Journal of Experimental Psychology: Animal Learning and Cognition
2015, Vol. 41, No. 2, 105–127

2329-8456/15/$12.00 http://dx.doi.org/10.1037/xan0000060

Dissociating Conscious Expectancies From Automatic Link Formation in Associative Learning: A Review on the So-Called Perruchet Effect

Pierre Perruchet
University of Bourgogne

A long-running debate in the literature on conditioning in humans focuses on the question of whether conditioned responses are the product of automatic link formation processes governed by the standard laws of simple associative learning, or the consequence of participants' inferences about the relationships between the 2 related events, E1 and E2, which would lead E1 to generate a conscious expectancy of E2. A paradigm aimed at dissociating the predictions of the 2 accounts was proposed by Perruchet (1985). In this paradigm, E2 randomly follows E1 only half of the time on average, a probability that is known to participants. When the preceding run goes from a long sequence of E1 alone to a long sequence of E1-E2 pairs, associative strength should increase, whereas conscious expectancy for E2 should decrease in keeping with the gambler's fallacy. This article reviews the studies making use of the paradigm in the classical conditioning domain, and the extension of the same logic to a few other experimental situations. Overall, overt behavior has been found to change in line with associative strength, and in opposition to conscious expectancy, attesting to an empirical dissociation of automatic and control processes within a single preparation. The paradigm, however, is endowed with a number of tricky methodological issues, which are examined each in turn. Although some of these issues call for further research, a tentative conclusion is that the effect provides evidence for automatic link formation processes, the existence of which has been recently denied in the "propositional" account of learning.

Keywords: associative learning, dissociation, conscious awareness, expectancy, propositional model

Supplemental materials: http://dx.doi.org/10.1037/xan0000060.supp

As is well known, the paradigm of classical conditioning involves two types of stimuli conventionally called the conditioned stimulus (CS) and the unconditioned stimulus (US) being sequentially paired. After repeated CS-US pairings, the CS comes to evoke a response, the conditioned response, which is generally similar to the unconditioned response. Despite the simplicity of the paradigm, the nature of the underlying processes involved in conditioning in humans remains an object of heated debate.

An appreciable part of the literature on human conditioning in the last half century has addressed the question: Can conditioned responses be acquired without participants' conscious awareness of the CS-US relationship? Available data has led several researchers to conclude that conscious awareness is necessary for conditioning to occur. A few influential reviews of the literature (Brewer, 1974; Lovibond & Shanks, 2002; Shanks & St. John, 1994) were the milestones on the path toward the view that consciousness is necessary for the establishment of conditioning, and that all the procedures aimed at providing evidence for the opposite (e.g., masking task, subliminal conditioning, conditioning in amnesiacs, or under anesthesia) have failed in their objective. This line of research has culminated in the formulation of a "propositional" model of learning (De Houwer, 2009, 2014; Mitchell, De Houwer, & Lovibond, 2009). In this framework, human participants infer the relationships between CSs and USs, which is represented in propositional or symbolic format. The production of conditioned responses is controlled by this representation, which modulate the conscious *expectancy* of the US based on presentation of the CS. To put it simply, when electrodermal responses occur to a tone after the pairing of the tone with an electric shock, it is because participants have inferred that the tone is followed by the shock based on their prior experience. This propositional knowledge leads to the conscious expectancy of the shock at the tone occurrence, which in turn generates an emotional response.

It is generally acknowledged that it would make no sense of denying overall the very existence of inferential processes, and hence the possible influence of expectancy and other cognitive factors in human learning. The current question is: Is a cognitive interpretation of conditioning sufficient? Does this interpretation replace, or only complement, a more traditional account relying on a link formation mechanism, whereby conditioned responses would emerge automatically as a function of event contingencies. In this traditional view, the frequency or amplitude of conditioned

This paper was invited by the editor.

This article was published Online First March 9, 2015.

I am especially indebted to the Editor, Ralph Miller, who invited me to write this review and provided invaluable assistance throughout the writing process. I also thank Benedicte Poulin-Charronnat for her comments on a first draft of this article. This work was supported by grants from the Centre National de la Recherche Scientifique (UMR 5022) and the University of Bourgogne.

Correspondence concerning this article should be addressed to Pierre Perruchet, Université de Bourgogne, Pôle AAFE, Esplanade Erasme, 21000, Dijon, France. E-mail: pierre.perruchet@u-bourgogne.fr

responses would vary as a function of the *strength* of the associations, which itself follows the standard laws of associative learning and conditioning known since Thorndike and Pavlov. For the advocates of a dual-process model of learning (e.g., Clark, Manns, & Squire, 2002; McLaren, Forrest, McLaren, Jones, Aitken, & Mackintosh, 2014), a full-blown account of conditioning should incorporate both propositional reasoning and link formation mechanisms.

The issue at hand is reminiscent of the old debate between Hull, Guthrie, and Thorndike on the one hand, and Tolman (1932) on the other hand, during the behaviorist era (for an overview, see Shanks, 2010). A direct filiation would run the risk of anachronism and oversimplification, but the keywords of the debate, namely "associative strength" and "expectancy," date back to this period. This review focuses on a paradigm that I proposed 30 years ago (Perruchet, 1985) as an attempt to solve the strength or expectancy controversy. This article was virtually ignored for more than 15 years. Clark, Manns, and Squire (2001) rescued the article from oblivion and since then, the number of studies using the paradigm has been continuously growing.

The present article presents first the different conditioning studies that made use of the paradigm, beginning with Perruchet (1985), then examines how the underpinning logic has been extended to other associative learning settings, mainly cued reaction time (RT) tasks. To anticipate, overt behavior has been found to depend on associative strength, and the results provide support to the action of automatic link formation processes. This outcome is now coined as the "Perruchet effect" after Weidemann, Tangen, Lovibond, and Mitchell (2009).[1] However, the paradigm is endowed with a number of tricky methodological issues, which are examined each in turn. I then end by suggesting some new directions in which the effect could be reconciled with the data on which the propositional approach is rooted, arguing that a dual-process view is not necessarily the ultimate theoretical option.

The Rationale of the Paradigm

The problem of separating accounts relying on associative strength and on expectancy is that in most cases, the two accounts generate the same predictions. The repetition of CS-US pairings potentially increases the strength of the CS-US link, but also increases the probability that a human participant discovers the relationships between stimuli and, on this basis, expects the occurrence of the US after the CS. Likewise, the repetition of CS-alone weakens the strength of the CS-US link in keeping with the law of extinction, but it is quite reasonable for a human participant to reduce expectation of the US in this condition. Both interpretations are also compatible with a huge number of much more subtle variations in procedure. The paradigm proposed in Perruchet (1985) was devised to give raise to opposite predictions. Because the effect has been studied with associative learning paradigms other than classical conditioning, the CS and the US will be designed hereafter as E1 and E2, respectively, for the sake of generalization, reflecting the fact that E1 (at least its onset) precedes E2.

The basic principle of the paradigm is the use of a random intermittent reinforcement schedule, with a reinforcement rate of 50%. This means that E1, the initially neutral event, is followed by E2, a stimulus eliciting an overt response, only half of the time on average. The participants are fully informed of this characteristic at the start of the session. Given (pseudo) randomization, the whole sequence comprises runs of E1-alone and runs of E1-E2 pairings of various lengths (note that a *run* is defined here as a sequence of consecutive trials of the *same type*). Let us consider how expectancy for E2 changes as a function of the preceding run of trials. Given that E2s occurs after half of the E1s on average, it seems natural to anticipate that after a long sequence of E1s alone, the next trial will be a E1-E2 pairing, and likewise, after a long sequence of E1-E2 pairings, the next trial will be an E1-alone. In the same way, it seems natural to anticipate a gradual change between these extremes, whereby expectancy for E2 on a given trial should decrease along a monotonic gradient when the run preceding the trial goes from a long sequence of E1 alone trials to a long sequence of E1-E2 pairings as indicated on the *x*-axis of Figure 1 (to ensure a correct comprehension of this axis, which will be identical for all subsequent figures, see Table 1). The resulting linear trend is usually referred to as a manifestation of the gambler's fallacy (Burns & Corpus, 2004). I borrow this terminology below, albeit in a purely descriptive way: The extent to which fallacious reasoning is actually involved will be addressed later.

The prediction of an expectancy theory of conditioning in this paradigm is straightforward: The probability of occurrence of conditioned responses should parallel the gambler's fallacy, thereby tracing a linear downward trend when the run preceding a given trial goes from a long sequence of E1-alone to a long sequence of E1-E2 trials. What makes the paradigm unique is that the associative strength theory predicts the exact opposite. Indeed, associative strength should increase with the repetition of E1-E2 trials, in keeping with the well-documented law of acquisition across repetitions of the CS-US pairing. Likewise, following the law of extinction with the presentation of nonreinforced CS, associative strength should decrease with the repetition of E1-alone trials. The final predictions of the two theories are plotted in Figure 1, where it is apparent that random intermittent reinforcement leads to opposite effects on conscious expectancy and associative strength.[2]

Eyeblink Conditioning Studies

Perruchet's (1985) Study

The original Perruchet (1985) study used a standard eyeblink conditioning paradigm. The CS was a tone of 1,000 ms, and the US

[1] This terminology, however, does not do justice to earlier seminal studies by William Prokasy (Prokasy & Kumpfer, 1969; Williams & Prokasy, 1977).

[2] Figure 1 presents idealized linear functions. One may wonder whether linearity is a reasonable hypothesis. In particular, assuming that associative strength changes linearly as a function of the preceding run could be construed as implying either that associative strength also increases linearly throughout the learning session, or that there is no long-term change (the gain in strength generated by E1-E2 trials being immediately offset by E1-alone trials). To address this issue, the Rescorla-Wagner (Rescorla & Wagner, 1972) model was used to generate a strength value on each trial. Simulations showed that associative strength increased across the session according to a negatively accelerated function typical of the Rescorla-Wagner model. However, when plotted as a function of the type and length of the preceding run, as in the Perruchet paradigm, associative strength increased roughly linearly, at least when the model parameters were set within a plausible range. These simulations are available on the following URL: http://www.apa.org/.

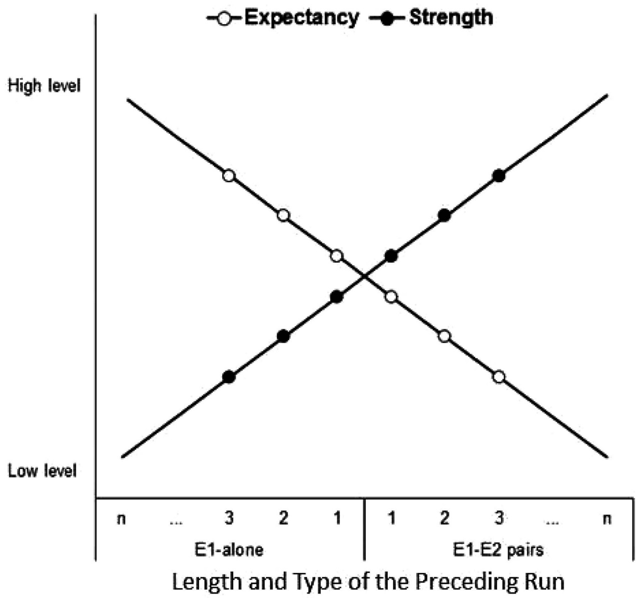

Figure 1. The length (i.e., the number of trials) and the type (i.e., E1-alone or E1-E2 pairs) of the run preceding any trial in a sequence should have opposite effects on the conscious expectancy of E2 and on the strength of the E1-E2 association. Both conscious expectancy and associative strength have the same status of hypothetical constructs, although conscious expectancy can be directly assessed through explicit expectancy ratings. The existence of automatic link formation processes varying in strength would be confirmed if overt performance follows the upward trend of the strength curve.

was a puff of nitrogen of 50 ms that occurred on half of the trials after a stimulus onset asynchrony (SOA)[3] of 950 ms (the two events coterminated). There was a mean intertrial interval (ITI) of 10 s (range: 6–14 s) between trials. CS-US pairings and CS-alone trials were pseudorandomized for each participant with a method borrowed from Nicks (1959), which has been adopted in all subsequent studies. The sequences were constructed by drawing randomly among a set of runs (and not a set of trials as usual) for which number and length were previously determined. For each type of trials (i.e., CS-US and CS-alone), there were 3 runs of four trials, 6 runs of three trials, 12 runs of two trials, and 24 runs of a single trial. The resulting sequences conformed exactly to a binomial distribution of two equally probable events, curtailed at the extremes. This means that, except after the longest runs, the mean probability of continuation and alternation was .50 after a given trial.

To check whether a gambler's fallacy occurred in these conditions, participants were instructed to rate their expectancy for the occurrence of the airpuff at the next trial during the ITI, on a 7-point scale. The results are reproduced in Figure 2. As anticipated, the overall trend was linear, with an exception after the shortest runs: Expectancy for the airpuff was stronger after a single CS-US pair than after a single CS-alone trial. I will return later on this inversion. Crucially, the pattern of conditioned responding was strictly opposite to the rated expectancy for the US. The probability of conditioned eyeblinks followed the upward function anticipated by a strength theory (see Figure 3). A decomposition of the main effect of runs in Perruchet (1985) showed that the linear trend, and only the linear trend, was statistically significant. In the terminology of Dunn and Kirsner (1988), these data provide evidence for a crossed double dissociation (also termed a "cross-over dissociation" after Shallice, 1988). Indeed, the same variable, namely the nature of the preceding run of trials, has opposite effects on two dependent variables, namely expectancy ratings and conditioned responses.

Other Delay Conditioning Studies

The results above were replicated in several eyeblink conditioning studies. When considered jointly with the experiment described above, the available data comprise 10 independent groups (Clark et al., 2001, Delay; Perruchet, 1985, Experiments 1 and 2; Weidemann et al., 2009, Experiments 1, 2, and 3; Weidemann, Broderick, Lovibond, & Mitchell, 2012, Delay 850, Delay 1250, Delay 1650).[4] To evaluate the magnitude of the effect, Figure 3 presents the probability of conditioned responses averaged over the 10 experimental groups ($N = 214$). The pooled curve appears to be somewhat less steep than in Perruchet (1985, Experiment 1). However, this comparison must not overshadow the main outcome: The magnitude of the effect on the pooled data remains impressive. The mean proportion of conditioned responses increases approximately by a factor of 0.5 (from 29% to 42%) when the preceding run goes from long (3- and 4-trial) sequences of CS-alone trials to long sequences of CS-US trials.

Does this large mean effect reflect the outcome of individual studies? It is remarkable that a statistically significant positive linear trend was reported for each of the 10 groups. However, a measure of between-groups variability is still useful. The variability of interest is related to the upward trend as a function of the preceding runs. Calculating a measure of variance on the raw values for each run length and each group would have conflated this source of variability with a second, irrelevant source, linked to the overall level of performance. Indeed, the overall level of performance may vary from one study to another as a function of a large number of parameters. To remove this second source of variability, for each group and each run length, the scores were transformed as a deviation from the mean of the group. For instance, the rate of conditioned responses in condition "4 CS-alone" in Perruchet (1985, Experiment 1) was 32.6%. As the mean score over conditions was 43.1% in this study, the score for 4 CS-alone was set to -10.5 (i.e., 32.6–43.1). The *SE*s of these difference scores, calculated over the 10 groups, are reported in Figure 3. It appears

[3] The usual acronym in the conditioning literature is ISI, which stands for "inter stimulus interval." However, in other fields of research, ISI designates the interval between the *end* of E1 and the beginning of E2. To avoid any ambiguity and given that the Perruchet effect is not limited to the conditioning area, the interval between the onset of E1 and the onset of E2 is referred to as SOA throughout the article.

[4] According to Lovibond and Shanks (2002), the effect was also replicated in an unpublished thesis (Bonic, 1989). I have not been able to gain access to this document, and as a consequence, these data were not included. Note that Prokasy and Kumpfer (1969) also reported data that, when reorganized as in Perruchet (1985), exhibited the same upward trend. However, the data were reported in a way that makes it impossible to infer whether the linear trend was significant. In addition, for some unknown reason, the mean probability of conditioned eyeblink responses (around 75%) was about twice the value observed in the more recent studies.

Table 1
The Beginning of a Training Sequence Is Given in the Upper Line, With the Corresponding Position of Each Trial on the X-Axis of Figures 1 to 11 (Except Figure 4) in the Lower Lines

Sequence	E1	E1-E2	E1-E2	E1-E2	E1	E1	E1-E2	E1	E1	E1	E1-E2	E1	
E1-alone		1				1	2		1	2	3		1
E1-E2 pairs			1	2	3			1				1	

Note. The values designate the number of trials composing the runs of E1 events (on the left hand of the figures) or the runs of E1-E2 pairs (on the right hand of the figures). The key point is that a given trial is defined by the run preceding it, irrespective of whether the current trial is E1-alone or E1-E2 pair.

that the error bars are quite small, reflecting the low variability of the upward trend across groups. However, a closer scrutiny suggests that, with the exception of the condition 4 CS-alone, the *SE*s tend to be larger for long runs than for short runs. This may reflect a genuine difference in slopes. However, it may also reflect the fact that the number of available data points for each study necessarily decreases as the length of the runs increased, generating more variability for the longer runs. As an aside, given that the longer runs are also the more informative in the logic of the procedure, the growing variability of the data with run length is undoubtedly a negative aspect of this experimental strategy.

The studies included in this survey were not intended to be exact replications of Perruchet (1985, Experiment 1). On the contrary, they introduced certain variations to examine whether the effect was impacted by the selection of parameters. Researchers' interest essentially focused on the SOA, which is known to be a crucial variable in conditioning studies. The SOA was set to 1,250 ms in Clark et al. (2001). Weidemann et al. (2012) compared SOAs of 850 ms, 1,250 ms, and 1,650 ms, and observed no difference with regard to the effect of concern. In these studies, the SOA was fixed for a given participant. Weidemann et al. (2009, Experiment 3) investigated the effect of SOA that varied from trial to trial between 400 ms and 1,200 ms for each participant. The linear trend across runs of CS-alone and CS-US trials was not statistically different from the one observed with a fixed SOA of 800 ms. Between-experiment comparisons also suggest that the effect does not depend on whether participants are instructed to rate their expectancy during the ITI (Clark et al., 2001; Perruchet, 1985, Experiment 1; Weidemann et al., 2012) or not (Perruchet, 1985, Experiment 2; Weidemann et al., 2009).

Trace Conditioning

All of the studies included in Figure 3 used a standard delay paradigm, in which the CS and the US overlapped in time, with the US occurring at the end of the CS and coterminating with it, as in

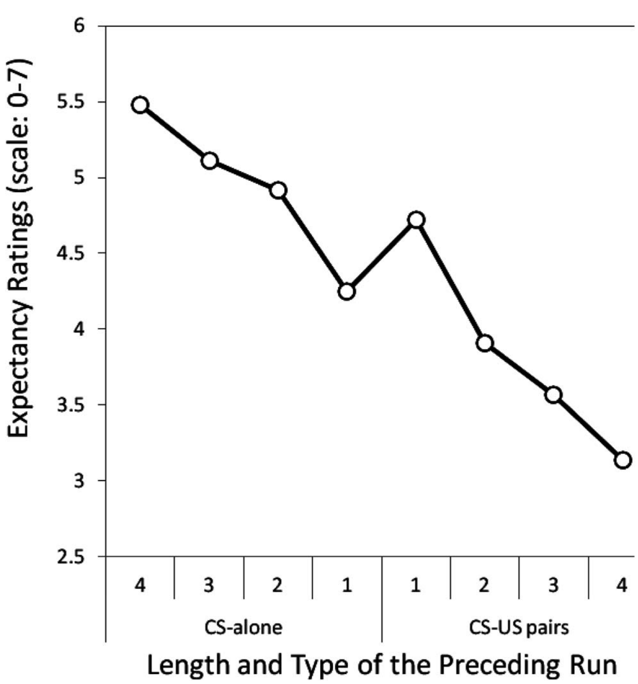

Figure 2. Mean subjective expectancy for the US as a function of the length (1 to 4 trials) and type (CS-alone/CS-US pairing) of the preceding run in Perruchet (1985, Experiment 1).

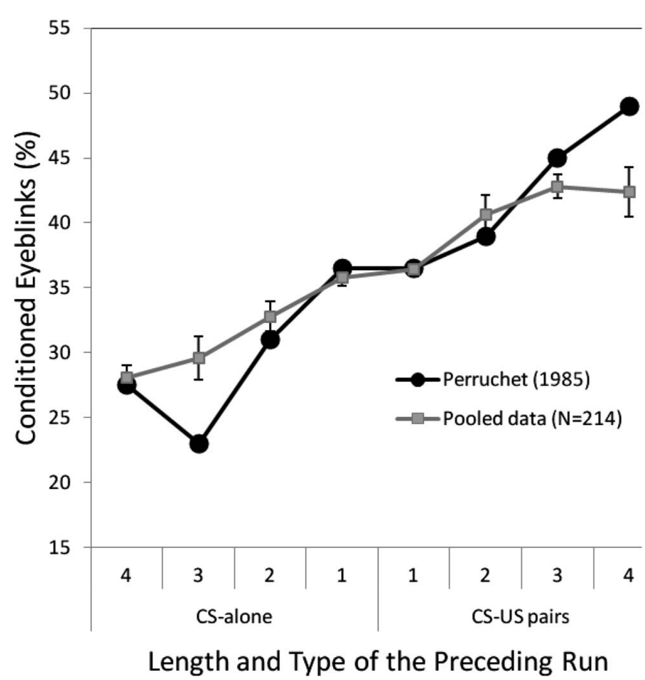

Figure 3. Percentage of conditioned eyeblinks as a function of the length and type of the preceding run. The figure shows the original results from Perruchet (1985, Experiment1), and the scores averaged over all similar studies (Delay eyeblink conditioning; 10 independent groups, $N = 214$). The error bars are *SE*s over the groups, after the effects of the between-groups variability in the overall level of performance are removed (see text for the details).

most studies on human classical conditioning. An alternative is trace conditioning, in which there is a temporal gap between termination of the CS and onset of the US. Clark et al. (2001) reported data showing the standard Perruchet effect in a delay eyeblink paradigm, as mentioned above, but the opposite was observed in a trace eyeblink paradigm. In the trace group, conditioning performance paralleled participants' expectancies. This difference between delay and trace paradigms makes sense if one considers that in a trace paradigm, the pathways devoted to the processing of the CS and the US are not coactivated, hence making it difficult to process the CS-US relationship in a reflexive (i.e., automatic) way. As a consequence, a declarative representation of the relationships would be required, a hypothesis that has received some independent experimental and neuropsychological support (Clark, Manns, & Squire, 2002).

However, the possibility that conditioned eyeblink responses would follow US expectancies in trace conditioning, in striking opposition to the results observed in delay paradigms, may be questioned on two points. First, a reanalysis by Shanks and Lovibond (2002) suggests that the Clark et al.'s conclusions are unreliable. Shanks and Lovibond pointed out that the differences between responses collected in delay and trace paradigms were restricted to the trials located at the endpoints of the curves, which, as noted above, are quite infrequent (in Clark et al., as in Perruchet, 1985, there were only three instances of each of the longest runs). As a consequence, only a few changed conditioned responses would have resulted in a quite different pattern. The analyses performed by Clark et al. did not take this source of variability into account. Clark et al. based their conclusions on the fact that the probability of the points of the curve being ranked in the observed order by chance reached (although just barely) the conventional significance threshold ($p = .05$). Shanks and Lovibond reanalyzed their data with a conventional analysis of variance (ANOVA), and they found no significant interaction between delay and trace groups.

Second, subsequent studies failed to replicate the Clark et al.'s dissociation between delay and trace conditioning. In particular, Weidemann et al. (2012) carried out a larger scale experiment including six groups of participants, differing in the paradigm (delay vs. trace) and, for each paradigm, the length of the SOA (three levels, one of them reproducing the temporal arrangement used by Clark et al.) was varied. The results were straightforward: The Perruchet effect was observed in both the delay and the trace groups, and there was no evidence of an effect of the SOAs. The absence of a difference between delay and trace preparations was also reported by Destrebecqz, Perruchet, Cleeremans, Laureys, Maquet, and Peigneux (2010) from a cued serial RT paradigm (see below). To conclude, the data reported by Clark et al. cannot be ignored, and invite us to be cautious with regard to the occurrence of the Perruchet effect in trace conditioning. However, the current weight of the evidence that trace conditioning provides an exception is rather weak.

Overall, this survey allows the conclusion that the Perruchet effect is remarkably robust. However, up to now this conclusion holds insofar as the original eyeblink conditioning procedure is concerned and it is of critical importance to examine whether the effect occurs in other associative learning settings. To my knowledge, no other motor reflex preparations have been used. This is unsurprising, given the focus of many human motor conditioning studies on eyeblink responses, which are easy to elicit and to capture. The tacit postulate is that any conclusions obtained from eyeblink conditioning settings can be safely generalized to other reflexes. However, other experimental paradigms have been explored.

How General Is the Effect?

Fear Conditioning

In fear conditioning paradigms, the US is an aversive event, such as an electric shock or a loud noise. The possibility that the Perruchet effect generalizes to fear conditioning is all but obvious. Indeed, the parameters are usually very different, notably regarding the timing of the events, which is considerably slowed down for various reasons. Moreover, on intuitive grounds, fear-elicited responses seem to be more dependent on subjective expectancies than are motor reflexes. With regard to our concern, two types of responses have been investigated. The first is the galvanic skin response, which is the main measure of autonomic conditioning collected in humans. The second is the steady state, visually evoked response potentials, which will be described in turn.

McAndrew, Jones, McLaren, and McLaren (2012) used galvanic skin responses. The CS was a tone, the US was a shock. The SOA was set to 4.5 s, and the ITI varied within a 30–40 s range. These values are much longer than in an eyeblink conditioning paradigms. To keep the experimental session within a reasonable duration, the number of trials was reduced, so that the longest runs were limited to three trials. The results presented a mixed picture. On the one hand, when the data were analyzed within the runs of a given type, there was a clear dissociation between the amplitude of the conditioned response and the ratings of expectancy. This dissociation is illustrated in Figure 4, where the run lengths were collapsed as if the left and the right halves of the *x*-values listed on the other figures (and in Table 1) were shifted to coincide with each other. This is consistent with the effect observed in eyeblink conditioning. However, there was a drop in the amplitude of conditioned responses from runs of E1-alone to runs of E1-E2 pairings, which hampered the appearance of an overall ascending effect. There were no significant differences between trials after negative and positive runs. This means that, when all the conditions are considered, there was evidence neither for the monotonic increase from the long runs of CS-alone to the long runs of CS-US that is predicted by associative strength theory, nor for the monotonic decrease predicted by expectancy theory.

McAndrew et al. (2012) briefly mention in their discussion that an earlier study by Williams and Prokasy (1977) reported different results. Indeed, Williams and Prokasy concluded that the probability of galvanic responses in a partial reinforcement paradigm "decreases across sequences of successively reinforced trials and increased across sequences of successively nonreinforced trials" (their abstract). However, they did not analyze their data as above. Williams and Prokasy included in their analysis what they called "run 0" trials. Run 0 trials are all the trials after a run of CS-US pairings when the analysis focuses on the runs of CS-alone, and all the trials after a run of CS-alone when the analysis focuses on the runs of CS-US pairs. In so doing, they conflate the effect of run length (on which McAndrew et al., 2012, focused) and the effect of whether the preceding run, whatever its length, was reinforced

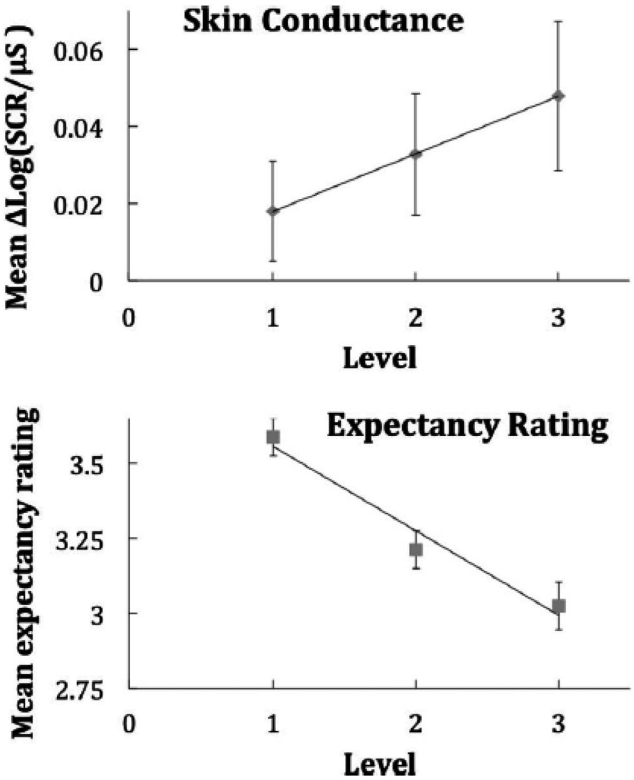

Figure 4. Amplitude of the conditioned response (top panel) and expectancy ratings (bottom panel) as a function of the preceding runs in McAndrew et al. (2012). Level 1 corresponds to the pooling of Conditions "3 E1 alone" and "1 CS-US pair" in Table 1, and likewise Level 2 corresponds to the pooling of Conditions "2 E1 alone" and "2 CS-US pair," and Level 3 corresponds to the pooling of Conditions "1 E1 alone" and "3 CS-US pair."

or not. Moreover, the experimental conditions were different from those used in recent studies. In particular, the reinforcement ratio was .33 for one group and .67 for another group, instead of .50, and the participants were (apparently) not informed of these values. Figure 5 displays the results for each group, and for two components of the conditioned galvanic responses defined by their latencies. The data, which come from Williams and Prokasy, Table 2, were rearranged as in the figures above. Overall, it is unquestionable that the dominant trend is negative, although the statistical analysis carried out by the authors do not allow the reader to confirm statistical significance. The authors note that the pattern of conditioned responses followed the level of expectancy for the US such as predicted by the gambler's fallacy, but they note also that the effect could partly reflect nonassociative aftereffects of the US. Indeed, galvanic skin responses are sensitive to habituation with repeated stimulations.

To sum up, data coming from electrodermal conditioning provide equivocal support for the Perruchet effect. There was no significant difference between negative and positive runs in McAndrew et al. (2012), and the data reported by Williams and Prokasy (1977) went descriptively in the wrong direction (note, however, that the procedure and notably the reinforcement rate, were different). The only result clearly in line with the Perruchet effect is the upward trend with run length observed by McAndrew et al. when performances were averaged over negative and positive runs.

The only fear conditioning study to investigate the Perruchet effect that did not involve electrodermal responses used evoked response potentials. Using neuroimaging methods in this objective appears a priori rather inappropriate. Indeed, as noted above, the more interesting trials are also the less frequent trials (because of the rarity of long runs), and usual neuroimaging methods require averaging the measures over a large number of trials (because of the low signal-to-noise ratio). However, an interesting exception is the use of steady state, visually evoked potentials. These potentials are natural responses to visual stimulation at specific frequencies: Oscillatory brain responses are generated at the same fundamental frequency as the visual stimulus. They can be recorded through magnetoencephalography (MEG) with a high signal-to noise ratio, even for a single trial (Keil, Smith, Wangelin, Sabatinelli, Bradley, & Lang, 2008). The rationale for using this method in the study of conditioning is that the activation of the primary sensory cortex that is related to the CS modality increases during a fear conditioning procedure (e.g., Knight, Cheng, Smith, Stein, & Helmstetter, 2004), presumably to ensure an efficient detection and processing of the fear-eliciting stimulus.

Moratti and Keil (2009) used this method to examine how the activation in visual cortex elicited by a visual CS evolves as a function of the preceding run of trials, following the conditions

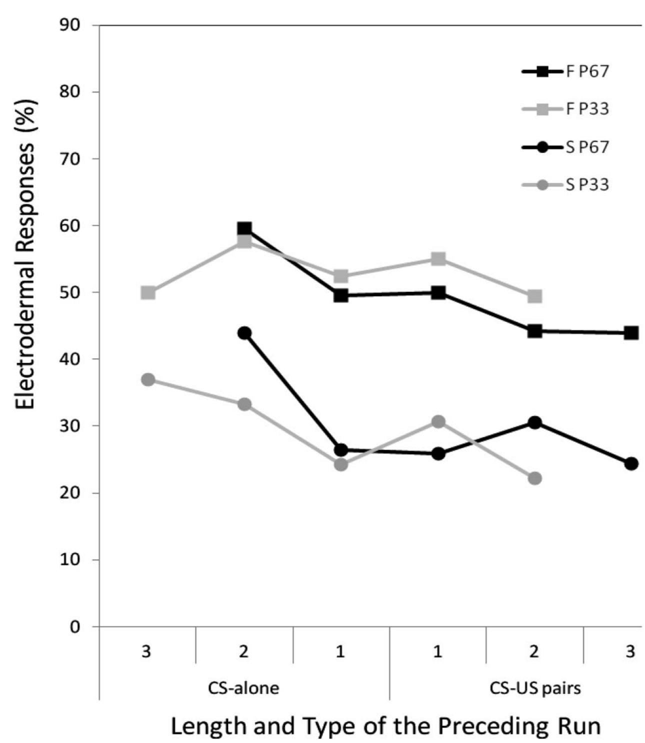

Figure 5. Percentage of conditioned electrodermal responses in Williams and Prokasy (1977, data rearranged from their Table 2). The data come from two groups of participants differing in the reinforcement rate (p = .33 and .67), and for each group, two components of responses were collected (First [F] and Second [S]). The dominant negative slope runs counter to the Perruchet effect.

listed on the *x*-axis of the figures above. The CS was displayed with a sinusoidal luminance variation at 12.5 Hz, the US was a loud white noise, and the SOA was 4 s. Expectancy ratings were also collected. The technical details regarding the processing of the MEG recordings are too complex to be reported here, but the final pattern is clear-cut: The magnetocortical activity evoked by the CS was positively related to associative strength and negatively related to subjective expectancies (see Figure 6). This study provides a nice illustration of the associative strength-expectancy dissociation with a neuroimaging technique.

Simple RT Tasks

It has long been known that when the imperative stimulus (E2) of an RT paradigm is preceded by a warning signal (E1), RTs to the imperative stimulus are decreased. The situation obviously differs from the standard conditioning paradigms described above in a number of respects. In particular, the response to E2 is a voluntary response instead of an "unconditioned" response, and training does not lead to the emergence of a "conditioned" response to E1 that would be similar to the response to E2 (save for occasional anticipatory responses). However, the shortening of the voluntary responses to E2 appears to be a consequence of learning the E1-E2 association, in the same way that the occurrence of a conditioned response does, and the phenomenon has been sometimes construed as a process of conditioning (e.g., Los, Knol, & Boers, 2001). Decreased RTs resulting from the presence of a preparatory signal is commonly attributed to the expectancy of E2.

What happens if E1 is randomly followed by E2 on half of the trials, as in a conditioning procedure of partial reinforcement? Perruchet, Cleeremans, and Destrebecqz (2006) designed experiments to address this issue, following the logic described above. The analysis of the results was performed as for the eyelid conditioning data, except that the relevant responses are now collected on E2 and not on E1. This difference entails that responses are available on only half of the trials (the E1-E2 pairings). However, this restriction affects equally all the points from the figures, given that the labels on the *x*-axes are defined by the run *preceding* the current trial, irrespective of whether E2 occurs on the current trial. As a consequence, the resulting curve may be interpreted in the same way as previously, with the obvious difference that, because the dependent variable is a latency, a better performance is now indexed by a lower score. Overall, RTs were consistent with the results obtained in the conditioning studies, with significant downward trends (Experiments 1, 3, and 4), whereas the expectancies followed the gambler's fallacy (Experiments 1 and 2).

The RT results were replicated in subsequent studies. Figure 7 shows the mean RTs collected in 12 independent groups (N = 321), coming from Perruchet et al. (2006, Experiments 1, 3, and 4, experimental group) and all the subsequent studies using a similar procedure (Barrett & Livesey, 2010, Experiment 1 single response; Destrebecqz et al., 2010, Experiments 1, 2, and 3, Delay groups; Livesey & Costa, 2014, Experiment 1; Mitchell et al., 2010, Experiments 1, 2, and 3, Experimental groups). The magnitude of the effect, assessed as a difference between the extreme runs (after pooling the 3-trial and 4-trial runs) as in the eyelid conditioning studies, was 27.6 ms.

The error bars in Figure 7 represent the variability of the downward trend over the groups, assessed as for eyeblink conditioning data by the *SE*s of the scores, after a transformation devised to remove the differences attributed to the overall level of performance. The origin of this variability is unclear. The analysis above compiles experiments in which expectancies were measured concurrently with RTs, and experiments in which expectancies were measured in a separate group or experimental phase (or not collected at all). When these experiments are teased apart, the magnitude of the effect is numerically smaller when measures are concurrent (23.8 ms) than separate (30.3 ms), suggesting that collecting RTs and expectancies concurrently is detrimental to the expression of the effect. However, this conclusion, which is drawn from between-experiment comparisons, is not confirmed by the

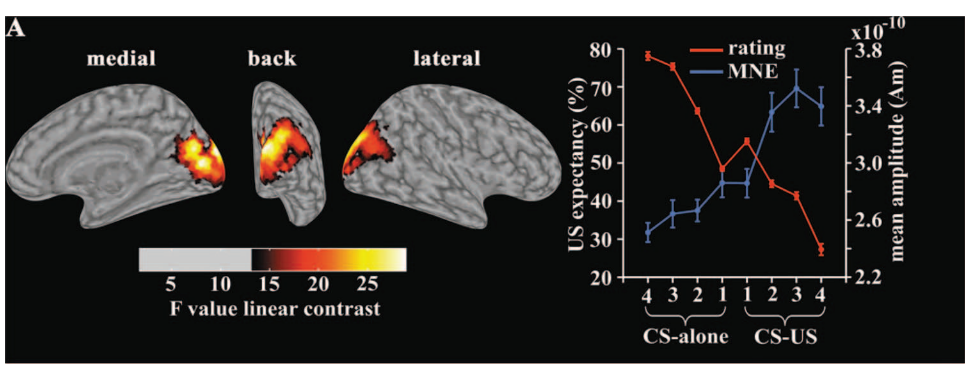

Figure 6. The left panel shows the visual cortex enhancement as a function of experience. Significant linear contrast F values are shown, representing linear enhancement of ssVEF amplitude during the first 2 s of viewing the CS, as it increased with the number of previously experienced CS-US pairs. F values are mapped onto the smoothed MNI brain. The right panel depicts the mean source strength across the occipital dipole cluster shown in the left panel (upward blue line). The (downward) red line in the right panel shows the US expectancy ratings of the subjects in the same conditions. Error bars indicate *SE*s. See the online article for the color version of this figure (reprinted from Moratti, S., & Keil, A., "Not what you expect: Experience but not expectancy predicts conditioned responses in human visual and supplementary cortex," *Cerebral Cortex* (2009),*19*(12), 2803–2809, by permission of Oxford University Press).

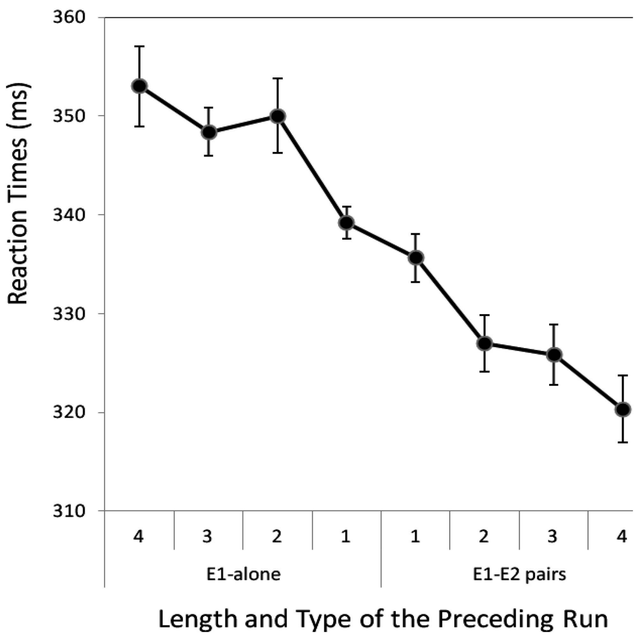

Figure 7. RTs as a function of the length and type of the preceding run. The scores in the available simple RT tasks have been averaged over 12 independent groups, $N = 321$. The error bars represent the *SE*s of the transformed scores; they reflect the between-groups variability.

single study in which concurrent and separate modes of data collection were directly compared (Livesey & Costa, 2014, Experiment 1). In this experiment, RTs were significantly slower under concurrent measurement, but there was no interaction with the downward linear trends, which were significant under both concurrent and separate conditions.

In all the experiments considered above, E2 overlapped with E1, as in a delay paradigm of conditioning. As noted in the section on eyeblink conditioning, Clark et al. (2001) obtained results opposite to the Perruchet effect in a trace paradigm, whereas Weidemann et al. (2012) observed the effect with both trace and delay paradigms. Destrebecqz et al. (2010) compared the effects observed in a cued RT task according to whether E1 overlapped with E2 (delay) or terminated before the onset of E2 (trace). In three experiments, they showed the very same results: The Perruchet effect was observed in all cases, without any interaction with the delay or trace condition. The weight of evidence suggests that the Perruchet effect does not depend on whether E1 overlaps with E2, and confirms that the counterevidence provided by Clark et al. was questionable, as claimed by Shanks and Lovibond (2002).

Choice RT Tasks

An interesting variation of the paradigm above was explored by Barrett and Livesey (2010, Experiment 1), Bertels and Destrebecqz (2013); Destrebecqz et al. (2010, Experiment 4), and with additional procedural modifications, Livesey and Costa (2014, Experiment 2). In the standard Perruchet paradigm, E1 is followed by E2 half of the time, and displayed alone in the remaining trials. In the variation examined here, E1 is followed by a given E2 (hereafter: $E2_1$) half of the times, and followed by another E2 (hereafter: $E2_2$) in the remaining trials, with $E2_1$ and $E2_2$ calling for different responses. Any trial is either an *alternation* or a *repetition* with regard to the preceding trials. An alternation is conceptually similar to E1-alone trials. For instance, in Figures 8 and 9, the reported RTs for a run of three trials in the left-hand side (Different) are the RTs obtained after a run of three $E2_1$ when the current trial is $E2_2$, or alternatively, after a run of three $E2_2$ when the current trial is $E2_1$. Likewise, a repetition is conceptually similar to E1-E2 trials: the RTs for a run of three trials in the right-hand side (Same) are the RTs obtained after a run of three $E2_1$ if the current trial is $E2_1$, or alternatively, after a run of three $E2_2$ if the current trial is $E2_2$. The reported level of expectancy for each point of the *x*-axis indicates the expectancy for the event that was actually presented (e.g., a low score means a greater expectancy for the other event). The predictions regarding both RTs and expectancies are the same as in the simple RT paradigms above.

The results, however, are more complex because they depend on whether expectancies and RTs are measured concurrently or in separate groups or phases. When the measures were taken separately (Barrett & Livesey, 2010; Bertels & Destrebecqz, 2013, Experiment 1; Livesey & Costa, 2014, Experiment 2, Condition Separate), the usual dissociation was observed. As an illustration,

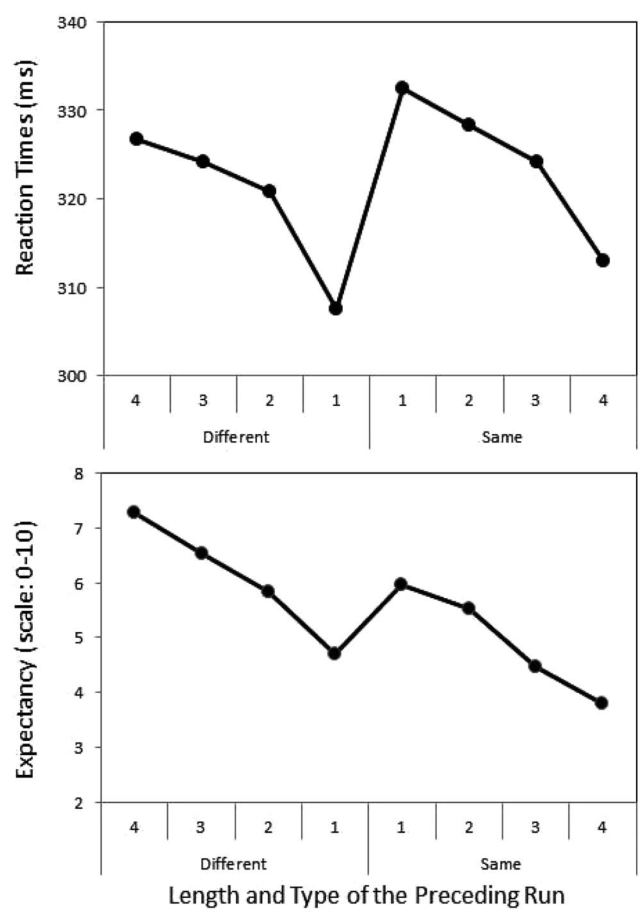

Figure 8. RTs and expectancy as a function of the length and nature of the preceding run in a choice RT task, with separate measurement of expectancies (Barrett & Livesey, 2010, Experiment 1).

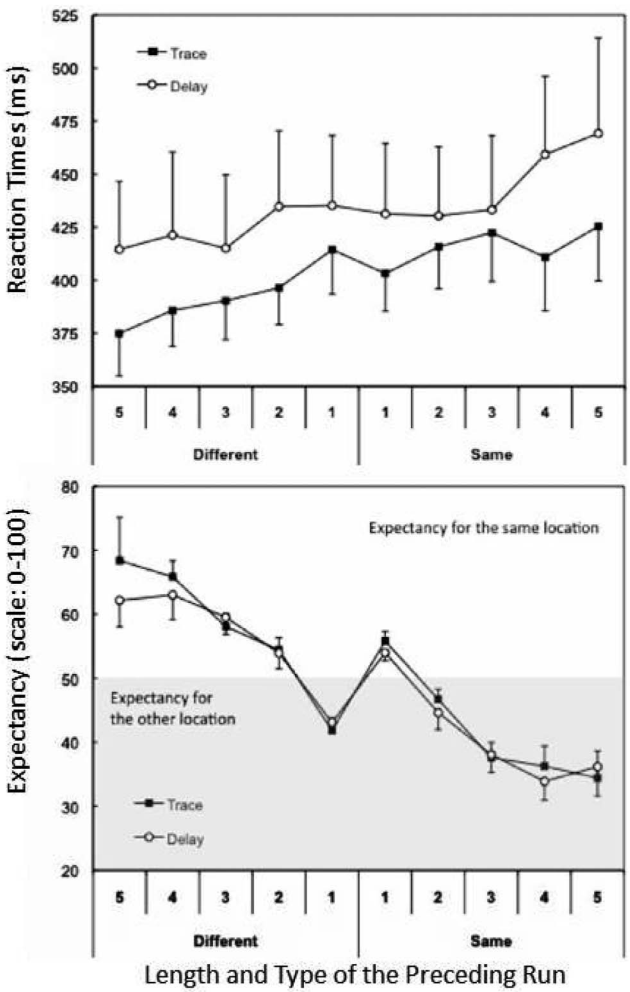

Figure 9. RTs and expectancy as a function of the length and nature of the preceding run in a choice RT task, with concurrent measurement of expectancies. Error bars indicate *SE* of the mean. See Experiment 4 from "The influence of temporal factors on automatic priming and conscious expectancy in a simple reaction time task," Destrebecqz, A., Perruchet, P., Cleeremans, A., Laureys, S., Maquet, P., & Peigneux, P., *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 2010, adapted by permission of the publisher (Taylor & Francis Ltd, http://www.tandfonline.com).

Figure 8 shows the results of Barrett and Livesey (Experiment 1), in which RTs were collected in the first two blocks of the experiment, and expectancies were measured in the last, third block. For both RT and expectancies, there was a strong first-order effect, which is an inversion in the curves around the shortest runs, to which we will return in the next section. However, when performances are considered within each type of runs, there was a striking dissociation between RTs and expectancies: RTs for a given E2 significantly improved whereas expectancies for this E2 decreased as a function of run length. Everything happens as if the preceding run strengthened the link between E1 and a specific E2, while concurrently extinguishing or interfering with the link between E1 and the other E2.

The same dissociation, however, is no longer observed when RTs and expectancy ratings are collected on the same trials. As shown in Figure 9, which reports the results of Destrebecqz et al. (2010, Experiment 4), conscious expectancies followed the standard gambler's fallacy in this condition, but the linear trends for RTs were now consistent with expectancies. The same pattern was obtained in two independent groups examined with a delay and a trace paradigm, respectively (see also Bertels & Destrebecqz, 2013, Experiment 2). A similar interaction resulted from a within-experiment comparison of the separate and concurrent conditions of measurement by Livesey and Costa (2014). In their Experiment 2, RTs exhibited the usual decreasing pattern when measures were separate, whereas RT trends in the concurrent condition showed some indication of reversal (although not significantly).

The fact that the usual direction of the Perruchet effect can be reversed under certain conditions is not, as such, particularly troublesome. The paradigm is designed to pit two opposite influences against each other, and the resulting effect presumably depends on a trade-off between these influences. Observing a prevalence of automatic activation, as is usually observed, does not mean that conscious expectancy has no influence of its own, but the converse is also true: A prevalence of expectancy does not imply that automatic activation has disappeared. There are certainly several reasons that may explain why the method of measurement of expectancies may reverse the pattern of results with a choice RT paradigm. This paradigm is undoubtedly more complex than the other experimental settings: There are two possible imperative stimuli, and both of them are cued by the same preparatory signal. This situation is conducive of associative (and/or response) competition, which could hinder or slow down the formation of associative links. Conversely, the effect of expectancy could be enhanced. When there is a single stimulus and a single response, participants might expect the onset of the imperative stimulus on each trial even though they actually had to respond in only half of the trials. Indeed, expecting a response signal that does not actually occur is not necessarily detrimental for performance. By contrast, in the choice RT task, whether participants expect the occurrence of the correct target or the wrong target necessarily has substantial consequences on the speed of the responses. The advantage will be manifest if the prediction is correct, but by the same token, the RTs should be considerably slowed down if the prediction is incorrect because preparation for a given response necessarily hampers the production of another response. In other words, there is a cost-benefit impact, whereby expecting the wrong E2 will have a detrimental effect on the RT to the signal that actually occurs (Jonides & Mack, 1984).

Other Paradigms

Other paradigms move increasingly away from classical conditioning paradigms and closely related experimental settings. Although their relation to associative learning may be looser than above, they are briefly mentioned here to illustrate that the rationale of the Perruchet's paradigm may be generalized to other areas of research. Moore, Middleton, Haggard, and Fletcher (2012) applied the Perruchet et al.'s (2006) strategy in a paradigm devised to measure the sense of agency. The sense of agency refers to the sense of initiating actions to influence external events. It can be assessed through retrospective reports in which participants rate the extent to which they felt their action caused the event, but also with an implicit measure called the intentional binding paradigm.

This paradigm exploits the fact that when an event is construed as being under one's own voluntary control, the initiating action and its consequence are perceived as closer in time compared with incidentally produced effects (for a review, see Moore & Obhi, 2012). In the Moore et al.'s experiment, E1 was a voluntary key press, and E2 was a tone, which occurred 250 ms later on 50% of trials at random. For the measure of intentional binding, a clock hand rotated rapidly on the computer screen, and participants had to report the position of the clock hand when they depressed the key (an implicit measure). Participants were also asked to judge the probability that their key press would cause the tone on the next trial (an explicit judgment).

The results are shown in Figure 10. Regarding the measure of action binding, there was a positive linear trend as a function of the preceding run of trials, consistent with the Perruchet effect. Regarding the explicit ratings, separate analyses for the runs of E1-alone and E1-E2 trials gave results consistent with the gambler's fallacy. However, overall, there was no significant linear trend, because of a very strong positive recency effect for the runs of one trial.

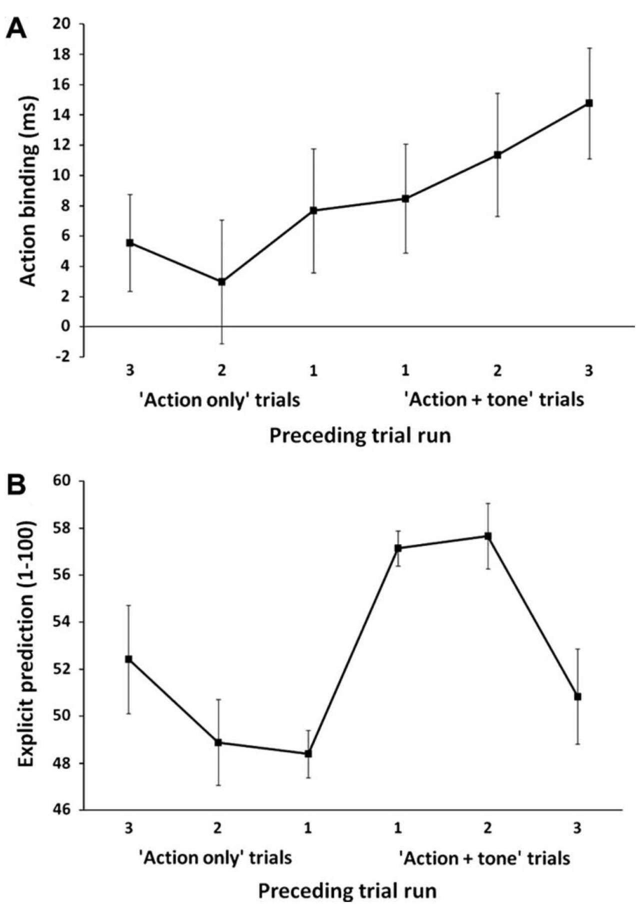

Figure 10. (A) Action binding plotted as a function of learning history. (B) Explicit predictions (1 = "definitely no tone"/100 = "definitely a tone") plotted as a function of learning history. Error bars represent *SEM*. Action binding refers to the difference in judgment error in the operant versus baseline conditions. The more positive the difference, the larger the action binding effect (from Moore et al., 2012).

Finally, Jimenez and Mendez (2013, 2014) used a variant of the Perruchet's paradigm to investigate the adaptation to conflict in a Stroop task. It is known that Stroop interference can be controlled to some extent; for instance, the amount of interference decreases when a cue informs participants about whether the next trial will be congruent or incongruent (Fernandez-Duque & Knight, 2008), but the nature of this control is still in debate. In Jimenez and Mendez's experiments, E1-alone and E1-E2 trials were replaced, respectively, by congruent trials (the color names are printed in the color ink they designate) and incongruent trials. Participants were informed that half of the trials would consist of congruent trials and the other half would be incongruent. Explicit ratings of expectancies were collected on certain blocks of trials to check that the gambler's fallacy occurs in this procedure. If conflict adaptation rests on explicit expectancy about the nature of the next trial, interference should increase when the preceding run of trials goes from a long sequence of congruent trials to a long sequence of incongruent trials. If conflict adaptation rests on a more automatic process such as envisioned in the conflict monitoring theory (Botvinick, Braver, Barch, Carter, & Cohen, 2001), the pattern should be the opposite.

The results were again clear-cut. The amount of Stroop interference reached its maximum after the longest sequence of congruent trials, even though the subjective expectancy for an incongruent trial was at its highest level, and conversely. This result was obtained when the trials occurred in immediate succession (Response Stimulus Interval, RSI = 0), a parameter that was set with the explicit aim of reducing the potential effect of expectancies (Jimenez & Mendez, 2013; see Figure 11). Worthy of note, the very same dissociation was observed when the RSI was set to 750 ms, a value that should be long enough to allow for the development of strategic operations (Jimenez & Mendez, 2014). However, the dissociation was observed only when RTs and expectancies were collected on separate blocks. In Jimenez and Mendez (2014), a change in the procedure used to collect expectancies (the use of a computer mouse in the earlier study was replaced by verbal reports) made it possible to analyze RTs on the blocks in which participants were also required to report their expectancies. In these conditions, the results were reversed, showing an association between expectancies and the Stroop effect. Overall, this pattern was very similar to the results described above regarding choice RT tasks.

To conclude this section, it appears that the empirical evidence for the Perruchet effect is stronger for certain paradigms than for others. The two paradigms that have given rise to a sizable number of studies are eyeblink conditioning and simple RT tasks. For these two paradigms, all studies have shown a large and robust effect. To the present time, studies exploring other situations are still in need of converging evidence, although the current data are suggestive of the generality of the effect over a broad range of procedures.

Open Issues and Challenges

The preceding section was silent concerning a number of methodological concerns. Although some of these concerns have been raised since 1985, most of them were put forward later by those researchers who found it difficult to encompass

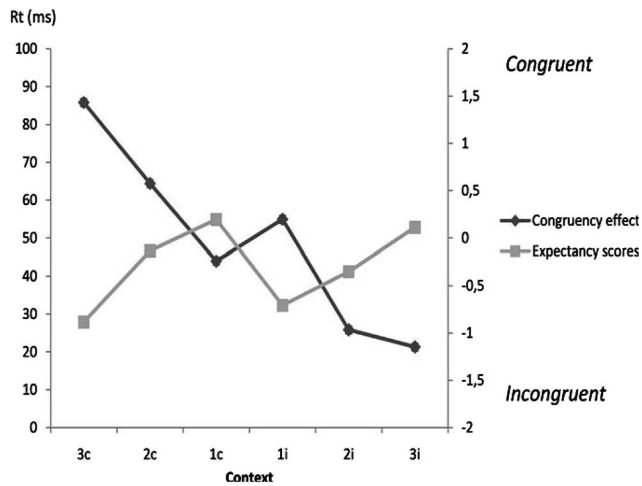

Figure 11. Dissociation between the effects of Context on the Congruency effects (represented on the left axis, as the difference in RT between responding to incongruent and to congruent trials) and on the explicit expectancies (represented on the right axis, as the average wagering on a congruent successor). The Context factor is ordered according to an increasing amount of conflict (from Jimenez & Mendez, 2013).

the results in their theoretical framework. All these issues could ultimately question the value of the paradigm and, thus, warrant careful scrutiny. The concerns regarding behavioral changes will be examined before the issues related to the measure of expectancy.

Nonassociative Processes? The Case for Classical Conditioning

One obvious objective of the paradigm is to enlighten the nature of associative learning. However, this objective could be undermined by a potential confound, which warrants examination first. Indeed, it turns out that the alternation of runs of E1-alone and E1–E2 pairings, which lies at the heart of the design, is confounded with the distribution of the temporal intervals between successive E2 events. Thus for instance, E2 has not been experienced for a long period after a long series of E1-alone trials; and conversely, E2 has been experienced several times in the recent past after a long series of E1–E2 trials. This confound could account for some of the data that run *against* the standard effect. As mentioned above, Williams and Prokasy (1977), who obtained results opposite to the usual observations with electrodermal conditioning, suggested that their results could be partly attributed to the repetitions of E2 alone. The repetition of the shocks would generate the habituation of the electrodermal responses, and would be responsible for the decreasing frequency of responses observed after long runs of E1-E2 trials. However, the question of major interest here is obviously related to the standard Perruchet effect: Is it actually the E1-E2 association that produces the usual results, as stated in the rationale of the paradigm, or the recency of the presentation of E2, with E1 being ineffective?

The nonassociative interpretations of the Perruchet effect relying on the recency of E2 differ for conditioning and RT paradigms. In eyeblink conditioning studies, it has long been shown that the repeated occurrence of a neutral stimulus and an airpuff was sufficient to elicit a blink to the neutral stimulus even though the two stimuli were never paired. The very same phenomenon occurs in fear conditioning, where the neutral event may elicit fear-related responses without any contingency between the neutral event and the fear-eliciting stimulus. This phenomenon has been called cross sensitization or more usually, pseudoconditioning (note that pseudoconditioning may be construed as conditioning to context. “Nonassociative” should be read here as “not relying on the E1-E2 association”).

The possibility that the effect of run was due to pseudoconditioning was considered by Perruchet (1985). To address this issue, Perruchet performed a second experiment, which applied the standard strategy to disentangle conditioning and pseudoconditioning, namely, running a control group trained with unpaired stimuli. More precisely, E1-E2 trials were simply replaced with E2-alone trials for half of the participants. The upward linear trend attesting to a Perruchet effect was observed only in the experimental group. Lovibond and Shanks (2002, p. 15) noted that the control in this study “is less than optimal in that the participants experienced unpredicted USs and never experienced a CS-US association.” Weidemann et al. (2009) carried out two additional experiments using a variant of the Perruchet’s (1985) control procedure to address this concern. Control participants also received runs of E2-alone, but these runs were intermixed with runs of E1-E2 trials in the same proportion as in the original study. If the linear trend were elicited by the recent exposure history of E2 and not by E1-E2 pairs, the same trend should be reproduced with E2-alone trials. This was not the case. The authors concluded that the linear trend observed in eyeblink conditioning could not be ascribed to nonassociative factors.

McAndrew, Weidemann, and McLaren (2013) reached a similar conclusion from two preliminary experiments using the electrodermal variant of the effect. There was no sign of an increasing trend in skin conductance responses in situations similar to that of McAndrew et al. (2012) except that there was no CS at all (Experiment 1) or only a few CS presentations located at strategic times during training (Experiment 2), that is, in conditions preventing the formation of associative links but normally conducive to pseudoconditioning. To my best knowledge, there is currently no evidence that the Perruchet effect using classical conditioning paradigms could be ascribed to nonassociative factors, although, for the sake of completeness, it is worth stressing that this conclusion relies exclusively on null results and therefore, should be taken with caution until confirmation by other experimental strategies, or at least, until further Bayesian analyses of the null results.

Nonassociative Processes? The Case for RT Tasks

The nonassociative accounts of the Perruchet effect are different for RT situations. Recall that in RT studies, the effect of learning is no longer measured as a response to E1, but as a modification of the response to E2. In the RT version of the Perruchet’s paradigm, the responses to E2 could be affected by the earlier presentations of E2 through a nonassociative process of sensitization. This effect of sensitization could be mediated by the transitory activation of the mental representation of E2, or alternatively by some kind of motor priming for a specific response. A beneficial effect of repeating the same imperative stimulus has indeed been reported in RT studies, although only for short RSIs. For instance, Bertelson (1961) found that the repetition effect, which was observed to occur with a 50-ms long RSI, disappeared when the RSI was increased to 500 ms in a two-choice RT tasks. Similar studies with simple RT tasks are seemingly missing, but the a priori likelihood that the effect substantially differs is weak at best. Indeed, subsequent studies have shown that a repetition effect can be observed with longer RSI, but only when the number of alternatives was *increased* (Bertelson & Renkin, 1966; Vervaeck & Boer, 1980). To reduce the possibility of a bias caused by a repetition effect, the RSIs used in Perruchet et al. (2006) and subsequent studies were in a range of values far beyond the intervals investigated in the literature on the repetition effect (with mean values equal to or larger than 6 s).

However, sensitization for a specific response is not the only version of a nonassociative account. Improved responding could be related to the increased level of arousal induced by a voluntary response. The presentation of E2 and the voluntary button press that follows could induce a temporary increase in the level of arousal, whereas participants would experience decreased arousal (and hence concentration) when the last stimulus is far back in time. This interpretation is all the more plausible as the sequence of stimuli occurs in the context of a rather boring background, which may amplify the influence of variations in task vigilance. In contrast to the sensitization of a specific response, the modulation of arousal could extend over longer periods than the 1-s scale, and even operate over different timescale.

Perruchet et al. (2006, Experiment 4) tested the influence of nonassociative factors, using a control procedure inspired by the procedure used in classical conditioning paradigm. Each experimental participant was yoked with a control participant who received exactly the same number of E2s with the same RSI. Control participants also received the same number of E1 as their experimental counterparts, except that the distribution of intervals between successive E1s was shuffled independently from presentations of E2. As a consequence, the two groups received very similar sequences of events when E1 and E2 were considered separately, but there was no contingency between E1 and E2 in the control group. The experimental group exhibited the usual linear trend with a negative slope, whereas the slope for the control group was positive (although not significantly different from zero). Additional evidence against nonassociative factors relies on the fact that if the crucial linear trend was dependent on the modulations of arousal induced by earlier E2 presentations, the linear trend should be steeper with short RSI than with long RSI, given that arousal vanishes with time. Destrebecqz et al. (2010) did not find any difference when using a mean RSI of 6.5 s and 10.5 s.

Mitchell et al. (2010), however, using a strategy similar to Perruchet et al. (2006), arrived at the opposite conclusion. They report three experiments in which an experimental group exposed to the standard paradigm was compared with a control group, which differed between experiments. In the first experiment, E1s were simply omitted throughout the session, in the second experiment, E1s were reinstated, but the E1-E2 trials were replaced by E2-E1 trials (i.e., the order of the events was reversed), and in the third experiment, E1 and E2 were unpaired, much like in Perruchet et al. (2006, Experiment 4). The standard Perruchet effect was obtained in each experimental group, but, crucially, a linear trend statistically indistinguishable from that of the experimental group was also reported for each control group. The results are especially surprising for Experiment 3, given that they contradict those of Perruchet et al., who observed no significant trend in a very similar control group. The authors do not account for the discrepancy. However, given that Mitchell et al. observed significant results whereas Perruchet et al. reported null results, the balance of the evidence undoubtedly tilts on the side of the former. These results, to quote Mitchell et al., "suggest that the linear trend in RT performance does not reflect E1-E2 (associative) learning, but rather the recency of E2 presentations. If the RT pattern does not reflect associative learning, then the dissociation between RTs and expectancy in the Perruchet effect does not constitute a challenge to an expectancy account of associative learning" (p. 371).

Mitchell et al.'s (2010) results are troublesome for the rationale of the Perruchet effect, and, to anticipate, these results are, understandably, regularly highlighted by those who do not find it easy to encompass the Perruchet effect in their framework. They could be nevertheless not so devastating as the authors suggest. The problem is that removing the predictive value of E1 (i.e., the E1-E2 contingency) in a control group may have other consequences on performance in addition to isolating the influence of E2 as intended. For instance, Barrett and Livesey (2010) noted that without a predictive E1, the context may serve as a substitutive cue for E2, with context conditioning providing an alternative associative explanation for the results. Along related lines, E1, although physically identical between experimental and control groups, subserves different functions and hence may generate different responses, which may in turn interfere with the responses to E2.

Let us examine how the latter claim could weaken the Mitchell et al. (2010)'s conclusions. Mitchell et al.'s account relies on variations in task vigilance. As noted above, this account is indeed plausible, but importantly, only insofar as E2 is the only arousal-eliciting stimulus in the task. This is obviously the case for the control group of Mitchell et al.'s (2010) Experiment 1, in which E1 has been totally withdrawn. To a lesser extent, this may also have been true of the control groups of Experiments 2 and 3, given that E1 was a tone that had no particular significance, and may even have served as a safety signal. The problem is that the arousal-eliciting value of E1 strikingly differed between control and experimental groups. For the experimental groups, E1 presumably acquired the value of a warning signal, and the Mitchell et al.'s results actually testify to the tone acquiring a high arousal-eliciting power: A rough approximation from their figures show that RTs were more than 100 ms shorter when E2 was preceded by E1 (i.e., in the experimental groups in comparison with the control groups). This suggests that the variations of arousal that were potentially responsible for the downward trend in the control groups in which E2 was the only alerting event could not account for the same trend in the experimental group in which another alerting event (E1) occurred at roughly equal intervals.

The above account has its own constraints. For instance, it needs to posit that the arousal-eliciting value acquired by E1 during training in the experimental group reaches some ceiling value, with E2 presentation and the related motor response being unable to increase participant's vigilance above and beyond this value. This assumption may seem unlikely. In addition, the proposed explanation is not particularly parsimonious, in that it involves different mechanisms producing the same downward linear trend in RTs in the control groups and the experimental groups. A final limitation is that this interpretation seems to be irrefutable. However, given the divergent results and the uncertainty of their interpretations, a conservative conclusion is that the strategy followed by Perruchet et al. (2006) and Mitchell et al. (2010) to test the nonassociative account of the Perruchet effect in RT tasks might not be the most conclusive approach.

Barrett and Livesey (2010) proposed a different strategy. Instead of creating a control situation to capture the alerting function of E2 in isolation, they modified the experimental situation in such a way that variations in arousal depending on the recency of E2 may no longer account for the pattern performance. The paradigm they proposed was the two-choice RT task, which was described in the section above. In this version of the task, one or the other of two possible E2s ($E2_1$ and $E2_2$) is present on each trial (each eliciting a motor response) and hence the modulation of vigilance induced by the recency of E2 or the related response presumably can no longer account for the trends in performance. Recall that a clear Perruchet effect was obtained in choice RT tasks, at least when expectancies were measured in a separate block of trials (Barrett & Livesey, 2010; Livesey & Costa, 2014, Experiment 2, Condition Separate). This task appears as the best way to address the problem raised by the confounds between the runs of E1-E2 and the recency of E2, at least insofar as variations of vigilance are involved. However, it could be argued that an interpretation in terms of sensitization, although rather implausible given the timing of events (see above), is not ruled out in a choice RT task, given that a sequence of E1-$E2_1$ remains confounded with a sequence of $E2_1$, and likewise for $E2_2$. Experiments 2 and 3 of Barrett and Livesey were devised to address this additional problem. Their conclusion was that some influence of the variations in the level of vigilance or sensitization cannot be definitely eliminated, but cannot be viewed as the only mechanism responsible for the Perruchet effect. Of course, the conclusions about the role of sensitization and vigilance coming from a given paradigm cannot be directly generalized to another paradigm, even though Barrett and Livesey (2010, p. 876) thought it likely that their conclusions regarding the choice RT paradigms "is also true of the original single-response paradigm, given the similarities between the two tasks."

To conclude, attempts to demonstrate the influence of nonassociative processes in classical conditioning settings have failed to do so, whereas evidence for the RT version of the paradigm remains more indecisive. A conservative conclusion is that it would be premature to definitely rule out the possibility that at least a part of the Perruchet effect is linked to nonassociative factors, especially variations in task vigilance. On the other hand, it looks unlikely that such factors would account for the whole effect.

About Expectancy Judgments

In most studies, expectancies are recorded using a graded scale displayed horizontally on the computer screen. The rating scale is typically divided into a small number of evenly spaced markers, for instance 5 (Perruchet et al., 2006, Experiment 2), 6 (Jimenez & Mendez, 2013), 7 (Perruchet, 1985), or 10 (Barrett & Livesey, 2010), and participants have to move an indicator bar by pressing one of two keys on the keyboard (usually the arrow keys). Each point of the scale is labeled, or more frequently, only the extreme values are. However, different procedures have been exploited on occasion. For instance, Weidemann et al. (2012) used a semicircular dial that was on the table in front of participants, and Jimenez and Mendez (2013) and Perruchet et al. (2006, Experiment 1) used a computer mouse and a linear potentiometer, respectively, to move an indicator bar on the scale. Purely verbal reports (collected by the experimenter) have also been used, with participants being asked to report aloud a number between 0 and 10 (Destrebecqz et al., 2010) or 0 and 100 (Moore et al., 2012). To my knowledge, there is no comparative evaluation, and the relative validity of different methods has never been assessed.

All of these methods have in common the explicit assessment of participants' expectancy. This shared principle seems to be in tune with the expectancy theory of associative learning, in which expectancy is indeed conscious, and hence communicable through symbolic tools. However, the exact meaning of a numerical assessment may be questioned. Lovibond and Shanks (2002) argued that the variations of expectancies induced by the preceding run are negligible, given there is no reliable basis for predicting whether or not the next E1 will be followed by E2. As a consequence, the "true" expectancy for E2 (the quotes are from the authors, p. 15) would remain close to 50% all along the sessions, and the conflict between opposite influences would be fictitious. The collected scores would be an exaggeration of small changes by participants to make use of the full rating scale. Weidemann et al. (2009, p. 175) and McAndrew et al. (2012, p. 207) made similar remarks.

This concern goes against a consensus about the effectiveness of the gambler's fallacy, but this is not a sufficient basis to ignore it. Addressing the point is difficult, given that separating the "true expectancy" from the measures collected on expectancy ratings seems to be out or reach of any experimental enquiry, unless alternative methods to assess the true expectancy are proposed. In the absence of alternative methods, a possible approach to determining the subjective meaning of a given interval on a rating scale would be to use a reinforcement rate differing from the conventional 50%. For instance, using a 75–25% ratio would provide an objective yardstick to use the rating scale, and the variations in expectancy specifically elicited by the preceding run could thus be assessed as a proportion of the variations elicited by the reinforcement rate. However, whether a Perruchet effect can be obtained in these conditions is an open question. We will return later on the consequences of assuming that expectancy remains almost stable across trials.

Are Changes in Expectancy Linear?

A brief scan of expectancy ratings in the figures above (Figures 2, 6, 8–11) is sufficient to reveal a break in linearity after the shortest runs. In contrast to the overall preference for alternations, expectancy for E2 is stronger after a single E1-E2 pair than after a single E1-alone trial. With the exception of Clark et al. (2001), all studies in which expectancy ratings were collected showed this effect. This is a relatively minor violation of the postulates underlying the Perruchet's paradigm such as laid down in Figure 1, but given its pervasiveness, it is important to examine whether this violation has detrimental consequences for the conventional interpretation of the data produced by the task. Perruchet (1985) and many others afterward noted that the phenomenon is akin to an effect described in the old literature on probability learning as a "positive recency effect" (e.g., Jarvik, 1951). However, the analogy is questionable because the procedures were different and, in any event, the probability learning literature is not really enlightening with respect to the psychological processes at play. I argue below that the break in the linear trend of expectancy does not attest to a genuine psychological effect, and instead could be nothing more than an artifact of the method of analysis.

To begin with, it is worth emphasizing that the observed pattern of expectancy does *not* entail that after a single guess between two equally probable alternatives, participants would prefer to repeat their first choice for their second guess. One knows from Skinner (1942) that this would be wrong. In the data he analyzed, comprising over one million responses from a radio experiment, the alternating predominated, although very slightly (52.9%). The difference between Skinner's data and the data above is that Skinner reported the first two guesses of a sequence, whereas the data above are related to the two guesses after a run, possibly long, of the alternative event. To make the point more concrete, consider the differences between two problems: What occurs after A, A or B? and what occurs after BBBA, A or B? It looks likely that the expectancy guesses will differ, with a greater proportion of A (repetition) for the second question. Of course, all the sequences preceding a 1-trial run in the Perruchet algorithm are not so extreme as in the example above, but this kind of sequence does occur. This analysis suggests that considering only the preceding *run* (i.e., a sequence of identical elements) in the Perruchet method of analysis may have misguiding consequences. Indeed, this leads to consideration of a variable number of prior trials, and in particular only one preceding trial when outcomes alternate, whereas it is likely that participants actually base their predictions on the last few preceding trials in all cases, whether the outcomes on those trials were identical or not.

The hypothesis above could be at least partly tested by comparing expectancy ratings after several types of sequences that are currently embedded into the same, one-trial run length type. The one-trial run length types are the two most heavily sampled runs, making this analysis relatively easy. For instance, the comparison could include 5-trial sequences comprising a various proportion of A and B, with the extremes instances being BBBBA and AAABA. In the present method of analysis, the predictions after these sequences are conflated. If participants took more than one element into account while being sensitive to a gamblers' fallacy, they should predict more often B in the first case and A in the second case. Such a result would indicate that participants' expectancies are governed by the same processes throughout the sequence, with the inversion around the middle of the curves being a by-product of the method of analysis.

The discussion above pertains to the most frequent outcome, in which the central inversion is restricted to the two points around the shortest runs. However, there are a few cases (e.g., Moore et al., 2012; Perruchet et al., 2006, Experiment 2) in which the central inversion is so striking that there is a substantial overlap between the left-hand and the right-hand sides of the curves, ultimately abolishing the negative linear trend over the runs. Is this pattern of expectancy still consistent with the conventional interpretation of the Perruchet effect?

This pattern of data is indeed troublesome. In Moore et al. (2012; see Figure 10), for example, participants expected a tone after two action-tones trials more than after two action-alone trials, thereby making it possible to view RTs variations as a direct consequence of this expectancy pattern. However, directly comparing the level of expectancy between distant points on the *x*-axis of the figures could be unwarranted. Interpreting a momentary level of subjective expectancy on a rating scale is certainly not an easy task, and it is likely that participants are not expressing anything except a judgment relative to the immediately preceding trial. If for whatever reason the inversion around the shortest trials is very wide, the guesses could remain affected by these initial values throughout the following run. This interpretation could be easily tested by asking an expectancy rating only on a selected subset of trials, to remove or reduce the influence of the just preceding rating.

Pending the results of further studies exploring the source of the inversion of the expectancy rating scores around the shortest runs, and sometimes beyond, a practical issue concerns the statistical processing of data. When the inversion is restricted to the two central points, the overall linear trend generally remains significant, and the descriptive break in linearity is commonly neglected. The issue is more critical when the inversion extends beyond the two central points. Perruchet et al. (2006, Experiments 3 and 4) removed the overlapping data from the analyses, and contrasted performance only for the longest, extreme runs. After all, the conventional interpretation of the Perruchet effect primarily concerns long runs; thus, this practice seems reasonable. However, a better method, which does not entail data deletion, has been used since then (e.g., Destrebecqz et al., 2010; McAndrew et al., 2012). Instead of considering the full array of preceding runs along a single dimension going from the longest runs of E1-alone to the longest runs of E1-E2 trials, the data are condensed as shown in Figure 4 above for the results of McAndrew et al. (2012). The conditions are collapsed, as if the left and the right halves of the *x*-values were shifted to coincide with each other. The mean difference between the types of runs that are collapsed (E1-alone vs. E1-E2 pair) is analyzed as a separate factor. The soundness of the procedure depends on the interpretation given to the central inversion. If one views this inversion as a deep violation of the postulates of the conventional interpretation as illustrated in Figure 1 above, a procedure masking what is construed as embarrassing data is obviously inappropriate. However, the prevalent view seems to be that, irrespective of the current mode of representation of the logic of the paradigm, a statistical analysis focusing on the effect of run length, and considering the effect of the type of runs separately, is valuable.

As a final point, it is worth adding that a central inversion similar to that observed for expectancy has been observed also for the performance data, although much less frequently. The phenomenon seems to be restricted to the RTs in the choice RT tasks (Barrett & Livesey, 2010; Livesey & Costa, 2014), as illustrated in Figure 8 above. The number of studies using this task is too small to draw firm conclusions. Nevertheless, it is interesting that the surface similarity between RTs and expectancy in Figure 8 is misleading: There is a break in linearity for both RTs and expectancy around the shortest runs, but in fact they reveal opposite trends. For instance, expectancy for $E2_1$ was higher after a single presentation of $E2_1$ than after a single presentation of $E2_2$, but RTs to $E2_1$ was faster after a single presentation of $E2_2$ than after a single presentation of $E2_1$. In other words, the dissociation between performance and expectancy revealed in the long runs stands true for the shortest runs as well. The break in linearity for choice RT tasks lends itself to the very same explanation used for expectancy, that is, as an artifact linked to an exclusive focus on the runs as units of analyses. It is indeed quite plausible that, for instance, the automatic activation for $E2_2$ is stronger after $E2_2$ $E2_2$ $E2_2$ $E2_1$ $E2_2$ than after $E2_1$ $E2_1$ $E2_1$ $E2_2$ $E2_1$, whereas the theoretical predictions depicted in Figure 1 (the relevant points are related to the shortest runs) are just the opposite.

Is the Gambler's Fallacy Actually Involved?

The changes in expectancy as a function of the nature of the preceding runs in the Perruchet's paradigm are routinely attributed to the gambler's fallacy. The gambler's fallacy (e.g., Burns & Corpus, 2004) is the mistaken propensity to believe that a just completed run of a particular outcome (e.g., heads on the toss of a coin) will be balanced by a tendency for the opposite outcome (e.g., tails) in the immediate future given a random sequence of events. Laplace (1902/1814) is credited as being the first to describe the phenomenon: "When a number in the lottery of France has not been drawn for a long time the crowd is eager to cover it with stakes. They judge since the number has not been drawn for a long time that it ought at the next drawing to be drawn in preference to others" (p. 161). A widely accepted interpretation of this belief has been proposed by Tversky and Kahneman (e.g., 1971) as an illustration of their "representativeness heuristic." According to this heuristic, people believe that short sequences of random events should be representative of longer ones, and more specifically, that deviation from the theoretical probability should balance out for the short runs as it does for the long runs. Is the gambler's fallacy really involved in the Perruchet effect?

In principle, the response is clearly negative for a simple reason: In an experimental session, short sequences are *actually* representative of longer ones. As detailed above, the specific procedure of randomization makes that the two outcomes match exactly a binomial distribution with $p = .5$. As a consequence, alternations and continuations are perfectly balanced over the whole sequence. (As an aside, Lovibond & Shanks, 2002, noted that the partial reinforcement schedule used in Perruchet, 1985, favors alternation over continuation of runs. They were right, because of truncation at the extremes of the binomial distribution: There were three runs of four trials and no longer run. This imbalance was no longer present in most subsequent studies.) However, the fact that alternations and continuations are perfectly balanced in number over the whole sequence does not mean that the probability of alternations and continuations is $p = .5$ at all points during the experimental session. In fact, $p = .5$ is true for the very first trial, but is generally incorrect afterward. As Laplace (1902/1814) noted after his description of the gambler's fallacy, "the extraction of a white ball from an urn which contains a limited number of white balls and of black balls increases the probability of extracting a black ball at the following drawing" (p. 162). Likewise, the probability of receiving an E1-E2 trial *actually* increases across a run of E1-alone, and conversely, decreases across a run of E1-E2 pairs. This does not imply that after a run of E1-alone, the probability of an E1-E2 trial is necessarily higher than the probability of repeating again E1-alone and conversely after a run of E1-E2 trials. The probability of the next event depends on the composition of *all* the preceding trials. The point, which does not seem to have been noticed in the literature on the Perruchet effect, is that this probability, whatever it is, fluctuates along a run in a direction consistent with the gambler's fallacy.

To recast the preceding argument in more technical words: In the Perruchet's paradigm, sampling is without replacement, whereas the gambler's fallacy applies only when successive events are truly random, as in sampling with replacement. Note that this does not result from methodological carelessness or inadvertence. This mode of sampling was chosen to deliver exactly the expected number of trials of each condition to each participant, an objective that appears especially suitable in the present case to ensure the presence of long runs in the sequence for each participant.

That being made clear, whether or not participants follow a faulty reasoning in their expectation of alternations does not depend on the objective structure of the materials, but on their representation of this structure. This raises a thorny issue. The exact instructions given to participants are not always reported, but it seems that participants are routinely told that E2 will follow E1 on only half of the trials. However, these instructions may be understood in two different ways. Participants may understand that on each trial, there is exactly a 50% chance to have E2 after E1, or alternatively, that over the whole sequence they will be exposed to exactly the same number of E1-alone and E1-E2 trials. The same trend in expectancy may be described as a gambler's fallacy in the former case, and underpinned by rational reasoning in the latter case. It is likely that the truth lies in-between, depending on participants and experimental parameters including instructions. However, it is likely that most participants guess that the succession of events in any experimental setting is not truly random. Thus, unreservedly ascribing their pattern of expectancy to the gambler's fallacy is unwarranted (nevertheless referring to the gambler's fallacy in a purely descriptive way remains a convenient shortcut; therefore, for internal consistency, this terminology will be used hereafter).

The issue of rationality is endowed with deep implications for the theoretical scope of the Perruchet effect. The broad objective of the paradigm is to pit conscious thoughts and automatic activation against each other. Conscious thoughts are commonly viewed as the site of logic and rationality, and the fact that expectancies here appear to be the product of fallacious reasoning may be viewed as a genuine limitation of the paradigm. It is conceivable that any argument aimed at showing that the observed pattern of expectancy has some rational grounds would make the Perruchet effect more compelling. In this respect, one objective of further research should be to vary the extent to which reported expectancies reflect rational choice. This could be done by manipulating the cover story and the instructions in the standard conditions, and/or by changing the experimental conditions. For instance, breaking the whole session into small subsets of trials each comprising the same number of E1-alone and E1-E2 pairs would provide some objective bases for predictions, making expectancies potentially more rational. Indeed, the objective probability of the different events changes more across a run when the total number of trials is small, and the probabilistic structure should be more easily perceived by the learner. If the Perruchet effect vanishes in such conditions, this would be indicative that the influence of expectancy in the standard paradigm could be limited by the fact that expectancies are the end-result of a fallacious reasoning, a condition that may be not representative of conscious thought in real-world settings. By contrast, observing a robust effect in conditions where the pattern of expectancy is grounded, at least partially, on correct reasoning, would strengthen the dissociation.

Are Anticipated Expectancy Ratings Valid?

In the logic of the paradigm depicted in Figure 1, a given E1 elicits both some amount of associative activation of E2 and some amount of conscious expectancy for E2. Given these two variables presumably act in opposition, the resulting effect on performance should reveal the winner. Everyone would certainly agree that for this logic to apply, the ideal condition would be that performance and reported conscious expectancy are captured exactly at the same time in the same participants. It is obvious that this condition cannot be fulfilled because of the possibility of mutual interference. Two main solutions have been exploited. Expectancy ratings are collected either before the possible occurrence of E2, or in separate blocks than performance (or even in another group of participants). Each method is endowed with specific advantages and drawbacks. This subsection deals with the problem of collecting expectancy ratings predating E2 occurrence.

In Perruchet et al. (2006), right-handed participants were instructed to continuously update their rating throughout the session, by moving the slider of a linear potentiometer with their left hand. The current position of the slider was indicated by a cursor on a continuous scale drawn on the computer screen. Shanks (2010) noted that the actual occurrence of E1 may change the level of expectancy for E2 from its level during the preceding intertrial interval, but participants may lack the time to update their rating during the brief E1-E2 interval (500 ms). This comment is relevant, but only for the first experiment in Perruchet et al. In their second experiment, Perruchet et al. modified the procedure in such a way that the ratings of expectancy for E2 could only be made after the onset of E1, during the E1-E2 interval, which was sensibly lengthened (750 ms). The cursor on the screen was reset in the middle location of a 5-point rating scale (i.e., on Point 3) at the onset of each E1, and participants moved it to indicate their expectancy for E2 by pressing the left or right arrows on the keyboard. It could be argued that the time interval was still too short to press the arrow keys. However, this would have resulted in a flat curve, which was not the observed result. Likewise, McAndrew et al. (2012) used an SOA of 4.5 s in electrodermal conditioning, and participants were asked to press one out of five buttons to rate their expectancies for E2, again after the onset of E1. In these two studies in which the problem raised by Shanks does not apply, a significant linear trend consistent with the gambler's fallacy was reported.

The comments above could be interpreted to suggest that sound expectancy ratings are only available for two studies, a conclusion that would severely restrict the empirical basis of the Perruchet effect. In fact, whether collecting expectancy ratings before occurrence of E1 is invalid or not is a matter of debate. In most studies, in keeping with Perruchet (1985), participants were instructed to rate during the intertrial interval (i.e., before the onset of E1) their expectancy that E2 will occur *after the next E1* (and not at the present time). Dismissing this measure amounts to considering that participants' expectancy of E2 after the next E1would change radically before and after the actual onset of E1. This hypothesis is not consistent with the fact that a between experiments comparison shows that ratings made before (e.g., Clark et al., 2001; Destrebecqz et al., 2010; Livesey & Costa, 2014; Perruchet, 1985; Weidemann et al., 2012) and after (McAndrew et al., 2012; Perruchet et al., 2006) the onset of E1 exhibit the same pattern. Moreover, at a more speculative level, the psychological processes that would motivate such a change in expectancy remain to be discovered. In any case, this hypothesis is inconsistent with a propositional framework, in which expectancies are based on explicit inferences about E1-E2 relationship. Indeed, the mere occurrence of E1 does not change in any way the informational basis on which an explicit inference can be drawn. It is conceivable that the mere occurrence of E1, even though E1 occurs on each trial, is sufficient to turn instantly the conscious representation of the E1-E2 relationships upside down, but this hypothesis does not seem to be in tune with the cognitive stance of the most active critics of the Perruchet effect.

Does Associative Strength Actually Overlook Expectancy?

A way to measure expectancy and performance in perfect synchrony while precluding potential interference consists of measuring these two components in separate blocks of trials or in different groups of participants. The underlying assumption is that independent ratings of expectancy provide a reliable picture of the expectancy reached when only performance is measured. However, a finding reported above casts doubt on this reasoning. In Bertels and Destrebecqz (2013, Experiment 2), Destrebecqz et al. (2010, Experiment 4), Jimenez and Mendez (2013), and Livesey and Costa (2014), performance reflected the theoretical variations in associative strength only when there was no concurrent measure of expectancy. When expectancy was assessed concurrently, there was no effect, or even a reverse effect, with performance being reversed to match the pattern of expectancy ratings.

The stronger influence of expectancy on performance when expectancy is assessed in the same phase as performance relative to when the measures are separately recorded may have a simple explanation. Indeed, participants may prepare a response that is consistent with their prediction if they just made one. Thus, the concurrent measurement of expectancy may serve as a strong motive for cognitive control of performance, and to align one's performance with one's conscious representations. This interpretation suggests that measuring expectancy and performance in separate phases or different groups of participants would be the best option to obtain an unbiased estimate of performance, and overall, the rational of the Perruchet effect remains intact. However, there is another interpretation, which is more damaging for the logic of the paradigm. Possibly explicit expectancies are not built by default, but only upon request (Jimenez & Mendez, 2013). In this view, the effect of associative strength would be manifest under separate measurement because there is no expectancy at all when an explicit rating of expectancy is not requested. However, as soon as expectancies are generated under the pressure of task demand, they would counterbalance or overpower associative strength. The question is how to know whether expectancy follows the gambler's fallacy without collecting any confirmatory evidence. This issue raises a conundrum in some ways analog to the Observer effect in physics, whereby, for instance, the path of an electron is altered by its interaction with some other particle or force that is required for its detection.

To reduce the impact of this objection, one could argue that a difference between concurrent and separate measures of expectancy observed in the articles cited above is not a common finding. These differences have been reported mainly for choice RT paradigms (Bertels & Destrebecqz, 2013; Destrebecqz et al., 2010; Livesey & Costa, 2014), but not for eyeblink conditioning or simple RT tasks. For instance, in Perruchet (1985), the upward trend of conditioned eyeblinks was comparable, if not steeper, in Experiment 1 in which expectancies were measured during conditioning, than in Experiment 2 in which expectancies were not collected. Moreover, most studies investigating the Perruchet effect with simple RT tasks have successfully used concurrent measurements of RTs and expectancies. Admittedly, Perruchet et al. (2006) observed a clearer decreasing effect in Experiments 3 and 4, without the online measure of expectancy, than in Experiment 1 in which the two measures were recorded concurrently, but the standard effect was nevertheless significant in Experiment 1. In addition, the single experiment implementing both concurrent and separate measurement in a simple RT paradigm reported no effect of this variable (Livesey & Costa, 2014, Experiment 1). These data could be viewed as evidence that at least in certain paradigms, there is a direct conflict between expectancy and associative strength in which associative strength is victorious in controlling performance.

There is a last objection, however, which challenges the validity of this conclusion. All the data reported up to now are averages computed over participants. A possibility is that some participants exhibit a gambler's fallacy without consequences for performance, whereas other participants are sensitive to variations of associative strength without any change in expectancy. Averaging over these participants would result in the observed dissociation, even though the postulated processes would be never pitted against each other. Barrett and Livesey (2010, Experiment 3) and in a more extensive way Livesey and Costa (2014) reported data supporting this hypothesis in RT tasks. The effect of prior runs was assessed for each participant by the slopes across run lengths for both expectancy ratings and RTs. If some participants are responsible for the RT trend and others for the expectancy trend, a negative correlation is expected between the two measures. With separate measurement, there was no significant correlation, and this was true for both simple RT tasks and two-choice RT tasks (Barrett & Livesey, 2010; Livesey & Costa, 2014). However, when expectancy ratings were collected concurrently with the RT task, the slope of the linear trends in RT and expectancy were negatively correlated, suggesting that participants exhibiting the stronger gambler's fallacy are less likely to show the common trend in RTs (Livesey & Costa, 2014). This conclusion applied to two-choice RT tasks, and also to simple RT tasks. Further analyses conducted after dividing participants into three groups according to their sensitivity to the gambler's fallacy confirmed these data. When RTs and expectancies were assessed on separate blocks, the standard downward trend in RTs was observed equally for each subgroup (and hence, even in participants exhibiting a consistent gambler's fallacy). However, when RTs and expectancies were assessed concurrently, no reliable trend in RTs was observed for participants exhibiting a consistent gambler's fallacy. An RT trend in a direction consistent with associative strength was only found for participants who did not show a consistent gambler's fallacy. Implications of this rather intricate pattern of data are examined just below.

What Remains as Unquestionable?

This section assesses the implications for the Perruchet effect of the various issues examined above, proceeding in a backward order. To begin with, it appears that a skeptical reader could argue that the literature offers no compelling demonstration that behavioral changes are actually evolving in opposition with conscious expectancies when both are measured jointly in the same participants. Evidence coming from independent measurement of performance and expectancy remains questionable because it is possible that true expectancy differs as a function of whether an explicit rating is required or not. The value of concurrent measures is also debatable because perfectly simultaneous measures is not possible; moreover, the dissociation observed on averaged data could stems from one fraction of the participants generating the expectancy curve and the other fraction the performance curve. Support for each of these possibilities is currently limited and comes mainly from choice RT tasks, but generalization to other tasks cannot be formally ruled out at this time. Detailed analyses of individual differences in tasks that do not involve voluntary responses should be a main objective of further research. An eyeblink conditioning paradigm, in which cognitive control would be a priori lower, would be a natural candidate for this objective.

Pending the outcome of future research, let us assume that for whatever reasons the existence of a genuine conflict between associative strength and expectancy in the Perruchet's paradigm cannot be conclusively demonstrated. In that case, the conclusion according to which automatic link formation processes are strong enough to counter the effects of expectancies in controlling performance would be no longer tenable. However, it is worth stressing that the Perruchet's paradigm would be any the less interesting. Indeed, it would remain that performance is sensitive to the preceding runs, and that changes in performance are going in the direction predicted by the action of automatic link formation processes (at least for some participants). To challenge the existence of automatic associations and advocate for the exclusive role of expectancy, the argument needs be that variations in expectancy are the actual causes of changes that are currently attributed to associative strength. This argument amounts to postulate that expectancies evolve in a direction *opposite* to the gambler's fallacy. In principle, this hypothesis makes sense. Gilovich, Vallone, and Tversky (1985) indeed observed what they called the "hot hand fallacy," whereby basketball players, coach, and fans tend to believe that a player's chance of hitting a shot is greater after a run of hits than after a run of misses, a belief that is not supported by statistical analyses. However, the problem is that even though the hot hand fallacy has been observed in domains other than basketball (e.g., Blanchard, Wilke, & Hayden, 2014), the expectancies collected in the literature on the Perruchet effect have always been consistent with the gambler's fallacy.

Other issues regarding expectancy are also worthy of consideration in planning further research. Notably, the question of whether expectancy curve reflects a genuine gambler's fallacy, as commonly assumed, or is based on some correct reasoning grounded on the idea that sampling is without replacement, needs to be clarified. However, whatever the finding (that may be in-between these extremes), there is no major implication for the validity of the approach.

Undoubtedly the most critical issue for the Perruchet's paradigm is whether the behavioral effect may be thought of as the consequence of nonassociative processes, such as sensitization or arousal induced by E2 (and the response to E2). The possibility of a confound between the associative effect of E1-E2 pairs and the nonassociative effect of E2 alone was assessed in Perruchet (1985) and in most of the subsequent studies, with results generally running counter to it. However, Mitchell et al. (2010) observed a trend in RT identical to the standard trend in three control groups in which there was no association between E1 and E2. I noted above that the arousal-eliciting value of E1 differs between experimental conditions in which E1 serves as a predictive signal for a biologically significant event, and in control conditions in which E1, if present, serves as a safety signal. This makes it difficult to apply to the experimental group the interpretation provided for the control groups (i.e., variations in arousal linked to the relative recency of E2). However, this account has its own weaknesses, and given the critical importance of the issue at hand, it is clear that future studies are needed to better confirm the associative nature of the processes involved in the Perruchet effect. This is a challenging objective because the nature of the nonassociative factors at play may vary from one paradigm to another, and there is certainly no "one-size-fits-all" solution to control for their diverse influences.

Theoretical Implications

The "Propositional" Model of Learning

In the propositional model of associative learning (e.g., De Houwer, 2009, 2014; Mitchell et al., 2009), human participants infer the relationships between E1 and E2 from their perceived contingency, and this relationship is consciously represented. Conditioned responses are due to the exploitation of this representation, which modulates the conscious expectancy of E2 given the presentation of E1. It is immediately evident that the Perruchet effect runs counter to a propositional view of learning. However, the effect has not been construed as sufficient to invalidate a propositional framework by the proponents of this approach, and the reasons for this are worth analyzing.

Unsurprisingly, the advocates of a propositional model have rightly pointed out that the main problem of the paradigm is the possible influence of nonassociative factors, with a pervasive reference, as empirical support, to the Mitchell et al. (2010)'s report of a Perruchet effect in control conditions preventing associative learning. They have also noted certain limits in the assessment of expectancy in the Perruchet paradigm, which have been scrutinized in the preceding section. To reiterate, this kind of argument is not compelling because to challenge the existence of automatic associations, the point should be that "true" expectancy changes in the direction opposite to the direction actually observed in participants' explicit ratings. Such a claim would hardly be tenable and in any case would be in contradiction with the Lovibond and Shanks's (2002) recommendation for using such rating scales for assessing consciousness. It is worth adding that the proponents of a single-process, expectancy-based theory of associative learning have not regarded their own criticisms of the conventional account of the Perruchet effect as leading to a definitive rebuttal. On the contrary, they have fairly acknowledged that the Perruchet effect still provides strong evidence against their view, and as an aside, they have heavily contributed to the empirical exploration of the effect.

The deep reason why the Perruchet effect is not construed as sufficient to reject a propositional model seemingly lies elsewhere than in the potential drawbacks of the procedure, as is seen in the following quotation at the end of Mitchell et al.'s (2009, p. 238) article:

> A close examination of the data reveals only one or two isolated phenomena that might indicate the presence of a nonpropositional (perhaps link-based) learning mechanism. These include the eyeblink version of the Perruchet effect (Perruchet, 1985) and the odor-taste learning work of Stevenson et al. (1998). As Dwyer et al. (Dwyer, Le Pelley, George, Haselgrove, & Honey, 2009) concede, evidence for the link-formation mechanism is not widespread. Thus, even the proponents of the dual-system approach accept that the link mechanism is of somewhat limited explanatory value. It seems to us that, if we do indeed possess two separate learning mechanisms, then we should see evidence for both mechanisms everywhere. Why, therefore, is the evidence for the second mechanism so weak and so vanishingly small? We keep an open mind, but there seems to be an obvious and almost unavoidable conclusion, that no such mechanism exists.

In other words, minimizing the implications of the Perruchet effect does not rest on the paradigm's intrinsic limitations, but rather on the paucity of converging evidence. Is this line of reasoning valid? It is worth noting first that the empirical evidence for the link formation mechanisms is not as scanty as Mitchell et al. (2009) state. A number of researchers would disagree that the Perruchet effect and a few other isolated phenomena are the sole evidence for a link formation mechanism. McLaren et al. (2014) present a defense of a dual-process account of learning which relies on a number of various empirical phenomena. Considerations regarding the level of parsimony, or the evolutionary plausibility of the respective positions are also relevant (see the commentaries in response to Mitchell et al.'s, 2009, BBS article). A full discussion of this issue is beyond the scope of this review. However, it must be realized that even if the eyeblink version of the Perruchet effect was the single piece of evidence in favor of automatic link formation, dismissing its relevance when evaluating the likelihood of a propositional approach would be premature.

Indeed, if low-level automatic associative processes do exist, then they should result in behavior that would correspond to a large extent with the optimal solution given by explicit reasoning and inference in the same situations. As McLaren et al. (2014, p. 186) wrote: "It must of course be the case that if it is to be adaptive and enable the animal to survive, in most circumstances the outcome of associative learning should parallel that to be expected from a rational propositional system." Assuming that basic learning processes established through natural selection would lead to irrational outcomes would be a contradiction in terms because we would not be here to talk about them. The range of behavioral responses that can be *exclusively* explained by automatic link formation is necessarily limited. However, the conclusion that evidence for link formation is "weak" and "vanishingly small" does not follow. This latter conclusion would hold only if one assumes that all behaviors that can be accounted for by either link formation or propositional reasoning (i.e., most behaviors) are in fact the end-product of propositional reasoning. As a consequence, Mitchell et al.'s conclusion implies this conclusion as a premise, which is the hallmark of circular reasoning.

A statistical analogy may be useful to make the point clearer, and to go a step further. If two strongly correlated variables, A and B, are predictive of the same dependent variable, C, and one is seeking to determine whether A or B is actually causal, then one may think of using hierarchical regression analysis. If a researcher believes that A is the genuine cause of C, he or she may enter A first as predictor, and, of course, B will be found to have a negligible residual contribution. The problem is that if another researcher believes that B is the cause, he or she may enter B first, with the opposite outcome. In both cases, the conclusion will depend on the initial belief, and hence cannot serve as a support for this belief. The Mitchell et al.'s (2009, p. 238) proposal that "the link mechanism is of somewhat limited explanatory value" is not an objective, theory-neutral assessment, it is nothing else than the corollary of considering that anything that fits well with the product of rational inferential processes is actually the product of these processes, even if a much simpler account is available. Faced with the same pattern of data, other researchers may find it more satisfying to publicize that associative strength defined as in the Rescorla and Wagner (1972) model of conditioning (for a review, see Miller, Barnet, & Grahame, 1995) accounts for complex adaptive behaviors that match nicely what would be the outcome of elaborate reasoning and rational considerations in similar conditions.

The final word is that dismissing counterevidence on the basis of the small number of phenomena that resist a propositional account of learning is questionable. Positing that a propositional account would be challenged only if such phenomena turn out to be quite common rests on the premise that, if automatic link formation processes exist, they should fail regularly to converge toward the same solution as that resulting from propositional reasoning. Now, such mismatches are certainly limited to infrequent arrangements, in which there is no, or very limited adaptive challenge. A truly random process is certainly an ideal illustration of these arrangements. Assuming two equally probable outcomes, selecting one or the other is neutral for adaptive purposes. In this case, automatic associative processes lead us to prepare to a repetition, whereas explicit expectancies lead us to prepare to an alternation, but no option is better than the other. As discussed above, the mode of sampling (without replacement) involved in the Perruchet's paradigm makes the situation a bit more complex than a truly random process, but similar enough to this ideal situation to apply the same line of reasoning.

About Alternative Models

As outlined in the introduction, the main alternative to the propositional framework is the dual-process model of learning (e.g., Clark et al., 2001; McLaren et al., 2014; also see Evans, 2010; Kahneman, 2011; Sloman, 2014 for related "dual-system" frameworks). The dual-process model of learning does not dismiss the existence of propositional reasoning, which still stands on the conscious side of mental life as in the propositional framework. However, a second process is postulated, generally viewed as a link formation mechanism. Researchers differ with regard to the specific relation of this second process to consciousness. The notion of automatic link processes is often conflated with the idea of unconscious learning (e.g., Clark et al., 2001). In keeping with this view, attempts to provide an existence proof of automatic link processes have often relied on conditions preventing the consciousness of contingencies. However, others investigators (e.g., McLaren, Green, & Mackintosh, 1994) have rejected the idea that the presence or absence of awareness of stimulus contingencies was a criterion discriminating learning systems. As an aside, the Perruchet effect is mute with regard to this issue. In all the experiments reported in the present review, participants were fully informed about the stimulus contingencies before the experimental session, and therefore, the Perruchet effect is in no way a demonstration that learning may occur without contingency awareness, nor is there evidence that consciousness is playing a causal role in learning.

Is the dual process model of learning the only alternative to the propositional view? This model seems especially well-suited to account for both the Perruchet effect (and a few other phenomena attesting to associative processes, McLaren et al., 2014) and the direct evidence we have that our behavior may sometimes be guided by conscious inferences. However, there are reasons for not to be entirely satisfied with this kind of model. Each of the processes composing the dual-process model seems to be endowed with antagonist limitations. The propositional system is composed of conscious representations and operations the existence of which may hardly be denied, but how these representations are built is left unspecified. By contrast, associative links define productive processes, but it is difficult to see how these processes can underpin complex human behavior, unless one considers that blinking to a conditioned tone or reacting faster to a predicted event is the best humans can learn. In a nutshell, the propositional system appears to be the end-product of unidentified mechanisms, while associative-link formation is a well-identified mechanism the end-product of which is uncertain. I would like to argue in this final section that this rather awkward view of mind could possibly give way to a more integrative account if the current simple snapshot of the forces at play was replaced by a temporally dynamic approach. I refer here to a view I have proposed with Annie Vinter (Perruchet & Vinter, 2002) that we call "self-organizing consciousness."

The basic principle of the self-organizing consciousness model is to join the two components isolated in the dual-process model of learning into a developmental, dynamic perspective, with the key proposal that conscious representations are the end-product of associative processes. Given the gap between the products of associative processes commonly studied in laboratory and the complexity inherent to conscious representations and reasoning, this proposal would have been a bit hazardous without some concrete evidence of feasibility. This was the primary objective of a computational model, PARSER (Perruchet & Vinter, 1998), which was applied to the word segmentation issue. This review is not the place to detail the model, but briefly, PARSER uses only very basic and ubiquitous processes of associative learning and memory, such as reinforcement with repetitions and interference. The model starts from a continuous sequence of syllables such as they may be perceived by infants, and extracts words, which can be reasonably conceived as the elementary bricks of thought involved in adults' cognitive activities. In so doing, PARSER builds a bridge, admittedly limited to a specific domain, between seemingly underexploited associative principles and a foundational component of linguistic abilities (for a recent use of PARSER, see, e.g., Perruchet, Poulin-Charronnat, Tillmann, & Peereman, 2014).

The obvious question is how the self-organizing consciousness model, which, as conveyed through its name, focuses on consciousness, can encompass the Perruchet effect, which seemingly reveals automatic link formation mechanisms, and more generally, how the model can account for the observed dissociation between explicit expectancies and overt behavior. In fact, such questions arise because for most people consciousness is exclusively linked to the notions of monitoring and control. Broadbent and colleagues dubbed as the "common sense" view of cognition the claim that "people act by consulting an internal model of the world, a database of knowledge common to all output processes, and manipulating it to decide on the best action" (Broadbent, Fitzgerald, & Broadbent, 1986, p. 77; also see Cleeremans & Jimenez, 2002). The self-organizing consciousness model does not take issue with the possible existence of such processes, but only as a facet, possibly minor, of the whole matter. Indeed, controlled reasoning and inferences exploit conscious representations that, as a consequence of their associative origin, are shaped by automatic link formation processes (see also Tzelgov, 1997). As such, the content of these representations is out of control (as the representations built in evaluative conditioning paradigms, at least when they are assessed through evaluative priming measures instead of self-reported evaluations, see Gawronski, Balas, & Creighton, 2014). This leads to a distinction between what has been coined as the *deliberative* (explicit) and *evocative* (implicit) mental episodes (Dulany, 1997, 2002). In deliberative mental episodes, mental contents would be related by explicit inferences and decisions, while in evocative mental episodes, conscious contents would be related by associative-activational operations. Under this framework, the Perruchet effect would reveal a dissociation between the (conscious) product of controlled inferential and reasoning processes, available through expectancy ratings, and the (conscious) representations directly evoked by associative processes, which generate overt responses.

More precisely, the self-organizing consciousness account of the Perruchet effect could proceed as follows. As the proponents of the propositional framework suggest, the associative history of E1 and E2 would make it possible to infer that E1 is sometimes followed by E2. Analysis of the sequence of trials would lead to assess the probability of E2 at a given point, and to translate this probability into expectancy ratings. These explicit ratings show that in a situation where two outcomes (E2 vs. not E2) occur randomly, expectancies are consistent with the gambler's fallacy, although it is difficult to determine whether this actually results from a fallacious reasoning, or from correct inferences based on the postulate that sampling certainly proceeds without replacement in a laboratory experiment. However, conscious representations evolve in a mandatory manner, following the standard laws of associative learning and memory. As a consequence, the associative history of E1 and E2 would also result in the automatic activation of the representation of E2 upon the presentation of E1. This automatic activation evolves in opposition to the gambler's fallacy, as a consequence of the experimental design.

Although the interpretation above may appear abstract and speculative, the basic dissociation is fully available to intuition. It is common experience to have quite vivid representations that pop-out in consciousness, and which are able to elicit strong responses of their own even if the represented events are in no way expected to occur in the near future. For instance, like the famous Proust's madeleine, a smell or a picture related to an emotionally loaded past event may automatically elicit emotional responses, even though one is quite certain that this event will not happen again. In contrast, one may expect the imminent arrival of an objectively arousing event without having strong reactions, maybe because the concrete representation of this event, and its resulting affective value, have declined over time. The dissociation observed in the Perruchet effect would be similar. After a long run of E1-E2 pairs, E1 would automatically evoke a vivid and precise representation of E2, even though E2 is no longer expected to occur again. By contrast, after a long run of E1alone, one would deliberately expect E2 on the next trial, but the quality of its representation (in the sense of Cleeremans & Jimenez, 2002) would have decreased because of forgetting and extinction.

Conclusion: Going Forward

Are all forms of associative learning in humans the consequence of cognitively grounded expectancies based on symbolic representations, or is there evidence for more basic automatic link formation processes? To address this issue, the prevalent approach has been to explore whether conditioned responding was possible without conscious awareness of the E1-E2 contingencies. After several decades of investigations, this possibility remains controversial, with the prevalent view being that there is no compelling evidence for it, but neither has there been a demonstration of its impossibility. Beyond its empirical intractability, this approach rests on the postulate that automatic link formation processes should have no conscious correlates. Although a compelling demonstration of unconscious learning would be certainly a strong support for automatic link formation, the lack of evidence for unconscious learning in no way undermines the existence, or even the prevalence of automatic link formation processes. Perhaps a major pitfall of past research has been to conflate the issue of the automaticity of associations and the inaccessibility of the process to consciousness, despite some cogent claims that such a position is not mandatory (McLaren et al., 1994).

The use of a partial reinforcement schedule in the Perruchet's (1985) paradigm opens a very different direction of research. The overall outcome is that immediately after a sequence of reinforced trials responses increase as predicted by the standard laws of reinforcement, whereas expectancy decreases. Conversely, after a sequence of nonreinforced trials, responses decrease, whereas expectancy increases. These results make a strong case for the concept of automatic association while remaining agnostic with regard to the issue of consciousness. The paradigm is certainly not an overpowered magic hand. Further studies are still necessary to definitely rule out the possibility that what is conceived as the associative influence of the E1-E2 pairings in the conventional interpretation of the task, is due instead to the nonassociative influence of E2 alone. It is also a challenge for future investigations to clarify whether the two potential sources of influence really conflict at any given time, or if the apparent dissociation is an artifact of averaging over trials and subjects. However, the promises of the paradigm, despite the current limitations, appear sufficient to go ahead and to extend its exploration.

The limited number of relevant studies leaves considerable room for extension and generalization. For instance, the cognitive involvement of participants could be manipulated in different ways. Among other possibilities, the instructions could convey more or less information to participants about the actual probability of the different outcomes. The timing of the events (SOA and ITI) could be modulated, with the general postulate that slower events and more time between trials more of a bias toward high-level thinking. Conversely, a masking task such as is frequently used in the conditioning literature could be exploited to decrease participants' involvement of higher-order cognitive processes. Individual differences regarding age and/or working memory capacities could also enlighten the processes at play in the task. Finally, the effect of greater variations in the procedure could be explored. For instance, all current studies (except Williams & Prokasy, 1977) have involved only two equiprobable outcomes. The limitation to two outcomes is certainly not a mandatory prerequisite for interpreting data from the task, nor is the equiprobability of these outcomes.

A still largely unexplored line of investigation that deserves a special mention concerns the electrophysiological correlates of the behavioral dissociation. A study by Xue, Lu, Levin, and Bechara (2011) provides a hint of the potential contribution of such investigations. The authors used a gambling task in which participants could gain or lose on each gamble. The situation substantially differed from Perruchet's paradigm, notably because the task was designed to measure the effect of only the just preceding gamble, and not the effect of increasingly longer runs of trials (gambles were presented as the repetition of pairs of "exposure" trial and "probe" trial, and not as a continuous sequence). However, the authors observed an effect similar to the gambler's fallacy: Participants were more risk-seeking after losing a gamble than after winning a gamble, as if their "luck" was going to turn. The relevant point for our concern is that functional magnetic response imaging (fMRI) data revealed that the behavioral decisions driven by the gambler's fallacy were associated with a consistent pattern, with an increased activation in the frontoparietal network, which is involved in cognitive control, and a decreased activation in the dorsal and ventral striatum. Other decisions were associated with the different pattern. Specifically, the authors related the activation in the dorsal and ventral striatum to the involvement of a "reinforcement learning mechanism," whereby wins are reinforced and losses lead to subsequent avoidance, which is the opposite of the gambler's fallacy. There is no doubt that this approach, as well as studies of patients with brain lesions, could enlighten the true nature of the Perruchet effect.

References

Barrett, L. C., & Livesey, E. J. (2010). Dissociations between expectancy and performance in simple and two-choice reaction-time tasks: A test of associative and nonassociative explanations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 864–877. http://dx.doi.org/10.1037/a0019403

Bertels, J., & Destrebecqz, A. (2013). *Does the Perruchet effect indicate a dissociation between learning and awareness?* Paper presented at the 18th ESOCP Conference (Budapest), Book of abstracts, p. 236.

Bertelson, P. (1961). Sequential redundancy and speed in a serial two-choice responding task. *The Quarterly Journal of Experimental Psychology, 13,* 90–102. http://dx.doi.org/10.1080/17470216108416478

Bertelson, P., & Renkin, A. (1966). Reaction times to new versus repeated signals in a serial task as a function of response-signal time interval. *Acta Psychologica, 25,* 132–136. http://dx.doi.org/10.1016/0001-6918(66)90007-2

Blanchard, T. C., Wilke, A., & Hayden, B. Y. (2014). Hot-hand bias in rhesus monkeys. *Journal of Experimental Psychology: Animal Learning and Cognition, 40,* 280–286. http://dx.doi.org/10.1037/xan0000033

Bonic, I. (1989). *A test for an expectancy theory of human eyeblink classical conditioning*. Unpublished honors thesis, University of New South Wales, Australia.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108,* 624–652. http://dx.doi.org/10.1037/0033-295X.108.3.624

Brewer, W. F. (1974). There is no convincing evidence for operant or classical conditioning in adult humans. In W. Weiner & D. Palermo (Eds.), *Cognition and symbolic processes* (pp. 1–42). Hillsdale, NJ: Erlbaum.

Broadbent, D. E., Fitzgerald, P., & Broadbent, M. H. P. (1986). Implicit and explicit knowledge in the control of complex system. *British Journal of Psychology, 77,* 33–50.

Burns, B. D., & Corpus, B. (2004). Randomness and inductions from streaks: "gambler's fallacy" versus "hot hand". *Psychonomic Bulletin & Review, 11,* 179–184. http://dx.doi.org/10.3758/BF03206480

Clark, R. E., Manns, J. R., & Squire, L. R. (2001). Trace and delay eyeblink conditioning: Contrasting phenomena of declarative and nondeclarative memory. *Psychological Science, 12,* 304–308. http://dx.doi.org/10.1111/1467-9280.00356

Clark, R. E., Manns, J. R., & Squire, L. R. (2002). Classical conditioning, awareness, and brain systems. *Trends in Cognitive Sciences, 6,* 524–531. http://dx.doi.org/10.1016/S1364-6613(02)02041-7

Cleeremans, A., & Jiménez, L. (2002). Implicit learning and consciousness: A graded, dynamic perspective. In R. M. French & A. Cleeremans (Eds.), *Implicit Learning and Consciousness* (pp. 1–40). Hove, United Kingdom: Psychology Press.

De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior, 37,* 1–20. http://dx.doi.org/10.3758/LB.37.1.1

De Houwer, J. (2014). Why a propositional single-process model of associative learning deserves to be defended. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual processes in social psychology* (pp. 560–577). New York, NY: Guilford Press.

Destrebecqz, A., Perruchet, P., Cleeremans, A., Laureys, S., Maquet, P., & Peigneux, P. (2010). The influence of temporal factors on automatic priming and conscious expectancy in a simple reaction time task. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 63,* 291–309. http://dx.doi.org/10.1080/17470210902888932

Dulany, D. E. (1997). Consciousness in the explicit (deliberative) and implicit (evocative). In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 179–211). Hillsdale, NJ: Erlbaum.

Dulany, D. E. (2002). Mentalistic metatheory and strategies. *Behavioral and Brain Sciences, 25,* 337–338. http://dx.doi.org/10.1017/S0140525X02290068

Dunn, J. C., & Kirsner, K. (1988). Discovering functionally independent mental processes: The principle of reversed association. *Psychological Review, 95,* 91–101. http://dx.doi.org/10.1037/0033-295X.95.1.91

Dwyer, D. M., Le Pelley, M. E., George, D. N., Haselgrove, M., & Honey, R. C. (2009). Straw-men and selective citation are needed to argue that associative-link formation makes no contribution to human learning. *Behavioral and Brain Sciences, 32,* 206–207. http://dx.doi.org/10.1017/S0140525X09000946

Evans, J. St. B. T. (2010). *Thinking twice: Two minds in one brain*. Oxford: Oxford University Press.

Fernandez-Duque, D., & Knight, M. (2008). Cognitive control: Dynamic, sustained, and voluntary influences. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 340–355. http://dx.doi.org/10.1037/0096-1523.34.2.340

Gawronski, B., Balas, R., & Creighton, L. A. (2014). Can the formation of conditioned attitudes be intentionally controlled? *Personality and Social Psychology Bulletin, 40,* 419–432. http://dx.doi.org/10.1177/0146167213513907

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology, 17,* 295–314. http://dx.doi.org/10.1016/0010-0285(85)90010-6

Jarvik, M. E. (1951). Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *Journal of Experimental Psychology, 41,* 291–297. http://dx.doi.org/10.1037/h0056878

Jiménez, L., & Méndez, A. (2013). It is not what you expect: Dissociating conflict adaptation from expectancies in a Stroop task. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 271–284. http://dx.doi.org/10.1037/a0027734

Jiménez, L., & Méndez, A. (2014), Even with time, conflict adaptation is not made of expectancies. *Frontiers in Psychology*. Retrieved from http://journal.frontiersin.org/Journal/10.3389/fpsyg.2014.01042/abstract

Jonides, J., & Mack, R. (1984). On the cost and benefit of cost and benefit. *Psychological Bulletin, 96,* 29–44. http://dx.doi.org/10.1037/0033-2909.96.1.29

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Strauss and Giroux.

Keil, A., Smith, J. C., Wangelin, B. C. C., Sabatinelli, D., Bradley, M. M., & Lang, P. J. (2008). Electrocortical and electrodermal responses covary as a function of emotional arousal: A single-trial analysis. *Psychophysiology, 45,* 516–523. http://dx.doi.org/10.1111/j.1469-8986.2008.00667.x

Knight, D. C., Cheng, D. T., Smith, C. N., Stein, E. A., & Helmstetter, F. J. (2004). Neural substrates mediating human delay and trace fear conditioning. *The Journal of Neuroscience, 24,* 218–228. http://dx.doi.org/10.1523/JNEUROSCI.0433-03.2004

Laplace, P. (1902). *A philosophical essay on probabilities*. New York, NY: Wiley & Sons (Translated by Truscott FW, Emory FL. First French ed.: 1814.).

Livesey, E. J., & Costa, D. S. (2014). Automaticity and conscious control in single and choice response time versions of the Perruchet effect. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 67,* 646–664. http://dx.doi.org/10.1080/17470218.2013.824014

Los, S. A., Knol, D. L., & Boers, R. M. (2001). The foreperiod effect revisited: Conditioning as a basis for nonspecific preparation. *Acta Psychologica, 106,* 121–145. http://dx.doi.org/10.1016/S0001-6918(00)00029-9

Lovibond, P. F., & Shanks, D. R. (2002). The role of awareness in Pavlovian conditioning: Empirical evidence and theoretical implications. *Journal of Experimental Psychology: Animal Behavior Processes, 28,* 3–26. http://dx.doi.org/10.1037/0097-7403.28.1.3

McAndrew, A., Jones, F. W., McLaren, R. P., & McLaren, I. P. L. (2012). Dissociating expectancy of shock and changes in skin conductance: An investigation of the Perruchet effect using an electrodermal paradigm. *Journal of Experimental Psychology: Animal Behavior Processes, 38,* 203–208. http://dx.doi.org/10.1037/a0026718

McAndrew, A., Weidemann, G., & McLaren, I. P. L. (2013). *Can US sensitization account for the electrodermal variant of the Perruchet effect?* Proceedings of the 35th Annual Conference of the Cognitive Science Society. Berlin.

McLaren, I. P. L., Forrest, C. L., McLaren, R. P., Jones, F. W., Aitken, M. R., & Mackintosh, N. J. (2014). Associations and propositions: The case for a dual-process account of learning in humans. *Neurobiology of Learning and Memory, 108,* 185–195. http://dx.doi.org/10.1016/j.nlm.2013.09.014

McLaren, I. P. L., Green, R. E. A., & Mackintosh, N. J. (1994). Animal learning and the implicit/explicit distinction. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 313–332). London, United Kingdom: Academic Press.

Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla-Wagner model. *Psychological Bulletin, 117,* 363–386. http://dx.doi.org/10.1037/0033-2909.117.3.363

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

Mitchell, C. J., Wardle, S. G., Lovibond, P. F., Weidemann, G., & Chang, B. P. I. (2010). Do reaction times in the Perruchet effect reflect variations in the strength of an associative link? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 567–572. http://dx.doi.org/10.1037/a0018433

Moore, J. W., Middleton, D., Haggard, P., & Fletcher, P. C. (2012). Exploring implicit and explicit aspects of sense of agency. *Consciousness and Cognition, 21,* 1748–1753. http://dx.doi.org/10.1016/j.concog.2012.10.005

Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: A review. *Consciousness and Cognition, 21,* 546–561. http://dx.doi.org/10.1016/j.concog.2011.12.002

Moratti, S., & Keil, A. (2009). Not what you expect: Experience but not expectancy predicts conditioned responses in human visual and supplementary cortex. *Cerebral Cortex, 19,* 2803–2809. http://dx.doi.org/10.1093/cercor/bhp052

Nicks, D. C. (1959). Prediction of sequential two-choice decisions from event runs. *Journal of Experimental Psychology, 57,* 105–114. http://dx.doi.org/10.1037/h0045193

Perruchet, P. (1985). A pitfall for the expectancy theory of human eyelid conditioning. *Pavlovian Journal of Biological Science, 20,* 163–170.

Perruchet, P., Cleeremans, A., & Destrebecqz, A. (2006). Dissociating the effects of automatic activation and explicit expectancy on reaction times in a simple associative learning task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 955–965. http://dx.doi.org/10.1037/0278–7393.32.5.955

Perruchet, P., Poulin-Charronnat, B., Tillmann, B., & Peereman, R. (2014). New evidence for chunk-based models in word segmentation. *Acta Psychologica, 149,* 1–8. http://dx.doi.org/10.1016/j.actpsy.2014.01.015

Perruchet, P., & Vinter, A. (1998). PARSER: A model for word segmentation. *Journal of Memory and Language, 39,* 246–263. http://dx.doi.org/10.1006/jmla.1998.2576

Perruchet, P., & Vinter, A. (2002). The self-organizing consciousness. *Behavioral and Brain Sciences, 25,* 297–330.

Prokasy, W., & Kumpfer, K. L. (1969). Conditional probability of reinforcement and sequential behavior in human conditioning with intermittent reinforcement schedules. *Psychonomic Science, 14,* 49–50. http://dx.doi.org/10.3758/BF03336421

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

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge, England: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511526817

Shanks, D. R. (2010). Learning: From association to cognition. *Annual Review of Psychology, 61,* 273–301. http://dx.doi.org/10.1146/annurev.psych.093008.100519

Shanks, D. R., & Lovibond, P. F. (2002). Autonomic and eyeblink conditioning are closely related to contingency awareness: Reply to Wiens

and Ohman (2002) and Manns et al. *Journal of Experimental Psychology: Animal Behavior Processes, 28*(3), 8–42. http://dx.doi.org/10.1037/0097-7403.28.1.38
Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences, 17,* 367–447. http://dx.doi.org/10.1017/S0140525X00035032
Skinner, B. F. (1942). The processes involved in the repeated guessing of alternatives. *Journal of Experimental Psychology, 30,* 495–503. http://dx.doi.org/10.1037/h0058201
Sloman, S. A. (2014). Two systems of reasoning: An update. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual processes in social psychology*. New York, NY: Guilford Press.
Stevenson, R. J., Boakes, R. A., & Prescott, J. (1998). Changes in odor sweetness resulting from implicit learning of a simultaneous odor-sweetness association: An example of learned synestesia. *Learning and Motivation, 29,* 113–132.
Tolman, E. C. (1932). *Purposive behavior in animal and man*. New York, NY: Appleton-Century-Crofts.
Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin, 76,* 105–110. http://dx.doi.org/10.1037/h0031322
Tzelgov, J. (1997). Automatic but conscious: That is how we act most of the time. In R. S. Wyer (Ed.), *The automaticity of everyday life: Advances in social cognition* (Vol. 10, pp. 217–230). Mahwah, NJ: Erlbaum Publishers.
Vervaeck, K. R., & Boer, L. C. (1980). Sequential effects in two-choice reaction time: Subjective expectancy and automatic after-effect at short response-stimulus intervals. *Acta Psychologica, 44,* 175–190. http://dx.doi.org/10.1016/0001-6918(80)90066-9
Weidemann, G., Broderick, J., Lovibond, P. F., & Mitchell, C. J. (2012). Both trace and delay conditioned eyeblink responding can be dissociated from outcome expectancy. *Journal of Experimental Psychology: Animal Behavior Processes, 38,* 1–10. http://dx.doi.org/10.1037/a0024411
Weidemann, G., Tangen, J. M., Lovibond, P. F., & Mitchell, C. J. (2009). Is Perruchet's dissociation between eyeblink conditioned responding and outcome expectancy evidence for two learning systems? *Journal of Experimental Psychology: Animal Behavior Processes, 35,* 169–176. http://dx.doi.org/10.1037/a0013294
Williams, W. C., & Prokasy, W. F. (1977). Classical skin conductance response conditioning: Effects of random intermittent reinforcement. *Psychophysiology, 14,* 401–407. http://dx.doi.org/10.1111/j.1469-8986.1977.tb02972.x
Xue, G., Lu, Z., Levin, I. P., & Bechara, A. (2011). An fMRI study of risk-taking following wins and losses: Implications for the gambler's fallacy. *Human Brain Mapping, 32,* 271–281. http://dx.doi.org/10.1002/hbm.21015

Received April 19, 2014
Revision received January 29, 2015
Accepted January 29, 2015 ■

Nicholas J. Mackintosh (1935–2015)

It is with great sadness that we inform our readers that Nicholas J. Mackintosh, a former editor of the Journal of Experimental Psychology: Animal Behavior Processes and the *Quarterly Journal of Experimental Psychology*, passed away on 8 February 2015, from a respiratory ailment after a short illness. Professor Mackintosh, FRS, was Head of the Department of Experimental Psychology at Cambridge University from 1981 to 2002. Following completion of his PhD at Oxford University in 1963, he successively served on the faculty at Oxford, Dalhousie University, and the University of Sussex, before moving to Cambridge where he was a Fellow of Kings College. Additionally, at various times, Professor Mackintosh held visiting professorships at the University of Pennsylvania, University of California, Berkeley, University of Hawaii, University of New South Wales, and Yale University. Mackintosh was renowned for his contributions to our understanding of both basic learning processes and human intelligence testing. His book, *The Psychology of Animal Learning*, published in 1974, set the tone for much of the research conducted in the field over the next 40 years, and is still widely used today. In 1975, he published a theory in *Psychological Review* concerning the role of attention in associative learning. Most subsequent theories of attention have used this as a starting point. In 2010, he and John Pearce published an important update of his 1975 model. Over his long and highly productive career, Professor Mackintosh mentored many of the major figures in the field of animal learning and cognition, both people who had the privilege of working with him in his laboratory and those who interacted with him through his numerous collaborations and several editorial roles. He will be missed, and his influence will be felt for many years to come.