CrossMark

ORIGINAL ARTICLE

Dissociating conscious expectancies from automatic-link formation in an electrodermal conditioning paradigm

Pierre Perruchet[1] · Laurent Grégoire[1] · Kevin Aerts[1] · Bénédicte Poulin-Charronnat[1]

Abstract The key point of a paradigm initially proposed by Perruchet (Pavlov J Biol Sci 20:163–170, 1985) to dissociate conscious expectancies from automatic-link formation in classical conditioning settings is the use of a partial reinforcement schedule, in which the unconditioned stimulus (US) follows the conditioned stimulus (CS) only half of the time on average. Given (pseudo) randomization, the whole sequence comprises runs of CS alone and runs of CS–US pairs of various lengths. When the preceding run goes from a long sequence of CS alone to a long sequence of CS–US pairs (via shorter sequences), associative strength should grow up, whereas conscious expectancy should decrease. Earlier studies have shown that, in most cases, conditioned performance parallels associative strength. As an exception, however, a few reports suggest that conditioned electrodermal responses (EDRs) would follow predicted changes in US expectancies. This paper presents an experiment that replicates this outcome. However, when the performances from a control group were taken as a baseline to control for response habituation, corrected conditioned EDRs were shown to follow associative strength. This suggests that the atypical pattern of conditioned EDRs in the Perruchet paradigm would be due to the fact that EDRs are more sensitive to habituation than responses involved in other associative learning settings. These results further challenge the recent "propositional" view of conditioning, which stipulates that conditioned responses in humans are the

consequence of participants' conscious inferences about the relationships between the CS and the US, which would lead the CS to generate conscious expectancy for the US.

Introduction

Despite its striking simplicity, the paradigm of classical conditioning is still the object of hot debates regarding the processes involved in humans. The most conventional interpretation is that conditioned responses (CRs) emerge automatically as a function of the contingencies between a conditioned stimulus (CS) and an unconditioned stimulus (US). The frequency or amplitude of CRs would be a function of the strength of the CS–US associations, which would follow the standard laws of simple associative learning known since Pavlov. However, an alternative, cognitively oriented interpretation is that human participants engaged in a classical conditioning setting quickly discover the relationship between the CS and the US. Afterwards, the presentation of the CS would trigger the conscious expectancy of the US, which would be responsible for the production of CRs.

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 (i.e., no longer paired with the US) 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

✉ Pierre Perruchet
pierre.perruchet@u-bourgogne.fr

[1] LEAD-CNRS, UMR5022, Université de Bourgogne, Pôle AAFE, 11 Esplanade Erasme, 21000 Dijon, France

US in this condition. Both interpretations are also compatible with a huge number of much more subtle variations in procedure.

As an exception, however, a paradigm initially proposed by Perruchet (1985) leads to opposite predictions. The key point of the Perruchet paradigm is the use of a random intermittent reinforcement schedule, with a reinforcement ratio of 50 %. This means that the CS is followed by the US only half of the time. Given (pseudo) randomization, the whole sequence comprises runs of CS alone and runs of CS–US pairings of various lengths (note that a run is defined hereafter as a sequence of consecutive trials of the same type). Associative strength should increase with the repetition of CS–US pairs, in keeping with the well-documented law of acquisition, and should decrease with the repetition of CS-alone trials, following the law of extinction (e.g., Pavlov, 1928). Simulations were run to examine how associative strength varies when the preceding run goes from a long sequence of CS alone to a long sequence of CS–US pairs (via shorter sequences, in the order listed on the x-axis of Fig. 1b). Associative strength was computed using the Rescorla–Wagner model of conditioning (Rescorla & Wagner, 1972; for a review, see Miller, Barnet, & Grahame, 1995). As expected, strength increased throughout training following a negatively accelerated curve (Fig. 1a; see "Appendix" for details). More relevant to the present concern, associative strength increased along a monotonous and roughly linear gradient when the preceding run goes from a long sequence of CS alone to a long sequence of CS–US pairs (Fig. 1b).

Crucially, US expectancy should predict the opposite outcome. Given that the US occurs after half of the CS on average, the common belief that alternations should occur more frequently than repetitions in random sequences of events (the so-called gambler's fallacy; e.g., Burns & Corpus, 2004) leads participants to anticipate a CS–US pair after a run of CS alone, and likewise, a CS alone after a run of CS–US pairs. In a nutshell, depending on prior trials, associative strength should increase, while US expectancy should decrease and conversely.

A number of studies exploiting this paradigm have shown that conditioned performance parallels associative strength. This result is now coined as the "Perruchet effect" after Weidemann, Tangen, Lovibond, & Mitchell (2009; see Perruchet, 2015, for a review). This pattern of results has been consistently observed in eyeblink conditioning studies (Clark, Manns, & Squire, 2001; Perruchet, 1985; Weidemann, Broderick, Lovibond, & Mitchell, 2012; Weidemann et al., 2009) and in cued reaction time tasks (Barrett & Livesey, 2010; Destrebecqz et al., 2010; Livesey & Costa, 2014; Mitchell, Wardle, Lovibond, Weidemann, & Chang, 2010; Perruchet, Cleeremans, & Destrebecqz, 2006).
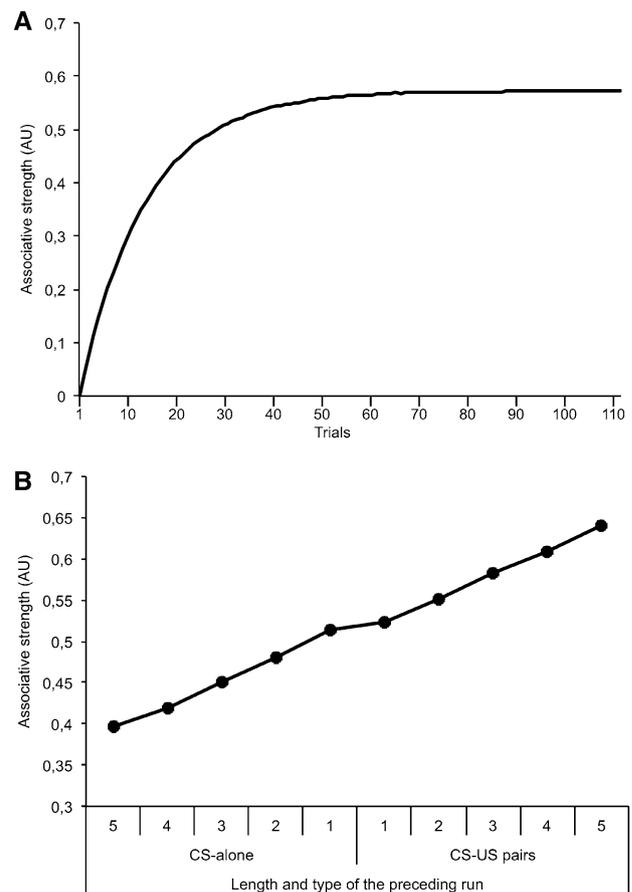
Fig. 1 Results from simulations. The Rescorla–Wagner model of conditioning predicts a negatively accelerated increase of associative strength across training (**a**). However, associative strength follows a roughly linear increasing trend when assessed as a function of the length (i.e., the number of trials) and the type (i.e., CS alone or CS–US pairs) of the run preceding any trial in the sequence (**b**). *AU* arbitrary units within a range from 0 to 1

However, a couple of studies in human autonomic conditioning has revealed different results. In McAndrew, Jones, McLaren, & McLaren (2012), there was a significant positive trend in electrodermal responses (EDRs) consistent with the Perruchet effect when analyses were carried out within each type of sequences (i.e., CS alone on the one hand and CS–US pairs on the other hand), but the authors observed no significant difference between trials following CS alone and trials following CS–US pairs (i.e., when the mean of all the trials falling into the left hand side of the figures are compared to the mean of all the trials falling into the right hand side). This goes against the predictions laid down in Fig. 1b. An earlier study by Williams & Prokasy (1977) reported still more adverse results. The experimental conditions were somewhat different from those used in recent studies devised to explore the Perruchet effect. In particular, the reinforcement ratio was .33 for one group and .67 for another group, instead of .50.
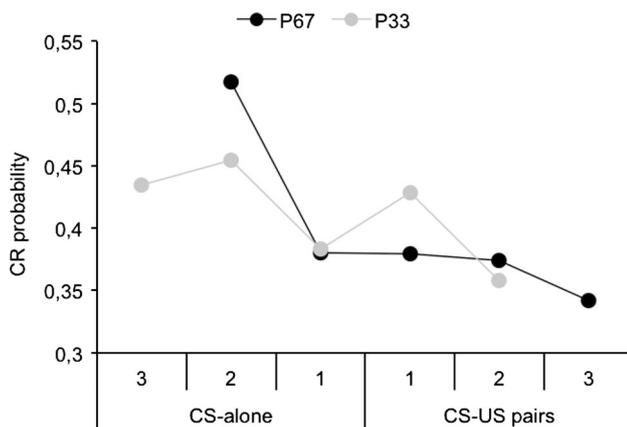
Fig. 2 Probabilities of conditioned EDRs in Williams & Prokasy (1977) for two groups of participants differing in the reinforcement ratio (p = .33 and .67). The data were rearranged from their Table 2. Williams and Prokasy distinguished two components of CRs, with a first component occurring just after the CS onset and the other just before the US onset. Although the overall frequency of the two components differed, they did not differ in their sensitivity to the preceding run of trials. As a consequence, the two components of responses were averaged in the figure above. The dominant negative slope runs counter the Perruchet effect

Figure 2 displays the results for each group. Overall, it is unquestionable that the dominant trend is negative, in opposition to the Perruchet effect, although the statistical analysis carried out by the authors do not allow confirming significance.[1]

A straightforward interpretation of this conflicting pattern of results is that associative strength would have a stronger influence on conditioned performance than US expectancy in eyeblink conditioning and cued reaction time tasks, whereas the conclusions would be more balanced in EDR conditioning (McAndrew et al., 2012), with expectancy sometimes prevailing over associative strength (Williams & Prokasy, 1977). However, there is an alternative interpretation for the pattern of EDRs, first suggested by Williams and Prokasy, which the present paper aims at testing. The effect observed in EDR conditioning would be due not to US expectancy, but instead to response habituation. Williams & Prokasy (1977) noted: "it is also possible that the sequential functions reflect nothing more that non-associative aftereffects of the UCS. Decreases across successive reinforcements may reflect suppression of performance in the presence of repeated stimulation with the UCS, while increases across successive non-reinforcements may reflect recovery of suppressed performance in the absence of stimulation with the UCS" (p. 406). It is indeed a fundamental property of habituation processes

that the response tends to recover over time, with the degree of spontaneous recovery depending on the length of the rest interval (e.g., Rankin et al., 2009). In this case, the CS alone would have no function of its own, but its occurrence attests that some time has elapsed (the duration of the inter-trial interval). For instance, assuming an inter-trial interval of 10 s, a US occurring after one, two, or three CS alone means that 20, 30, or 40 s (respectively) elapsed since the last US, leaving increasing time to recover, at least partially, the amplitude of the initial response.

It remains to explain why the result pattern in the Perruchet paradigm is reversed with EDRs when compared with other associative learning tasks. A possibility is that EDRs would be particularly sensitive to habituation. Taken as a general and definitive assessment, this hypothesis would be hardly defensible. For instance, when the investigation bears on the long-term (e.g., with one exposure by day), and/or involves very intense stimuli, habituation of EDRs may not occur or occur at a slow rate, and moreover, may not exhibit spontaneous recovery. Conversely, habituation of skeletal responses is commonly reported (e.g., Ornitz & Guthrie, 1989). However, a brief overview of the experimental literature is sufficient to assert that in the conditions involved in the usual paradigms of conditioning in humans (and notably within a single experimental session and moderately intense stimuli), habituation of autonomic responses (and especially EDRs) is more rapid and/or more pronounced than habituation of skeletal responses, such as eyeblink reflex and voluntary responses.

How may the influence of EDR habituation be controlled in the Perruchet paradigm? It is of course impossible to prevent habituation altogether, and therefore, our experimental strategy was to use a control group allowing the capture in isolation of the influence of EDR habituation on the short-term effect of the preceding runs of trials. The control group was submitted to the same sequence of events as the experimental group, except that CS-alone trials were replaced by a blank interval of the same length. As a consequence, all trials were of the same type (they were all CS–US pairs), and therefore the concept of "run" defined as a short sequence of consecutive trials of the same type is no longer relevant for the learner. However, it remains possible to analyze the data of this group as if the CS alone were displayed, creating "virtual" runs. Given that virtual runs exist only from the experimenter's standpoint, they cannot modulate learner's associative strength and expectancy (both of them depend on the actual presence of CS-alone trials). If EDRs still change as a function of the preceding (virtual) run in the control group, this should be ascribed to US habituation (which does not depend on the actual presence of CS, but only on the time interval that is necessarily involved in a CS-alone trial). In this way, the changes in EDRs as a function of the

---

[1] The single effect that is reported as significant is an interaction between two factors with two modalities, which are labeled as [CS alone; 1, 2] and [CS–US; 1, 2] in Fig. 2, respectively.

preceding run that are due to habituation can be assessed from the control group, then subtracted[2] from the performance observed in the experimental group to obtain an unbiased estimate of the combined influence of conscious expectancies and associative strength. What grounds the validity of a direct subtraction is that the rate and the timing of US occurrences, which govern habituation processes, were, at any moment, identical in the two groups.

Method

Participants

Forty-six undergraduate students (age range 18–26 years) from the University of Bourgogne took part in the experiment for partial fulfillment of a course requirement. They were randomly assigned to one of the two conditions (experimental versus control) with 23 participants per condition.

Apparatus

The CS was a 55 dB, 7-s tone of 600 Hz, and the US was an aversive 1-s burst of white noise at 85 dB, which arrived 6 s after the CS onset, and co-terminated with it. The electrodes used to capture the EDRs were Biopac TSD203 Ag–AgCl finger electrodes, which were placed on the fingertips of the index and the middle finger. The responses were measured by a Biopac MP150 device in combination with a Biopac GSR100C electrodermal amplifier. As an isotonic conductant, we used Biopac GEL101. An exosomatic direct current measurement was used, with a sampling rate of 200 Hz. We used the high-pass filter .05 Hz on the GSR100C machine, so electrodermal levels would be removed from the electrodermal reactions (i.e., a relative measure of electrodermal activity). The low-pass filter was set to 10 Hz and the gain level to 20 µS/V.

Procedure

The participants were informed that they were going to hear two sounds during the experiment, and that the first sound would be followed in 50 % of the cases (for the experimental group) or in all cases (for the control group)

---

[2] It is worth pointing out that the control group was strikingly different from the control groups routinely involved in conditioning settings. Control participants are usually assigned to conditions that prevent associative learning, to assess whether, by comparison, conditioning actually occurs in the experimental group. By contrast, removing CS-alone trials from the original paradigm leads to turn the 50 % reinforcement ratio into a 100 % reinforcement ratio, with the unusual consequence that conditioned performance should be better in the control group than in the experimental group and hence, that subtracting the former from the latter should return negative scores.

Table 1 The organization of trial types

	CS alone or blank interval				CS–US pairs			
Run length	4	3	2	1	1	2	3	4
Frequency	2	4	6	8	8	6	4	2

CS conditioned stimulus, *US* unconditioned stimulus

by the second sound. The sounds were presented to them to make sure they knew which sounds to expect.

In most studies using the Perruchet paradigm, the sequences were designed to conform exactly to a binomial distribution of two equally probable events (CS alone and CS–US pairs). As a consequence, the number of available datapoints decreases as the length of the runs increases, quickly reaching zero. In Williams & Prokasy (1977) and McAndrew et al. (2012), the maximum length of the runs was limited to three trials. This is damageable, because the longer runs are also the more informative in the logic of the procedure. In the present experiment, each run was displayed twice to collect data for four-trial runs, while keeping the total number of trials manageable. The resulting number of data points available for each run length is shown in Table 1. The sequences were constructed by randomly selecting lengths of CS–US runs and CS-alone runs alternately, following a method borrowed from Nicks (1959). Two CS–US trials were added, one after each four-trial run of CS–US pairs, to take measurements after these runs in the control group. No measure was taken on the trials following these extra-trials.

All experimental participants were exposed to the same sequence of runs. The inter-trial interval varied within a range of 30–40 s. They were randomized for each participant. Control participants were exposed to the same sequence, except that the CS-alone trials were replaced by a blank interval of 7 s (for the sake of convenience, this interval will be designated hereafter as "no-CS", although from the participant's standpoint, there was no event at all). Experimental and control participants were yoked, in such a way that overall, the temporal distribution of the CS–US trials was exactly the same for the two groups.

During the whole experiment, which lasted about 45 min, Chinese ambient music was played in the background, to help participants to remain relaxed without falling asleep. Participants were told to avoid movements, in particular those of the right arm from which the electrodermal activity was recorded.

Results

The CRs could be measured on CS-alone trials in the experimental group, but not in the control group where such trials were absent. To compare group performances on

Fig. 3 Mean skin conductance for the CS–US pairs, taken between 5 s before the CS onset and 15 s after the CS onset for both the experimental and control groups. The *filled symbols* are the values that fall into the latency range defining the conditioned EDRs (from 2 to 7 s)
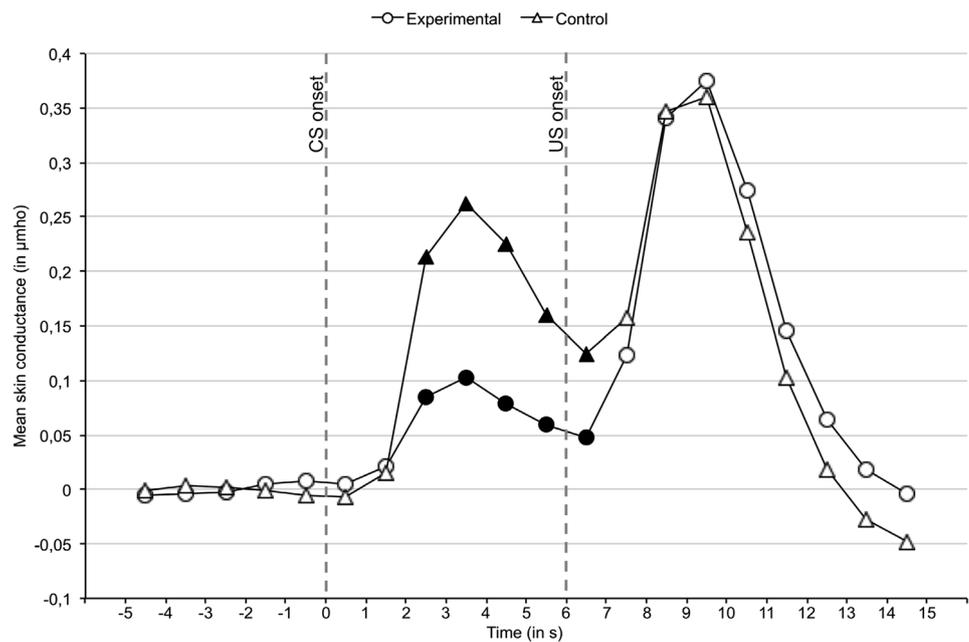
the very same trials, only the anticipatory responses in the CS–US trials were considered for both groups. To define the latency range of the CRs, the level of electrodermal activity was averaged over the 22 CS–US trials within a window going from 5 s before the CS onset to 15 s after the CS onset, separately for the experimental and the control groups. It appears on Fig. 3 that the mean anticipatory response was larger for the control group than for the experimental group, in keeping with the well-known superiority of total reinforcement over partial reinforcement. However, response latencies did not differ between groups. In addition, for both groups, there was no hint toward the existence of the two components of responses distinguished by Williams & Prokasy (1977; see caption of Fig. 2) and as a consequence, no subdivision was carried out, as in McAndrew et al. (2012). The CRs were defined as any responses occurring in the 2–7 s latency range. To take into account the conductance baseline at the CS onset, the amplitude of the CRs was measured, for each trial, as a difference between the maximum level reached in this latency range, and the maximum level reached during the 5 s preceding the CS onset.

The mean amplitude of the CRs as a function of the preceding run is depicted in Fig. 4. Recall that in the control group, the preceding runs were only virtual, given that all trials were of the same type (CS–US pairs) in this group. Any effect of runs was expected to reflect response habituation. As observed on the averaged curves, control participants outperformed experimental participants. More importantly, the CRs slightly decreased as a function of runs in the experimental group, hence going in a direction opposite to the standard Perruchet effect. Finally, a

stronger decrement was observed in the control group. A repeated-measures mixed ANOVA on CR amplitude performed with Group (Experimental versus Control) as a between-subject variable and Run (8 levels) as a within-subject variable confirmed these observations. The main effect of groups was significant, $F(1, 44) = 5.18$, $p = .028$, $\eta = .105$. The main effect of Run was also significant, $F(7, 308) = 3.83$, $p < .001$, $\eta = .08$, as well as the Group $\times$ Run interaction, $F(7, 308) = 2.26$, $p = .029$, $\eta = .049$.

Planned comparisons revealed that the downward linear trend for the experimental group was quite close to statistical significance, $F(1, 22) = 4.27$, $p = .051$, $\eta = .163$. However, the downward linear trend for the control group, which was clearly reliable, $F(1, 22) = 27.63$, $p < .001$, $\eta = .557$, was still steeper, as shown by the interaction between groups and the linear component of the effect of runs, $F(1, 44) = 13.51$, $p = .001$, $\eta = .235$. To make the implication of this interaction more explicit, Fig. 4 represents the scores of the experimental participants, which presumably conflate the genuine effect of runs with the effect of US recency on EDRs, after subtracting the scores of the control group, which presumably capture only the effect of US recency. The resulting trend is clearly positive,[3] as in the other experimental settings using the Perruchet (1985) paradigm.

---

[3] The fact that the values are in the negative range reflects nothing else than the advantage of total reinforcement (control group) over partial reinforcement (experimental group). Only the relations between the points of the curve reflect the size of the effect of the preceding runs, which is of concern here.
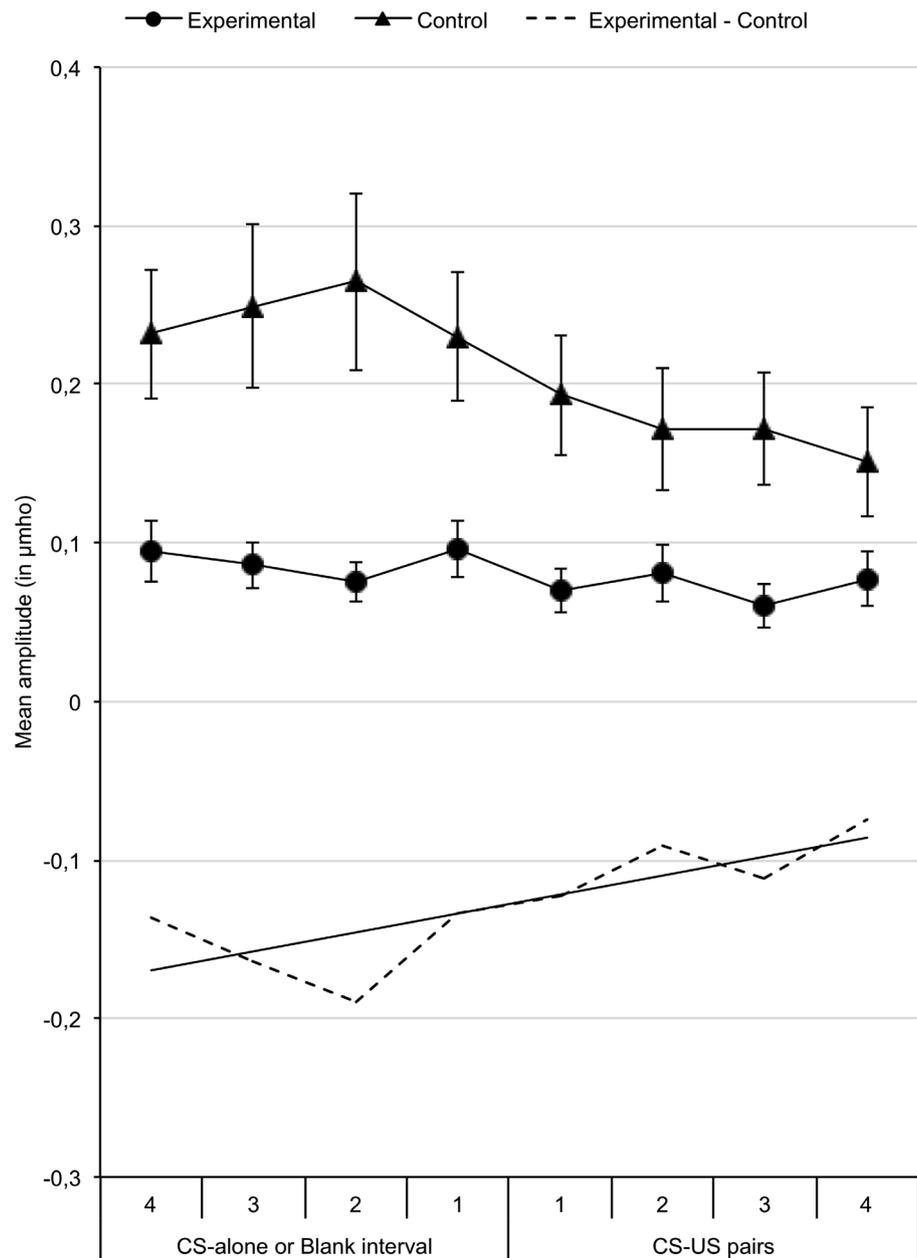
One could argue that slope differences are a mechanical consequence of differences in the mean amplitude of the responses. Because CRs in the control group were larger on the average than those in the experimental group, they could vary along a larger span. To address this issue, we adapted a correction proposed by Lykken & Venables (1971; see Boucsein, 1992, p. 154). For each participant, the CR amplitude for a given condition was divided by the maximum value reached by this participant over all conditions. As a consequence, for each participant, the score for the condition generating the strongest CRs was set to 1, and the scores for the other conditions were expressed as a proportion of this value. In addition to controlling for between-group differences, this correction has the advantage of reducing inter-individual variance within each group. A repeated-measures ANOVA was performed as above on the corrected scores. The Group × Run interaction was only marginally significant, $F_{(7, 308)} = 1.75$, $p = .098$, $\eta = .038$, but planned comparisons[4] still revealed a significant interaction between groups and the linear component of the effect of runs, $F_{(1, 44)} = 6.56$,

---

[4] Planned comparisons are sometimes claimed to be inappropriate when the overall F is not significant. However, a number of experts in statistics have convincingly argued that this general advice was unwarranted (e.g., Howell, 2010; Ryan, 1959; Wilcox, 1987).

$p = .014$, $\eta = .130$. The downward linear trend remained steeper for the control group than for the experimental group (and hence the between-group differences still revealed a positive linear trend typical of the Perruchet effect), even when the scores were expressed as proportions to remove undesirable scaling effects.

Discussion

Williams & Prokasy (1977) reported that the probability of conditioned EDRs in a partial reinforcement paradigm "decreases across sequences of successively reinforced trials and increased across sequences of successively non-reinforced trials" (their abstract) as predicted by the gambler's fallacy. Our own results are consistent with this observation, and hence may be taken, at first glance, as a confirmation of the failure to observe the Perruchet effect in electrodermal conditioning.

However, Williams & Prokasy (1977) also acknowledged that their result pattern could reflect nothing more than non-associative aftereffects of US repetition on EDRs. To capture these potential effects, we ran a control group in which the timing of the CS–US pairs was identical to that of the experimental group, but the CS-alone trials were removed and replaced by a blank interval to make irrelevant the predictions related to the balance between conscious expectancy and associative strength in the logic of the Perruchet paradigm. The CRs of the control group followed a downward trend as in the experimental group. Importantly, this effect cannot be imputed to the gambler's fallacy: Expecting that a sequence of CS–US pairs should be followed by a CS-alone trial would have been nonsensical for control participants, for whom all trials were reinforced. Moreover, this downward trend was significantly steeper than in the experimental group. When the scores of the control participants were subtracted from the scores of the experimental participants to isolate the effect of runs from the effects due to US-elicited habituation, conditioned performance followed the upward linear trend predicted by variations in associative strength, as observed in most other paradigms.

McAndrew et al. (2012) reported results in partial contradiction with both those of Williams & Prokasy (1977) and ours. They found no significant difference between negative and positive runs (which means that presumably, the overall linear trend, if calculated, would have been flat). Moreover, their raw data (i.e., without any correction for response habituation) followed the Perruchet effect when analyzed within a given type of runs. We see no clear explanation for this discrepancy. A procedural difference is that McAndrew et al. was the only study to require a concurrent rating of conscious expectancy. Because this rating implied to press a button on each CS, a possibility is that the effect of this movement on EDRs partially overshadowed the expected effects of the preceding run of trials. Whatever the source of the discrepancy, however, McAndrew et al. are in broad agreement with our general conclusion, which confirms the Perruchet effect in paradigms of EDR conditioning.

This experiment makes increasingly clear that any inference drawn from the data collected in the Perruchet paradigm imperatively needs a comparison with a control condition capturing the non-associative effects of US repetition. Earlier studies showed that this is true when the raw data seemingly support the Perruchet effect, because the observed trend may be caused by various alternative processes. In eyeblink conditioning, the effect may be due to pseudo-conditioning (Perruchet, 1985; Weidemann et al., 2009). In cued reaction time tasks, the effect may be due to sensitization, a process mediated by the transitory activation of the mental representation of US, or alternatively by some kind of motor priming for a specific response. A still other view is that the presentation of the response signal and the voluntary button press that follows would induce a temporary increase in the level of arousal, whereas participants would lose concentration when the last stimulus is far back in time (Mitchell et al., 2010; Perruchet et al., 2006). A large array of procedures has been used to control for these potential effects, such as decoupling CS and US (Mitchell et al., 2010; Perruchet, 1985; Weidemann et al., 2009), reversing the usual CS–US order (Mitchell et al., 2010), removing the CS altogether (McAndrew et al., 2012), and using a differential conditioning procedure (Barrett & Livesey, 2010). The present paper suggests that a control condition is also necessary when the raw data seemingly go against the Perruchet effect. Indeed, our experiment shows that the effect of US recency can overshadow, and even reverse the genuine effect of runs on associative strength. This is because, by contrast with skeletal responses, EDRs undergo fast habituation in the event the same stimulus is repeated over and over, and the alternating short-term effects of habituation and recovery from habituation run in a direction opposite to the changes in associative strength.

The upsurge of studies exploiting the Perruchet paradigm in the last few years (for a review, see Perruchet, 2015) is presumably due to the fact that it provides one of the strongest counterarguments to the so-called "propositional" model of learning, recently proposed by Mitchell, De Houwer, & Lovibond (2009; see also De Houwer, 2009, 2014). In this model, the conscious expectancy of the US at the presentation of the CS is the only factor responsible for the production of CRs, instead of being superimposed on the conventional automatic-link formation mechanism. For instance, when an EDR to a tone occurs after the pairing of

the tone with an aversive event, it would be because participants have inferred that the tone is followed by the aversive event based on their prior experience, and this propositional knowledge would lead participants to consciously expect the aversive event at the tone occurrence.

It could be argued that because we did not measure expectancies in our experiment, our results are inconsequential for the propositional view. However, it must be realized that advocating for the exclusive role of expectancy in this experiment would imply that US expectancies changed as a function of the preceding trials as expected from associative strength, that is in a direction opposite to the gambler's fallacy. This hypothesis is highly implausible. The gambler's fallacy has been observed in all the earlier studies involving the Perruchet paradigm, irrespective of parametric variations. For instance, McAndrew et al. (2012) confirmed the presence of the gambler's fallacy in an EDR conditioning paradigm using time parameters very similar to ours. To conclude, the available data strongly suggest that conditioned EDRs are governed by associative strength as other CRs, thus giving a new existence proof to the automatic-link formation mechanism that the proponents of a propositional view of learning deny.

Appendix

The Rescorla–Wagner model of classical conditioning (Rescorla & Wagner, 1972; for a review, see Miller, Barnet, & Grahame, 1995) assumes that the gain in associative strength due to the pairing of a CS with a US is proportional to the difference between the maximum strength of association that the US can support, and the current associative strength of all the CSs present on this trial. As a consequence, the rate of change in conditioned responding to a CS decreases as the number of training trials increases, resulting in a negatively accelerated curve of learning across training.

More formally, two equations describe how associative strength of a given CS (CS_x) changes on each trial:

$$\Delta V_X^{n+1} = a_X \beta (\lambda - V_{\text{tot}}) \tag{1}$$

and

$$V_X^{n+1} = V_X^n + \Delta V_X^{n+1} \tag{2}$$

where ΔV_X^{n+1} is the change in the associative strength (V) of CS_x on Trial $n + 1$; α_x is the associability of CS_x (range 0–1), β is the associability of the US (range 0–1); λ is the maximum associative strength that the US can support; and V_{tot} is the sum of associative strengths of all CSs (including

X) present on Trial n. In Eq. 2, V_X^{n+1} is the associative strength of CS_x after Trial $n + 1$; V_X^n is the associative strength of CS_x immediately before Trial $n + 1$ (i.e., after trial n); and ΔV_X^{n+1} is the change in the associate strength of CS_x as a result of Trial $n + 1$, as assessed by Eq. 1.

In the Perruchet paradigm, there is a single CS and a single US. In the simulation below, α was arbitrarily set to 1; and β was given the function of a learning rate parameter. The paradigm comprises both CS–US pairs (acquisition) and CS alone trials (extinction). λ was set to 1 for acquisition trials and zero for extinction trials, and β took different values for the different type of trials. The Rescorla–Wagner model assumes that β for extinction is a number smaller than β for acquisition but larger than zero. β was set to .08 for CS–US pairs and .06 for CS-alone trials. These values were empirically selected in such a way that the learning curve reaches an asymptote within the set of trials usually involved in studies investigating the Perruchet effect.

The training sequences were constructed by drawing randomly among a set of runs from which number and length were previously determined. For each type of trials (i.e., CS–US and CS alone), there were one run of five trials, two runs of four trials, four runs of three trials, eight runs of two trials and 16 runs of a single trial. The resulting sequences comprised 114 trials [i.e., $2(1 \times 5 + 2 \times 4 + 4 \times 3 + 8 \times 2 + 16 \times 1)$] and conformed exactly to a binomial distribution of two equally probable events.

The results shown in Fig. 1 have been averaged over 10.000 randomly generated sequences. Figure 1a shows that the associative strength increased following a negatively accelerated function across the 114 trials. Note that the asymptotic value is far from 1 (λ), which is typical of the prediction of the model with a partial reinforcement paradigm. Figure 1b shows the variation of associative strength as a function of the length and type of the preceding run of trials. The curve now follows a roughly linear positive trend.

It is worth stressing that predictions are related to associative strength, and not to overt behavior. Rescorla & Wagner (1972) themselves noted that acquired responding "would necessarily depend on a large number of 'performance' variables" (p. 77). This cautionary note is especially relevant for EDR conditioning, because EDRs are strongly dependent on habituation.

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, 3*, 864–877. doi:10.1037/a0019403.

Boucsein, W. (1992). *Electrodermal activity*. New York: Plenum Press.

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

Clark, R. E., Manns, J. R., & Squire, L. R. (2001). Trace and delay eyeblink conditioning: contrasting phenomena of declarative and non-declarative memory. *Psychological Science, 12*, 304–308.

De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning and Behavior, 37*, 1–20.

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. 530–541). New York: Guilford.

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. *Quarterly Journal of Experimental Psychology, 63*, 291–309.

Howell, D. C. (2010). *Statistical methods for psychology* (8th ed.). Belmont: Cengage Wadsworth.

Livesey, E., & Costa, D. (2014). Automaticity and conscious control in single and choice response time versions of the Perruchet effect. *The Quarterly Journal of Experimental Psychology, 67*, 646–664.

Lykken, D. T., & Venables, P. H. (1971). Direct measurement of skin conductance: a proposal for standardization. *Psychophysiology, 8*, 656–672.

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.

Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla–Wagner model. *Psychological Bulletin, 117*, 363–386.

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences, 32*, 183–198. doi: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. doi:10.1037/a0018433.

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

Ornitz, E. M., & Guthrie, D. (1989). Long-term habituation and sensitization of the acoustic startle response in the normal adult human. *Psychophysiology, 26*, 166–173.

Pavlov, I. P. (1928). Lectures on conditioned reflexes (trans: Gantt, W.H.). London: Allen and Unwin.

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

Perruchet, P. (2015). Dissociating conscious expectancies from automatic link formation in associative learning: a review on the so-called Perruchet effect. *Journal of Experimental Psychology: Animal Learning and Cognition, 41*, 105–127.

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. doi:10.1037/0278-7393.32.5.955.

Rankin, C. H., Abrams, T., Barry, R. J., Bhatnagar, S., Clayton, D. F., Colombo, J., & Thompson, R. F. (2009). Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. *Neurobiology of Learning and Memory, 92*(2), 135–138. doi:10.1016/j.nlm.2008.09.012.

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

Ryan, T. A. (1959). Comments on orthogonal components. *Psychological Bulletin, 56*, 394–396.

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. doi: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*, 169–172. doi:10.1037/a0013294.

Wilcox, R. R. (1987). New designs in analysis of variance. *Annual Review of Psychology, 38*, 29–60.

Williams, W. C., & Prokasy, W. F. (1977). Classical skin conductance response conditioning: effects of random intermittent reinforcement. *Psychophysiology, 14*, 401–407. doi:10.1111/j.1469-8986.1977.tb02972.x.