 2016 in preparation.
- Holland PC, Thornton JA, Ciali L. The influence of associability changes in negative patterning and other discriminations. *Journal of Experimental Psychology: Animal Behavior Processes*. 2000; 26:462–476. [PubMed: 11056886]
- Kamin, LJ. Attention-like processes in classical conditioning. In: Jones, MR., editor. *Miami Symposium on the Prediction of Behavior: Aversive Stimulation*. University of Miami Press; Coral Gables, Florida: 1968. p. 9-32.
- Lee HJ, Gallagher M, Holland PC. The central amygdala projection to the substantia nigra reflects prediction error information in appetitive conditioning. *Learning and Memory*. 2010; 17:531–538. [PubMed: 20889725]
- Lee HJ, Youn JM, Gallagher M, Holland PC. Temporally-limited role of substantia nigra-central amygdala connections in surprise-induced enhancement of learning. *European Journal of Neuroscience*. 2008; 27:3043–3049. [PubMed: 18588542]
- Lee HJ, Youn JM, O MJ, Gallagher M, Holland PC. Role of substantia nigra-amygdala connections in surprise-induced enhancement of attention. *Journal of Neuroscience*. 2006; 26:6077–6081. [PubMed: 16738251]
- Le Pelley ME. The role of associative history in models of associative learning: a selective review and a hybrid model. *Quarterly Journal of Experimental Psychology*. 2004; 57B:193–243. [PubMed: 15204108]
- Lubow RE, Wagner M, Weiner I. The effects of compound stimulus preexposure of two elements differing in salience on the acquisition of conditioned suppression. *Animal Learning & Behavior*. 1982; 10:483–489.
- Mackintosh NJ. A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*. 1975; 82:276–298.
- Maddux, JM. The roles of rat medial prefrontal cortex subregions in attention for action and attention for learning. Johns Hopkins University; 2008. Ph.D. thesis

- Maddux JM, Holland PC. Dissociations between medial prefrontal cortical subregions in the modulation of learning and action. *Behavioral Neuroscience*. 2011; 125:383–395. [PubMed: 21517147]
- Maddux J-M, Kerfoot EC, Chatterjee S, Holland PC. Dissociation of attention in learning and action: Effects of lesions of the amygdala central nucleus, medial prefrontal cortex, and posterior parietal cortex. *Behavioral Neuroscience*. 2007; 121:63–79. [PubMed: 17324051]
- McGaugh JL. Memory- a century of consolidation. *Science*. 2000; 287:248–251. [PubMed: 10634773]
- Pearce JM, Hall G. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*. 1980; 87:532–552. [PubMed: 7443916]
- Pearce, JM.; Kaye, H.; Hall, G. Predictive accuracy and stimulus associability: Development of a model for Pavlovian learning. In: Commons, ML.; Herrnstein, RJ.; Wagner, AR., editors. *Quantitative Analyses of Behavior*. Ballinger; Cambridge, MA: 1982. p. 241-255.
- Pearce, JM.; Mackintosh, NJ. Two theories of attention: a review and a possible integration. In: Mitchell, CJ.; LePelley, ME., editors. *Attention and learning*. Oxford University Press; Oxford: 2010. p. 11-39.
- Petersen SE, Posner MI. The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*. 2012; 35:73–89.
- Reep RL, Corwin JV. Posterior parietal cortex as part of a neural network for directed attention in rats. *Neurobiology of Learning and Memory*. 2009; 91:104–113. [PubMed: 18824116]
- Rescorla RA. Summation and overexpectation with qualitatively different outcomes. *Animal Learning & Behavior*. 1999; 27:50–62.
- Rescorla RA, Holland PC. Behavioral studies of associative learning in animals. *Annual Review of Psychology*. 1982; 33:265–308.
- Rescorla, RA.; Wagner, AR. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: Black, AH.; Prokasy, WF., editors. *Classical Conditioning*. Appleton-Century-Crofts; New York: 1972. p. 64-99.
- Roesch MR, Calu DJ, Esber GR, Schoenbaum G. Neural correlates of variations in event processing during learning in basolateral amygdala. *Journal of Neuroscience*. 2010; 30:2464–2471. [PubMed: 20164330]
- Schiffino, FL. Pearce-Hall attention for learning parameter: memory substrates. Johns Hopkins University; 2015. Ph.D. thesis
- Schiffino FL, Holland PC. Consolidation of a memory that guides attention for learning depends upon amygdala central nucleus activity. 2016 Submitted for publication. (a).
- Schiffino FL, Holland PC. Perturbed activity of rat extrastriate cortex disrupts the expression, but not encoding, of a memory that guides attention for learning. 2016 Submitted for publication. (b).
- Schiffino FL, Zhou V, Holland PC. Posterior parietal cortex is critical for the encoding, consolidation, and retrieval of a memory that guides attention for learning. *European Journal of Neuroscience*. 2014; 39:640–649. (a). [PubMed: 24236913]
- Schiffino FL, Zhou V, Trageser J, Holland PC. Unit activity in rat posterior parietal cortex tracks Pearce-Hall attention-for-learning. 2014:SS27. Society for Neuroscience poster 841.05. (b).
- Schultz W, Dickinson A. Neuronal Coding of Prediction Errors. *Annual Review of Neuroscience*. 2000; 23:473–500.
- Sutton RS, Barto AG. Toward a modern theory of adaptive networks: expectations and prediction. *Psychological Review*. 1981; 88:135–170. [PubMed: 7291377]
- Wheeler DS, Holland PC. Effects of reward timing information on cue associability are mediated by amygdala central nucleus. *Behavioral Neuroscience*. 2011; 125:46–53. [PubMed: 21319887]
- Wheeler DS, Wan S, Miller A, Angeli N, Adileh B, Hu W, Holland PC. Role of lateral hypothalamus in two aspects of attention in associative learning. *European Journal of Neuroscience*. 2014; 40:2359–2377. [PubMed: 24750426]
- Wilson PN, Boumphrey P, Pearce JM. Restoration of the orienting response to a light by a change in its predictive accuracy. *Quarterly Journal of Experimental Psychology*. 1992; 44B:17–36.

Highlights

- Prediction error (the difference between received and expected events) can affect associative learning in many ways
- Events that were accompanied by prediction error in the past have higher associability (they acquire new learning more readily) than events that were good predictors of their consequences
- Amygdala and midbrain subregions may be involved in prediction error-induced computations of altered associability
- The posterior parietal cortex may be critical for the storage and retrieval of altered associability values
- Function of a range of structures including the basal forebrain, dorsal striatum, and cortical regions, is essential for the expression of altered associability in faster learning

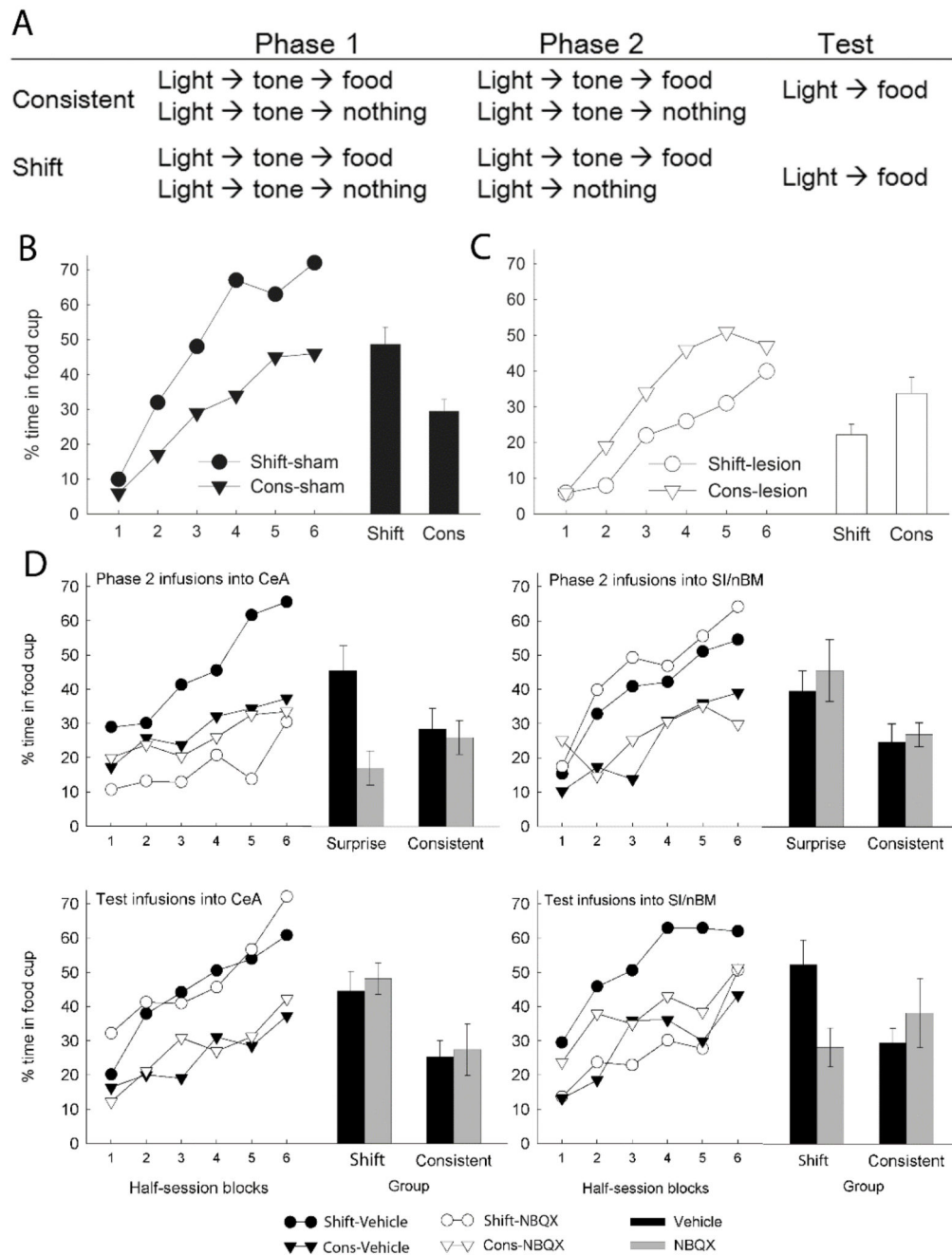


Figure 1. The serial prediction task. (A) Outline of SPT procedures. In the first phase, tone→light expectancies are established in both groups. In the second phase, a negative prediction error (PE) is induced in the Shift group by omitting the tone on half of the trials, while the Consistent group continues to receive light→tone pairings on every trial. Associability of the light is assessed in the test phase by examining its acquisition of direct light→food associations. (B) Typical SPT test result with sham-lesioned rats. Rats in the Shift group acquire conditioning to the light more rapidly than rats in the Consistent (Cons) group (a

“Shift advantage”). The bars indicate performance averaged across all sessions. (C) SPT test result with rats with lesions of the amygdala central nucleus (CeA). Rats in the Shift group acquire conditioning to the light more slowly than rats in the Consistent (Cons) group (a “Shift disadvantage”), and more slowly than sham-lesioned rats in the Shift condition (panel B). Learning of CeA-lesioned and sham-lesioned rats in the Consistent condition did not differ. (D) Effects of transient suppression of AMPA receptor activity by NBQX prior to Phase 2 or Test sessions in the SPT. The two left graphs show the effects of NBQX or vehicle infusions into CeA and the two right graphs show the effects of those infusions into the substantia innominata/nucleus basalis (SI). Shift advantages were found when only vehicle was infused, but were eliminated when NBQX was infused into CeA before Phase 2 sessions (when PE was induced) and when NBQX was infused into SI before test sessions. The data in panels B and C are from Holland and Gallagher (1993a); Panel D is adapted from Holland and Gallagher (2006).