The role of common elements in the redundancy effect

Metin Uengoer*, Harald Lachnit, Philipps-Universität Marburg, Germany

John M Pearce, Cardiff University, UK

Running title: The Redundancy Effect

* Corresponding Author

Author Note:

Metin Uengoer

Faculty of Psychology

Philipps-Universität Marburg

Gutenbergstraße 18

35032 Marburg

Germany

Email: uengoer@staff.uni-marburg.de

Abstract

In two experiments, participants received a predictive learning task in which the presence of one or two food items signalled the onset or absence of stomach ache in a hypothetical patient. Their task was to identify the cues that signalled the occurrence, or non-occurrence of this ailment. The two groups in Experiment 1, and the single group in Experiment 2, received a blocking treatment, where cue A and a combination of cues A and X both signalled stomach ache, A+ AX+. These groups also received a simple discrimination where the outcome was signalled by one compound but not another, BY+ CY-. Subsequent test trials revealed the so-called redundancy effect, where X was regarded as a more reliable predictor of the outcome than Y. This result occurred when the trials with A+ preceded those with AX+ (Group E, Experiment 1, and Experiment 2), and when the trials with A+ and AX+ were intermixed (Group C, Experiment 1). The results challenge theories based on the assumption that cues presented together must compete for a limited pool of associative strength. Rather, they are said to support theories that assume changes in attention determine what is learned when two or more cues are presented together.

Allan Wagner, in his chapter for the book Fundamental Issues in Associative Learning (Mackintosh & Honig, 1969), focussed on two methods for investigating the role of what he termed cue validity in associative learning. One method was blocking (e.g. Kamin, 1969), the other was referred to as relative validity (Wagner, Logan, Haberlandt & Price, 1968). In the first part of Wagner's chapter, two blocking experiments are described, both of which involved an experimental group that received a single cue paired with an outcome, A+, intermixed among trials in which the same cue along with another cue was paired with the same outcome, AX+. A control group just received the compound trials, AX+. Subsequent test trials with X by itself revealed a considerably stronger response in the control group than the experimental group. The next part of the chapter dealt with relative validity. In each of the three reported experiments, two groups received two compounds, both of which contained a unique cue and a common cue, AY and BY. For an experimental group, AY, but not BY was paired with the outcome, AY+ BY-, for a simple discrimination, while for a control group, each compound was paired with the outcome on half the trials, and nothing on the remaining trials, AY+/- BY+/- for a pseudo-discrimination. Even though Y was paired with the outcome on half the trials in both groups, the experiments consistently revealed a stronger response to this cue in the control than the experimental group. This difference is referred to as the relative validity effect.

At the time, Wagner (1969) regarded these findings as evidence that "relative informativeness", or "cue validity", was critically important for determining the extent to which each of two cues gain control over responding when they are presented together. He then went on to suggest that the increment in the signal value of a cue, B, is "a function of the degree to which reinforcement is predictable *on the basis of the entire configuration of cues among which B is included.*" (p. 115). Several years later these ideas were presented more formally as an equation by Wagner and Rescorla (1972, see also Rescorla & Wagner, 1972).

In this equation, the increment in the strength of the association between cue A and an outcome on any trial, ΔV_A, is given by Equation 1, where λ is the asymptote of learning set by the magnitude of the outcome, V_T is the sum of the associative strengths of all the cues present on the trial in question, and α and β are learning rate parameters with values between 0 and 1 that reflect, respectively, properties of the cue and the outcome.

$$\Delta V_A = \alpha_A.\beta.(\lambda - V_T) \qquad 1$$

The explanation for blocking follows readily from Equation 1. With sufficient intermixed training trials a blocking treatment of the form A+ AX+ will result in the associative strength of A reaching the asymptotic value, λ, while that for X will be zero. Turning to the relative validity experiments, the analysis is rather more complex. The simple discrimination given to the experimental group, AY+ BY-, is predicted to result in the associative strength of Y having a positive value, which will be protected from extinction on BY- trials by inhibition (negative associative strength) acquired by B. The partial reinforcement schedule involving AY and BY in the control group is also predicted to result in Y gaining associative strength, and provided the learning rate parameter, β, has a higher value on reinforced than nonreinforced trials, it then follows from the equation that the associative strength of Y in the control group will be greater than of Y in the experimental group. Experiments by Rescorla (2002) have shown that the foregoing proviso is reasonable.

Blocking and the relative validity task both played a seminal role in theorising about the circumstances that promote associative learning. Rather little interest was shown, however, in the relative effects of these treatments on the response to the redundant cues in the two tasks. This neglect is surprising because, as we show shortly, the eventual fate of these cues has important theoretical implications. In Wagner's (1969) blocking condition, A+ AX+, cue A reliably signals when the outcome will occur, thus rendering X redundant as a cue for providing information about the trial outcome. Turning to the simple discrimination

of the relative validity design, AY+ BY-, cues A and B provide accurate information about when the outcome will occur and thus render Y redundant. Given the lack of informativeness, or low cue validity of X and Y, according to the speculations of Wagner (1969) these cues might be treated in the same way, and elicit responses of similar magnitude after the different training schedules. A rather different prediction is made by the more formal analysis offered by the Rescorla-Wagner (1972) theory. As we have just seen, provided sufficient trials are given, it follows from Equation 1 that a blocking treatment of the form A+ AX+ will result in the associative strength of X eventually being zero, whereas in the case of the simple discrimination, AY+ BY-, the associative strength of Y is predicted to be greater than zero. A direct prediction from the theory, therefore is that, at asymptote, the associative strength of Y will be greater than of X.

In fact, when the effects of these different treatments were eventually compared by Pearce, Dopson, Haselgrove and Esber (2012), the above predictions were found to be incorrect. A single group of pigeons was presented with a blocking treatment, A+ AX+, and a simple discrimination, BY+ CY-. In contrast to the prediction from Equation 1, subsequent test trials then revealed a stronger response to X than Y. As well as being observed with pigeons, this effect, which is referred to as the redundancy effect, has been observed with rats (Jones & Pearce, 2015), and humans (Jones & Zakseite, 2018; Jones, Zakseite, & Mitchell, 2019; Uengoer, Lotz, & Pearce, 2013).

The speculations of Wagner (1969), and the Rescorla and Wagner (1972) model, are not the only theoretical accounts to be challenged by the redundancy effect. A crucial assumption of Equation 1 is that associative learning is governed by a pooled error term. That is, the change in associative strength of a cue is determined by the discrepancy between the asymptote of conditioning, λ, and the sum of the associative strengths of all the cues present on the trial in question, V_T. As Uengoer et al. (2013) point out, a number of theories

share this assumption and are therefore led into making the same erroneous prediction concerning the redundancy effect as the Rescorla-Wagner theory (e.g. Gluck & Bower, 1988; Le Pelley, 2004; Pearce & Mackintosh, 2010; Pearce, 1994; Pearce & Hall, 1980).

A long-standing feature of Wagner's theorising was the ingenious way in which he made use of elemental cues to represent components of individual stimuli, or to represent combinations of stimuli. This strategy was originally adopted in order to enable the Rescorla-Wagner (1972) theory to explain how certain discriminations are solved when the outcome is signalled by more than a single stimulus – for example, the conditional discrimination studied by his student, Saavedra (1971), or negative patterning (e.g. Whitlow & Wagner, 1972). In the latter, the outcome is signalled by two cues when presented separately, but not when presented together, A+ B+ AB-. Although the Rescorla-Wagner model predicts this discrimination will never be solved, with responding to AB predicted to be consistently stronger than to A or B alone, there are numerous demonstrations of this discrimination being mastered. In order to accommodate this kind of result, Wagner (1971) proposed that when two or more cues are presented together, they create a "configurational" element that is able to enter into associations in the same way as conventional cues. Negative patterning can then be explained because the configurational cue created by the combination of A and B will gain negative associative strength that will counteract the excitatory influence of A and B.

Subsequent articles explored the implications of further developments of these principles. In his model of automatic memory processing in animals, Wagner (1981) proposed that the presentation of a stimulus will activate a set of elements that decay through two different states of activation before returning to being inactive. While Wagner (2003), in his context-sensitive elementary theory, suggested that the elements excited by a stimulus are either context independent, and excited regardless of the presence of other stimuli, or context

dependent, and either excited or inhibited by the presence of another element (see also Brandon & Wagner, 1998; Wagner & Brandon, 2001).

By adopting the foregoing proposals, Wagner was able to explain a wide range of phenomena, including occasion setting (Brandon & Wagner, 1998), the role of temporal factors in conditioning (Wagner, 1981), and various aspects of stimulus generalization (Wagner, 2003). Of particular concern to the present discussion, however, is the proposal that it is possible to explain the redundancy effect with the Rescorla-Wagner (1972) theory by appealing to elements that are common to a set of experimental stimuli. In a theoretical note, Vogel and Wagner (2017) offered the simple suggestion that during an experiment, all the experimental stimuli share a common feature, K, because they are all similar to each other in some way (see also Haselgrove, 2010). Thus, if a single group is trained with the trials A+ AX+ BY+ CY-, Vogel and Wagner suggested this treatment can be represented as AK+ AXK+ BYK+ CYK-. Applying Equation 1 to this characterisation leads to the prediction that test trials with XK will lead to a stronger response than to YK, which is the redundancy effect. For their formal derivation of this prediction, Vogel and Wagner assigned a value of .4 to all the cues, including K, while the values of β on trials with and without an outcome, respectively, were .2 and .1. Vogel and Wagner (2017) acknowledge that with a different set of assumptions concerning these values, the common-cue version of the Rescorla-Wagner theory is unable to explain the redundancy effect. Despite this shortcoming, the fact that with a plausible set of parameter values this version of the theory can explain the redundancy effect means that the proposal of Vogel and Wagner should be taken seriously. Vogel and Wagner further acknowledge that their explanation was developed for within-subject demonstrations of the redundancy effect, but when different groups are trained with A+ AX+ and BY+ CY-, then differences in the associative strength of K might account for the

successful demonstrations of the redundancy effect in these circumstances (Jones & Pearce, 2015).

Thus far, demonstrations of the redundancy effect have involved a blocking treatment where the trials with A+ and AX+ were intermixed throughout the same stage of training. A more common methodology for investigating blocking has been to present the trials with A+ prior to those with AX+ (e.g. Kamin, 1965). The purpose of the present experiments was to determine if the redundancy effect can be found with this different methodology for demonstrating blocking. Not only will the experiments permit the circumstances in which the redundancy effect can be observed to be explored further but, as we shall see, the experiments also allow for a test of the explanation offered by Vogel and Wagner (2017) for this effect. The overall design of the experiments was based on a predictive learning task that was used with humans by Uengoer et al. (2013).

Experiment 1

There were two groups in the first experiment, the design of which can be seen in Table 1. An experimental group, Group E, received trials in which the blocking cue was paired by itself with the outcome in Stage 1, A+, and accompanied by X when paired with the same outcome in Stage 2, AX+. The group also received trials with BY+ CY- in Stage 2. A control group, Group C, received similar training except that A was not presented during Stage 1, but the trials with A+ and AX+ were intermixed in Stage 2. On the basis of previous results (e.g. Uengoer et al. 2013), it was expected for Group C that during subsequent test trials, the response to X will be stronger than to Y. Of more interest are the results from Group E. If the redundancy effect can be found when the blocking treatment takes place in separate stages, then the test trials will again reveal a stronger response to X than Y. The additional cues shown in the table were included in order to ensure that both groups received

the same pattern of trials with and without an outcome in both stages, and to ensure that the number of cues present on a trial did not indicate the outcome of the trial.

Table 1 about here

In order to determine the predictions concerning the experiment, from the account of the redundancy effect offered by Vogel and Wagner (2017), a computer simulation was conducted based on the designs for the two groups shown in Table 1. The simulation was based on Equation 1, with the parameter values used by Vogel and Wagner mentioned above, and with a single common cue, K, present on every trial. The upper panel of Figure 1 shows the predicted overall associative strength of X and Y during the second stage of training for Group C. These values reflect the sum of the predicted values of X and K for cue X, and of Y and K for cue Y. It is apparent that as training progresses the overall associative strength of X is predicted to be stronger than of Y. That is, the results for Group C are predicted to provide a further demonstration of the redundancy effect. The lower panel of Figure 1 shows the equivalent predictions for X and Y for Group E. On this occasion, the response to X is predicted to be weaker than to Y, and the redundancy effect should not be observed. In essence, despite the assumed presence of K, the pretraining with A+ is predicted to result in cue X gaining very little associative strength in Stage 2, so that the response to X is predicted to be determined almost entirely by the associative strength of K. If this pattern of results should be found, not only will it lend support to the proposals of Vogel and Wagner, it will also point to an important boundary condition under which the redundancy effect might not be observed.

Figure 1 about here

Method

Participants. Seventy-two students of Philipps-Universität Marburg (of which 46 were females) participated in the experiment and received either course credit or payment.

Their age varied between 17 and 30 years, with a median of 22.5 years. Participants were randomly allocated to the two experimental groups and were tested individually. They gave informed written consent to participate in the experiment. The experimental procedure was approved by the ethics committee of the Psychology Department of the Philipps-Universität Marburg.

Apparatus and procedure. Instructions, stimuli, and further necessary information were presented on a computer screen. Participants responded by using a computer mouse. Pictures of the following foods served as cues: apple, banana, broccoli, carrot, cherry, grapes, kiwi, lemon, orange, pear, and pineapple. Assignment of foods to cues was randomized for each participant. Cues were followed by the occurrence of stomach ache (+) or by its absence (-). Participants were initially asked to read the following instruction (in German):

"This study is concerned with the question of how people learn about relationships between different events. In the present case, you should learn whether the consumption of certain foods leads to stomach ache or not. Imagine that you are a medical doctor. One of your patients often suffers from stomach ache after meals. To discover the foods the patient reacts to, your patient eats specific foods and observes whether stomach ache occurs or not. The results of these tests are shown to you on the screen one after the other. You will be told what your patient has eaten. Please look at the foods carefully. Thereafter you will be asked to predict whether the patient suffers from stomach ache. For this prediction, please click on the appropriate response button. After you have made your prediction, you will be informed whether your patient actually suffered from stomach ache. Use this feedback to find out what causes the stomach ache your patient is suffering from. Obviously, at first you will have to guess because you do not know anything about your patient. But eventually you will learn which foods lead to stomach ache in this patient and you will be able to make correct

predictions. For all of your answers, accuracy rather than speed is essential. Please do not take any notes during the experiment. If you have any more questions, please ask them now. If you do not have any questions, please start the experiment by clicking on the Next button."

Each training trial featured one or two food item pictures shown in the centre of the screen on a black background. In the case of two food pictures, they appeared side by side, with the left-right allocation determined randomly on each trial. The sentences "The patient ate the following food(s)" and "Which reaction do you expect?" were presented above and below the food(s), respectively. Participants made their predictions by clicking one of two response buttons shown side by side on the bottom half of the screen. The button on the left was labelled "no stomach ache", and the one on the right "stomach ache". Immediately after they responded, a feedback window appeared in the centre of the screen informing the participant whether the patient actually suffered from stomach ache or not. After clicking on the feedback window, the next trial started.

Stage 1 consisted of 40 trials. Ten of these trials consisted of A+ in Group E, and Q+ in Group C. The remaining trials in both groups consisted of ten trials with each of P-, EF+, and GH-. Stage 2 comprised 60 trials. Ten with Q+ in Group E, and ten with A+ in Group C, as well as ten trials with each of AX+, BY+, CY-, P-, and GH- in both groups. The trials of each stage were divided into five blocks with two presentations of each trial type in each block. The order of trials was determined randomly for each block and participant. After participants completed Stage 2, they received a series of test trials. The test was introduced with the following instruction:

"Now, your task is to judge the probability with which specific foods cause stomach ache in your patient. For this purpose, single foods will be shown to you on the screen. In this

part, you will receive no feedback about the actual reaction of the patient. Use all the information that you have collected up to that time."

Each test trial comprised one food picture shown in the centre of the screen. Above the food picture, the question "What is the likelihood that the food causes stomach ache?" was presented. Participants gave their ratings using a scale ranging from 0 (certainly not) to 10 (very certain). The rating scale was presented in the bottom half of the screen. The 11 values of the rating scale appeared side by side and participants chose one value by clicking on it. After participants confirmed their choice by clicking on an OK button presented below the rating scale, the next test trial started. Participants did not receive any feedback during this stage. Each of the cues A, B, C, X, and Y appeared twice, in a random sequence. For each cue, the two ratings were averaged for data analysis.

Results and Discussion

For this and the subsequent experiment, the .05 level of significance was used in all statistical tests. Stated probability levels were based on the Greenhouse and Geisser (1959) adjustment of degrees of freedom where appropriate. We used partial eta squared (η_p^2) as the measure of effect size.

The left-hand panel of Figure 2 shows the mean percentage of trials on which the outcome was predicted to occur for cues A and Q in Groups E and C, respectively, across the five training blocks in Stage 1. Participants rapidly mastered the task, so that by the end of this stage both groups were performing with a high degree of accuracy; during the final block the mean percentage of trials on which the outcome was predicted correctly was 95.83% (SEM = 3.07) in Group E, and 94.44% (SEM = 2.66) in Group C. A Block (1 – 5) × Group

(E vs. C) ANOVA yielded a significant main effect of block, $F(4, 280) = 61.49$, $p < .001$, $\eta_p^2 = .47$, while the main effect of group and the interaction were not significant, both Fs < 1.

For ease of presentation, the results from the second stage of training are presented in separate panels. The panel in the centre of Figure 2 shows the course of acquisition of the simple discrimination, BY+ CY-, for the two groups. In support of the observation that the discrimination was acquired at a similar rate by the two groups, a Cue (BY vs. CY) × Block (1 – 5) × Group (E vs. C) ANOVA revealed a significant main effect of cue, $F(1, 70) = 411.33$, $p < .001$, $\eta_p^2 = .86$, and a significant Cue × Block interaction, $F(4, 280) = 57.89$, $p < .001$, $\eta_p^2 = .45$. The main effect of block did not reach significance, $F(4, 280) = 2.38$, $p = .08$. The main effect of group and the three interactions including this factor were not significant, all Fs < 1.91, all ps > 1.3.

Figure 2 about here

The right-hand panel of Figure 2 shows the results from the trials with Q+ AX+ in Group E and A+ AX+ in Group C across the five blocks in Stage 2. For the first block of this stage, it is noteworthy that as a result of the Stage-1 training, the response to AX was stronger in Group E than in Group C, but as training progressed, this group difference vanished. For AX trials, a Block (1 – 5) × Group (E vs. C) ANOVA revealed a significant main effect of block, $F(4, 280) = 8.99$, $p < .001$, $\eta_p^2 = .11$, and a Block × Group interaction, $F(4, 280) = 3.56$, $p = .02$, $\eta_p^2 = .05$. The main effect of group was not significant, $F(1, 70) = 2.86$, $p = .095$. Unpaired t tests showed stronger responding to AX in Group E than in Group C for Block 1, $t(70) = 2.54$, $p = .03$, but not for Block 5, $t < 1$ (p-values were adjusted for two comparisons according to Benjamini & Hochberg, 1995). A Block × Group ANOVA comparing the results from the trials with Q in Group E and those with A in Group C yielded a significant main effect of block, $F(4, 280) = 57.52$, $p < .001$, $\eta_p^2 = .45$, while the main effect of group and the interaction were not significant, both Fs < 1.

The results from the test trials for the two groups are displayed in Figure 3. In contrast to the predictions derived from the proposals of Vogel and Wagner (2017), both groups displayed the redundancy effect by responding with a higher rating to X, the blocked cue, than Y, the irrelevant cue from the simple discrimination. Not surprisingly, given the nature of the training, both groups also awarded a high mean rating to A and B, and a low mean rating to C. The results also suggest that overall the ratings were higher for Group C than Group E. For the informative cues A, B, and C, a Cue × Group ANOVA revealed significant main effects of cue, $F(2, 140) = 293.28$, $p < .001$, $\eta_p^2 = .81$, and of group, $F(1, 70) = 4.73$, $p = .03$, $\eta_p^2 = .06$; the interaction was not significant, $F < 1$.

For the blocked cue X and the irrelevant cue Y, a Cue × Group ANOVA yielded a significant main effect of cue, $F(1, 70) = 38.82$, $p < .001$, $\eta_p^2 = .36$, indicating the redundancy effect; the main effect of group, $F(1, 70) = 2.21$, $p = .14$, and the interaction, $F < 1$, were not significant. Given the specific within-group predictions, we also confirmed that ratings were higher for X than Y in Group E, $t(35) = 3.52$, $p < .01$, and in Group C, $t(35) = 5.41$, $p < .01$ (p-values were adjusted for two comparisons according to Benjamini & Hochberg, 1995).

Figure 3 about here

The higher rating for X than Y in Group E, constitutes the first demonstration of the redundancy effect when the blocking treatment, A+ AX+, takes place in separate stages with trials with A+ preceding rather than being intermixed with trials with AX+. Before exploring the theoretical significance of this finding, a second experiment is described which was, in part, conducted in order to determine the reliability of the above findings.

Experiment 2

According to Equation 1, blocking is a consequence of cues competing for a limited pool of associative strength. Given training of the kind A+ AX+, the greater the associative

strength of A, the weaker will be the associative strength of X. It is possible that the failure to observe a weak response to X, relative to Y in both groups of Experiment 1 was due to participants failing to take account of the presence of A when learning took place about X in Stage 2. If this were the case, then blocking with X would not be expected, and it could be argued that the results from Experiment 1 do not challenge the Rescorla-Wagner (1972) theory, because participants did not approach the stimuli in the manner specified by the theory. One purpose of the second experiment, which was based on the design for Group E in Experiment 1, was to test this argument. A single group of participants received trials with A+ and D- in Stage 1 followed by trials with two compounds paired with the same outcome, AX+ and DZ+, in Stage 2 (see Table 2). If learning about individual cues in Experiment 1 is unaffected by the cues they are paired with, then the ratings awarded to X and Z at the end of the present experiment should be similar. Given such an outcome, the redundancy effect observed in Experiment 1 might not then be regarded as a serious challenge to the Rescorla-Wagner theory.

Table 2 about here

Experiment 2 was also used as an opportunity to confirm the reliability of the redundancy effect that was observed in Group E of Experiment 1. It is evident from Table 2 that Stage 2 of the experiment not only contained trials with AX+ and DZ+, but also with BY+ CY-. On the basis of the results from the first experiment, it was anticipated that subsequent test trials would reveal that the ratings awarded to X were higher than to Y. The remaining cues specified in Table 2 were included for the same reasons as their counterparts in Experiment 1.

Figure 4 shows the results of a computer simulation, based on the same principles as those adopted for the simulation in Experiment 1, for the single-group design summarised in Table 2. That is the simulation was based on Equation 1 with the assumption that a common

cue, K, was present on every trial. The simulation revealed that the response to the irrelevant cue from the simple discrimination, Y, is predicted to be stronger than to X. Thus, despite the changes in design from Experiment 1, the simulation again predicts the opposite of the redundancy effect. In addition, the response to X is predicted to be weaker than to Z, which is consistent with the assumption that cues compete for a limited pool of associative strength.

Figure 4 about here

Method

Participants, apparatus, and procedure. A group of 32 students of Philipps-Universität Marburg (17 males), whose age varied between 19 and 30 years (median of 22 years), participated in Experiment 2. They were tested in the same way and received the same recompense as the participants of Experiment 1.

The apparatus and procedure of Experiment 2 were the same as those used in Experiment 1, unless stated otherwise. In addition to the eleven food pictures used in Experiment 1, the following food pictures served as cues in Experiment 2: cheese, corn, lettuce, melon, plum, strawberry, and tomato. Stage 1 comprised ten trials with each of the trial types A+, D-, EF+, EG-, HI+, and JK-. Stage 2 consisted of ten trials with each of the trial types AX+, DZ+, BY+, CY-, LM-, and NO-. The test featured two presentations of each of the cues A, D, X, Y, and Z.

Results and Discussion

The left-hand panel of Figure 5 shows the mean percentage of trials on which the outcome was predicted to occur for cues A and D across the five blocks in Stage 1. By the end of this stage, the A+ D- discrimination had been solved. In the final block of trials, the mean percentage of trials on which the outcome was predicted correctly was 96.88% (SEM = 3.13) for cue A and 90.63% (SEM = 4.16) for cue D.

The results for the trials with AX+, DZ+, BY+, and CY- in Stage 2 are displayed in the right-hand panel of Figure 5. By the end of this stage, it was clear that participants were again predicting the correct outcome of the trials with a high degree of accuracy; during the final block of training the mean percentage of correct predictions was 96.88% (SEM = 2.17) for AX, 96.88% (SEM = 3.13) for DZ, 93.75% (SEM = 4.35) for BY, and 87.5% (SEM = 5.94) for CY-.

Figure 5 about here

The results from the test trials are presented in Figure 6. The high rating for A reflects the pairing of this cue with the outcome in Stage 1, and the low rating for D, likewise reflects the pairings of this cue with the absence of the outcome in Stage 1. The influence of these different treatments can be seen in the considerably higher rating given to Z than to X, as a consequence of the trials with AX+ and DZ+ in Stage 2. The difference between the ratings to X and Z was significant, $t(31) = 5.59$, $p < .001$, which confirms that learning about these stimuli was influenced by the cues that accompanied them during Stage 2. The rating for X was higher than for Y, which provides a further demonstration of the redundancy effect. This difference was significant, $t(31) = 2.14$, $p < .05$. The two p-values were corrected for multiple comparisons according to Benjamini and Hochberg (1995).

Figure 6 about here

General Discussion

The two reported experiments demonstrate for the first time that the redundancy effect can be observed when training with the blocking cue by itself, A+, takes place prior to the trials in which the blocking cue is presented in compound with the blocked cue, AX+. The experiments thus demonstrate both the reliability and the generality of the redundancy effect.

The principal concern of the present experiments was to evaluate an explanation for the redundancy effect based on the Rescorla-Wagner (1972) theory. It has been pointed out on several occasions (Jones & Pearce, 2015; Pearce et al, 2012; Uengoer et al., 2013) that the theory predicts the response to a blocked cue, X, after a blocking treatment, A+ AX+, will be weaker than to the irrelevant cue, Y, from a simple discrimination, BY+ CY-. In order to enable the Rescorla-Wagner theory to come to terms with the redundancy effect, Vogel and Wagner (2017) suggested it is important to acknowledge that the experimental cues might be similar to each other, which can be captured with the assumption that any cue, say A, is in fact a compound, AK, composed of a unique component, A, and a component shared with other cues, K. Although this approach enables the Rescorla-Wagner theory to predict the redundancy effect when the trials for the blocking treatment are intermixed throughout training, computer simulations revealed that the theory does not predict the redundancy effect when training with the blocking cue, A+, precedes the blocking treatment, AX+. The present experiments thus suggest that the proposals of Vogel and Wagner fall short when it comes to accounting for all demonstrations of the redundancy effect.

Having said that, we must acknowledge that the computer simulations used to derive predictions from the proposals of Vogel and Wagner (2017) were restricted to the parameter values that they employed. In order to determine whether there is something special about these values, additional simulations were conducted, with two constraints. First, the value of β for trials with an outcome was greater than for trials without an outcome. As noted in the Introduction, this relationship is required if the Rescorla-Wagner (1972) theory is to explain the relative validity effect, which has been demonstrated in both animals (e.g. Wagner et al., 1969) and humans (e.g. Uengoer, Lachnit, & Pearce, 2019). The second constraint was that sufficient Stage-1 trials were employed to enable the predicted associative strength of A to

reach asymptote. Such a stipulation is necessary if A is to be regarded as an effective cue for blocking.

Simulations incorporating the above two conditions with a wide range of parameter values consistently revealed that the training with Group E of Experiment 1, and with the single group of Experiment 2, will result in the opposite of the redundancy effect. If, however, either of the above two conditions is violated, then the redundancy effect can be predicted, but in some cases the effect is transient, and of small magnitude. It would appear, therefore, that in plausible circumstances the proposals of Vogel and Wagner (2017) are unable to explain the results from both experiments.

Vogel and Wagner (2017) derived from their proposals an intriguing prediction concerning the redundancy effect. They considered two different training schedules for two different groups. What will be referred to as the lean schedule comprised trials with A+ AX+ BY+ CY- G- H- I- J-, while a rich schedule comprised trials with A+ AX+ BY+ CY- G+ H+ I+ J+. These different schedules were intended to result in K having a higher associative value in the rich than the lean schedule, and thereby make the redundancy effect more likely to be evident in a group trained with the rich than the lean schedule. Jones et al. (2019) tested this prediction using a similar task to that used for the present studies and found evidence of the redundancy effect with both schedules. This outcome thus joins the present studies by posing a further challenge to the proposals of Vogel and Wagner.

As an alternative explanation for the redundancy effect, when observed with humans, Jones et al. (2019) pointed to the manner in which information about the ratings of individual cues was collected. In their experiment, which employed a similar methodology to the present design, participants were asked during testing to judge the probability with which specific foods cause stomach ache. They then assigned a value between 0 and 10 to indicate their degree of certainty that stomach ache was caused by the food under consideration. If a

participant was certain that this outcome would not occur, a score of 0 was expected, whereas if they were certain that stomach ache would occur then the score expected was 10. On this scale a score of 5 was said by Jones et al. to be ambiguous. Such a score might mean that participants were reasonably confident the food would cause stomach ache, or it might mean that participants were uncertain about whether or not the outcome was associated with the food under scrutiny. They further argued that with the blocking treatment, A+ AX+, participants would be uncertain about the causal significance of X and award it a score of 5. In addition, because Y occurs on trials without an outcome with BY+ CY- training, they will be confident that it is not a cause of illness and give it a rating close to zero. Thus the explanation for the redundancy effect in humans is that rather than reflect a difference in the associative strength of X and Y, it reflects a difference in the degree to which participants are confident about the causal properties of X and Y. To test this analysis, participants were asked how confident they were about the accuracy of the ratings awarded to the different stimuli, and it was found that they were more confident about those for Y than X. Of course, as the authors acknowledge, an alternative possibility is that confidence ratings are derived from the associative properties of each cue. On this basis, a score of 5 on the rating scale would reflect the cue has an intermediate associative strength, and thus lead the participant to be uncertain as to whether it is a signal for the presence or the absence of the outcome.

Turning to the present study, the results from the test trials with D in Experiment 2 might be of some relevance to the proposals of Jones et al. (2019). D was presented by itself and did not signal the outcome in Stage 1, D-, whereas in Stage 2 it was accompanied by a novel cue Z, and paired with the outcome, DZ+. It might be thought that when presented with D by itself for the test trials, participants would be uncertain about its significance and award it a rating close to 5, but this was not the case. Instead, the rating for D was low, and significantly lower than that to X, $t(31) = 2.21$, $p = .03$. Furthermore, Jones et al. proposed

that participants solve the BY+ CY- discrimination by reasoning that Y and C are equally non-causal of the outcome. In contrast to this proposal, we observed during the test in Experiment 1 that Y received a significantly higher rating than C, $t(71) = 3.95$, $p < .001$. In view of these findings, and the ambiguity over the interpretation of the confidence ratings recorded by Jones et al., it would seem appropriate to seek an alternative explanation for the redundancy effect to the one that they proposed.

An obvious alternative is that the redundancy effect reflects the stronger associative strength of X, after A+ AX+, than of Y after BY+ CY-. Given this conclusion, it then becomes necessary to explain how these different treatments result in X gaining more associative strength than Y. According to Equation 1, the change in associative strength of a single cue is determined by the discrepancy between the combined associative strength of all the cues present on a trial, V_T, and the asymptote of learning, λ, which leads it to predict the opposite of the redundancy effect. An alternative possibility is to assume that learning is governed in the manner depicted in Equation 2, where the discrepancy between the associative strength of a single cue and λ determines the change in associative strength of that cue.

$$\Delta V_A = \alpha_A.\beta.(\lambda - V_A) \qquad 2$$

This equation readily predicts the redundancy effect because a blocked cue is always followed by the outcome, whereas the irrelevant cue from a simple discrimination is intermittently followed by the outcome. The obvious weakness with Equation 2, however, is that it fails to predict effects such as blocking. Given sufficient training trials, it follows from this equation that the associative strengths of both the blocked and the blocking cue will reach the same asymptotic value, λ. One strategy for overcoming this problem is to assume that when two or more stimuli are present on a trial, the presence of one will influence the amount of attention that is paid to the other. According to Mackintosh (1975), the cue that is

the best predictor of the outcome will gain in the attention it receives, while the poorer predictor will be paid progressively less attention. If α in Equation 2 reflects the amount of attention paid to the cue, then it follows that learning with the blocking cue will reach an asymptote of λ, while the asymptote of learning for the blocked cue will be less than this value, and be reached when α for the blocked cue is equal to zero and no further attention is paid to it. Turning now to a simple discrimination, BY+ CY-, as training progresses, Y will become a poorer predictor of the outcome on trials when it is accompanied by both B and C, with the consequence that attention will drop to Y and increase to B and C. According to this analysis, therefore, attention during training will drop to both the blocked cue, X, and the irrelevant cue Y. However, for so long as attention persists to these cues, learning about their relationship with the outcome will continue, and given the different reinforcement schedules associated with X and Y, the associative strength of X will ultimately be greater than of Y.

A simple way in which these ideas can be expressed formally is to identify on each trial the cue that is the best predictor of the outcome, that is the cue whose associative strength is closest to the asymptote of learning, λ, set by the outcome. Attention to this cue, as represented by the value of α ($0 < \alpha < 1$), could then be increased according to Equation 3a, while attention to any other cue could be reduced according to Equation 3b.

$$\Delta\alpha = \theta \,.\, (1 - \alpha) \qquad 3a$$

$$\Delta\alpha = \theta \,.\, (0 - \alpha) \qquad 3b$$

The rate at which changes in attention take place is determined by the value of θ ($0 < \theta < 1$) and, once they have taken place, any changes in attention will influence subsequent changes in associative strength according to Equation 2. A series of computer simulations based on Equations 2, 3a and 3b revealed that the redundancy effect is correctly predicted for both groups of Experiment 1, and for the single group of Experiment 2. Moreover, this

prediction holds true for a wide range of starting values of α, as well as for a wide range of values for θ and β.

In Experiment 1, Group E received a sequential blocking design, A+ in Stage 1 and AX+ in Stage 2, whereas Group C received a simultaneous blocking treatment with both A+ and AX+ trials taking place in Stage 2. At the outset of Stage 2, therefore, the associative strength of A should be greater in Group E than C, and it might be thought that this difference would result in blocking being more effective in Group E than Group C. According to the above account, however, this difference between the associative properties of A in the two groups should have little impact. The first AX trial in Group E, will result in a decline in attention to X that will be effective from the second AX+ trial onwards, but a similar drop in attention in Group C will follow closely behind. All that is needed is for one A+ trials to be experienced by this group for there to be a decline in attention to X on the subsequent AX+ trial. Given that the rate of decline is given by Equation 3b, it follows that the loss of attention to X will be similar in both groups, with the decrement in Group C lagging no more than a few trials behind Group E. Given this analysis, it would be unlikely for a statistically significant difference between the groups to be evident when X is presented by itself for testing, which is in keeping with the findings of Experiment 1.

The foregoing analysis is rather different, and perhaps simpler, to the proposals put forward by Mackintosh (1975) for determining changes in attention during associative learning. He proposed that the magnitude of the change in attention to a cue was related to how well it predicted the outcome on the trial in question, V_A, relative to how well all the other cues combined predicted the outcome, V_O. Le Pelley (2004, see also Le Pelley, Mitchell, Beesley, George & Wills, 2016) presented this proposal in the form of Equation 4, which enables precise calculations of changes in attention to be determined on a trial by trial basis.

$$\Delta\alpha_A = \theta \,.\, (|\lambda - V_O| - |\lambda - V_A|) \qquad 4$$

In fact, when Equation 4, together with Equation 2 is applied to the present experiments, the correct outcomes can be predicted, but only in restricted circumstances. By way of example, the redundancy effect can be predicted for both Group E and Group C of Experiment 1, and for Experiment 2, if the starting value of α is .6, the value of β is .4, and the value of θ is .2. If the value of θ is increased to .4, say, or the starting value of α is reduced to .2, then the opposite of the redundancy effect is predicted when the trials for Group E in Experiment 1, or the single group of Experiment 2 are considered. Given this dependency on a restricted set of parameter values if the theory of Mackintosh (1975) is to explain the present results, it may be worth exploring further the possibility that changes in attention during predictive learning are controlled by Equations 3a and 3b, rather than by Equation 4.

Support for the proposal that the associability of a blocked cue will ultimately be less than of the blocking cue during a blocking treatment can be found in the results from a number of human learning experiments (Kruschke & Blair, 2000; Le Pelley, Beesley & Griffiths, 2014; Le Pelley, Beesley, & Suret, 2007; Luque, Vadillo, Gutiérrez-Cobo, & Le Pelley, 2018; Uengoer, Dwyer, Koenig, & Pearce, 2019). To our knowledge, however, there is no direct evidence with humans supporting the additional proposal made above that the associability of the redundant cue from a simple discrimination (BY+ CY-) will ultimately be lower than of the other cues, especially the one that signals the outcome (but see Dopson, Esber and Pearce, 2010, for evidence with pigeons). Results suggesting that such a difference might be found can be seen in the learned predictiveness effect, which is observed with training of the kind, AX+ AY+ BX* BY*. In this task, cues A and B reliably signal outcomes + and * respectively, while X and Y can be said to be uninformative, or redundant, in this respect. Subsequent testing has revealed that the associability of the relevant cues is

greater than of the irrelevant cues (Le Pelley & McLaren, 2003; Livesey, Thorwart, De Fina, & Harris, 2011; Lochman & Wills, 2003; see Feldmann-Wüstefeld, Uengoer, & Schubö, 2015, for evidence with neurophysiological markers of attentional allocation). Given the similarity between this training and a simple discrimination, it is tempting to speculate that the associability of relevant cues from a simple discrimination will likewise be greater than of the irrelevant cues, but this remains to be confirmed.

The account of attentional changes captured by Equations 3a and 3b can also be applied successfully to a phenomenon described by Uengoer, Lachnit, and Pearce (2019) as the outcome ratio effect. Participants were presented with two simple discriminations that differed in their outcome ratio. One discrimination was of the form 3AX+ BX-, while the other was CY+ 3DY-. Theories such as Rescorla and Wagner (1972) predict these treatments will result in no difference in the response to X and Y during test trials, whereas the rating given to X was found to be greater than to Y. Moreover, this prediction remains true, even when a common cue is assumed to be present on every trial, as envisaged by Vogel and Wagner (2017). In contrast, it follows from Equations 3a and 3b that during training attention to both X and Y will fall as training progresses but the richer reinforcement schedule associated with X than Y will, according to Equation 2, result in the former gaining more associative strength than the latter.

The present article is not the first occasion on which it has been suggested that the redundancy effect can be understood by referring to attentional processes. To explain their demonstration of this effect, Uengoer et al. (2013) suggested that the blocking treatment resulted in more attention being paid to the blocked than the irrelevant cue throughout training, and thereby enabled the former to acquire more associative strength than the latter. However, experiments to test this proposal, either by comparing the associability of the two cues (Uengoer, Dwyer, Koenig, & Pearce, 2019), or by comparing the time spent looking at

them (Jones & Zaksaite, 2018), failed to find any evidence of more attention being paid to the blocked than the irrelevant cue. In the present interpretation, attention is assumed to be similar to both cues throughout the experiment, and to fall off as training progresses. It is then the different reinforcement schedules associated with the two cues that is responsible for the blocked cue ultimately gaining more associative strength than the irrelevant cue.

In contrast to the foregoing support for the assumptions on which Equations 3a and 3b are based, additional findings suggest that our proposals will not provide a complete account of the factors that govern changes in the associability of cues during human learning. Livesey et al. (2011) for example, have shown that the associability of two cues repeatedly presented together will be greater if they are consistently followed by the same outcome (AB+, AB+) than if they are followed unpredictably by two different outcomes (AB+, AB*). Given that the two cues would be expected to gain the same associative strength within each treatment, according to Equations 3a and 3b, there should be no opportunity for either treatment to result in a change in associability. In a rather different study, Le Pelley, Turnbull, Reimers, Knipe, and Murphy (2010) have shown that the associability of a single cue paired repeatedly with the same outcome is ultimately greater than for a cue repeatedly paired with different outcomes. Such a result would not be expected if changes in associability are restricted to training in which two cues are presented together. To complicate matters further, Uengoer and Lachnit (2012) have shown that the associability of cues can be modified if they are used for a biconditional discrimination which, again, would not be expected according to the principles underlying Equations 3a and 3b (see also Livesey, Don, Uengoer, & Thorwart, 2019).

Given such a complex pattern of results, it would be unreasonably optimistic to expect the simple explanation offered for the present results to provide a complete account for all the changes in stimulus associability that have been recorded in studies of human

learning. Nonetheless, it is conceivable that Equations 3a and 3b capture one set of circumstances that result in a change in associability in human learning, and that these changes influence the ultimate associative strength of a cue in the manner determined by Equation 2.

When discussing the role of relative validity in associative learning, Wagner (1969, see also Wagner, 2003) identified himself as a modified continuity theorist. Such a stance led him to advocate that associative learning was governed by the extent to which the outcome of a trial was surprising, as defined by the Rescorla-Wagner (1972) equation. He also distanced himself from modified non-continuity theorists such as Mackintosh (1965, 1975), who assumed that a process akin to attention can influence the course of associative learning. The theoretical proposals put forward in the present article clearly belong to the second of these schools of thought. Time will tell if an alternative set of proposals can be developed to explain the redundancy effect, the roots of which are based firmly in the theorising of Allan Wagner, and which would belong more to the continuity than the non-continuity school of thought.

Acknowledgements

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 316803389 – SFB 1280.

References

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate–A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B, 57*, 289-300.

Brandon, S. E., & Wagner, A. R. (1998). Occasion setting: Influences of conditioned emotional responses and configural cues. In N. A. Schmajuk & P. C. Holland (Eds.), *Occasion setting: Associative learning and cognition in animals* (pp. 343-382). Washington, DC, US: American Psychological Association.

Dopson, J. C., Esber, G. R., & Pearce, J. M. (2010). Differences in the associability of relevant and irrelevant stimuli. *Journal of Experimental Psychology. Animal Behavior Processes*, *36*(2), 258–267. https://doi.org/10.1037/a0016588

Feldmann-Wüstefeld, T., Uengoer, M., & Schubö, A. (2015). You see what you have learned. Evidence for an interrelation of associative learning and visual selective attention. *Psychophysiology*, *52*(11), 1483–1497. https://doi.org/10.1111/psyp.12514

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, *117*(3), 227–247. https://doi.org/10.1037/0096-3445.117.3.227

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95–112. https://doi.org/10.1007/BF02289823

Haselgrove, M. (2010). Reasoning rats or associative animals? A common-element analysis of the effects of additive and subadditive pretraining on blocking. *Journal of Experimental Psychology. Animal Behavior Processes*, *36*(2), 296–306. https://doi.org/10.1037/a0016603

Jones, P. M., & Pearce, J. M. (2015). The fate of redundant cues: Further analysis of the redundancy effect. *Learning & Behavior*, *43*(1), 72–82. https://doi.org/10.3758/s13420-014-0162-x

Jones, P. M., & Zaksaite, T. (2018). The redundancy effect in human causal learning: No evidence for changes in selective attention. *Quarterly Journal of Experimental Psychology (2006)*, *71*(8), 1748–1760. https://doi.org/10.1080/17470218.2017.1350868

Jones, P. M., Zaksaite, T., & Mitchell, C. J. (2019). Uncertainty and blocking in human causal learning. *Journal of Experimental Psychology. Animal Learning and Cognition*, *45*(1), 111–124. https://doi.org/10.1037/xan0000185

Kamin, L. J. (1965). Temporal and intensity characteristics of the conditioned stimulus. In W. F. Prokasy (Ed.), *Classical conditioning: A symposium* (pp. 118-147). New York: Appleton-Century-Crofts.

Kamin, L. J. (1969). Selective attention and conditioning. In N. J. Mackintosh & W. K. Honig (Eds.), *Fundamental issues in associative learning* (pp. 42–64). Halifax, Nova Scotia, Canada: Dalhousie University Press.

Kruschke, J. K., & Blair, N. J. (2000). Blocking and backward blocking involve learned inattention. *Psychonomic Bulletin & Review*, *7*(4), 636–645. https://doi.org/10.3758/BF03213001

Le Pelley, M. E. (2004). The role of associative history in models of associative learning: A selective review and a hybrid model. *The Quarterly Journal of Experimental Psychology: Section B*, *57*(3), 193–243. https://doi.org/10.1080/02724990344000141

Le Pelley, M. E., Beesley, T., & Griffiths, O. (2014). Relative salience versus relative validity: Cue salience influences blocking in human associative learning. *Journal of*

Experimental Psychology: Animal Learning and Cognition, *40*(1), 116–132. https://doi.org/10.1037/xan0000006

Le Pelley, M. E., Beesley, T., & Suret, M. B. (2007). Blocking of human causal learning involves learned changes in stimulus processing. *The Quarterly Journal of Experimental Psychology*, *60*(11), 1468–1476. https://doi.org/10.1080/17470210701515645

Le Pelley, M. E., & McLaren, I. P. L. (2003). Learned associability and associative change in human causal learning. *The Quarterly Journal of Experimental Psychology: Section B*, *56*(1), 68–79. https://doi.org/10.1080/02724990244000179

Le Pelley, M. E., Mitchell, C. J., Beesley, T., George, D. N., & Wills, A. J. (2016). Attention and associative learning in humans: An integrative review. *Psychological Bulletin*, *142*(10), 1111–1140. https://doi.org/10.1037/bul0000064

Le Pelley, M. E., Turnbull, M. N., Reimers, S. J., & Knipe, R. L. (2010). Learned predictiveness effects following single-cue training in humans. *Learning & Behavior*, *38*(2), 126–144. https://doi.org/10.3758/LB.38.2.126

Livesey, E. J., Don, H. J., Uengoer, M., & Thorwart, A. (2019). Transfer of associability and relational structure in human associative learning. *Journal of Experimental Psychology. Animal Learning and Cognition*, *45*(2), 125–142. https://doi.org/10.1037/xan0000197

Livesey, E. J., Thorwart, A., Fina, N. L. de, & Harris, J. A. (2011). Comparing learned predictiveness effects within and across compound discriminations. *Journal of Experimental Psychology. Animal Behavior Processes*, *37*(4), 446–465. https://doi.org/10.1037/a0023391

Lochman, T., & Wills, A. J. (2003). Predictive history in an allergy prediction task. In F. Schmahlofer, R. M. Young, & G. Katz (Eds.), *Proceedings of EuroCogSci (pp. 217-222).* Mahwah, NJ: Erlbaum.

Luque, D., Vadillo, M. A., Gutiérrez-Cobo, M. J., & Le Pelley, M. E. (2018). The blocking effect in associative learning involves learned biases in rapid attentional capture. *Quarterly Journal of Experimental Psychology (2006)*, *71*(2), 522–544. https://doi.org/10.1080/17470218.2016.1262435

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*(4), 276–298. https://doi.org/10.1037/h0076778

Mackintosh, N. J., & Honig, W. K. (Eds.). (1969). *Fundamental issues in associative learning*. Halifax: Dalhousie University Press.

Pearce, J. M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review*, *101*(4), 587–607. https://doi.org/10.1037//0033-295X.101.4.587

Pearce, J. M., Dopson, J. C., Haselgrove, M., & Esber, G. R. (2012). The fate of redundant cues during blocking and a simple discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, *38*(2), 167–179. https://doi.org/10.1037/a0027662

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*(6), 532–552. https://doi.org/10.1037//0033-295X.87.6.532

Pearce, J. M., & Mackintosh, N. J. (2010). Two theories of attention: A review and a possible integration. In C. J. Mitchell & M. E. Le Pelley (Eds.), *Attention and associative*

learning: From brain to behaviour (pp. 11–39). Oxford, UK: Oxford University Press.

Rescorla, R. A. (2002). Comparison of the rates of associative change during acquisition and extinction. *Journal of Experimental Psychology: Animal Behavior Processes*, *28*(4), 406–415. https://doi.org/10.1037//0097-7403.28.4.406

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 theory and research* (pp. 64–99). New York, NY: Appleton-Century-Crofts.

Saavedra, M. A. (1971). Pavlovian compound conditioning. Unpublished doctoral dissertation. Yale University.

Thorwart, A., Schultheis, H., König, S., & Lachnit, H. (2009). ALTSim: A MATLAB simulator for current associative learning theories. *Behavior Research Methods*, *41*(1), 29–34. https://doi.org/10.3758/BRM.41.1.29

Uengoer, M., Dwyer, D. M., Koenig, S., & Pearce, J. M. (2019). A test for a difference in the associability of blocked and uninformative cues in human predictive learning. *Quarterly Journal of Experimental Psychology*, *72*(2), 222–237. https://doi.org/10.1080/17470218.2017.1345957

Uengoer, M., & Lachnit, H. (2012). Modulation of attention in discrimination learning: The roles of stimulus relevance and stimulus–outcome correlation. *Learning & Behavior*, *40*(2), 117–127. https://doi.org/10.3758/s13420-011-0049-z

Uengoer, M., Lachnit, H., & Pearce, J. M. (2019). The fate of redundant cues in human predictive learning: The outcome ratio effect. *Quarterly Journal of Experimental Psychology*, *72*(8), 1945–1960. https://doi.org/10.1177/1747021818820042

Uengoer, M., Lotz, A., & Pearce, J. M. (2013). The fate of redundant cues in human predictive learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *39*(4), 323–333. https://doi.org/10.1037/a0034073

Vogel, E. H., & Wagner, A. R. (2017). A theoretical note in interpretation of the "redundancy effect" in associative learning. *Journal of Experimental Psychology. Animal Learning and Cognition*, *43*(1), 119–125. https://doi.org/10.1037/xan0000123

Wagner, A. R. (1969). Stimulus validity and stimulus selection in associative learning. In N. J. Mackintosh & W. K. Honig (Eds.), *Fundamental issues in associative learning* (pp. 90-122). Halifax: Dalhousie University Press.

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

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

Wagner, A. R. (2003). Context-sensitive elemental theory. *The Quarterly Journal of Experimental Psychology: Section B*, *56*(1), 7–29. https://doi.org/10.1080/02724990244000133

Wagner, A. R., & Brandon, S. E. (2001). A componential theory of Pavlovian conditioning. In R. R. Mowrer & S. B. Klein (Eds.), *Contemporary learning: Theory and application* (pp. 23–64). Mahwah: Erlbaum.

Wagner, A. R., Logan, F. A., Haberlandt, K., & Price, T. (1968). Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology*, *76*(2, Pt.1), 171–180. https://doi.org/10.1037/h0025414

Wagner, A. R., & Rescorla, R A. (1972). Inhibition in Pavlovian conditioning: Application to a theory. In R. A. Boakes & M. S. Halliday (Eds.), *Inhibition and learning* (pp. 301-336). London: Academic Press.

Whitlow, J. W., & Wagner, A. R. (1972). Negative patterning in classical conditioning: Summation of response tendencies to isolable and configural components. *Psychonomic Science*, *27*(5), 299–301. https://doi.org/10.3758/BF03328970

Table 1

Design of Experiment 1

Group	Stage 1	Stage 2	Test
E	A+	Q+ AX+ BY+ CY-	A B C X Y
	P- EF+ GH-	P- GH-	
C	Q+	A+ AX+ BY+ CY-	A B C X Y
	P- EF+ GH-	P- GH-	

Table 2

Design of Experiment 2

Stage 1	Stage 2	Test
A+ D-	AX+ DZ+ BY+ CY-	A D X Y Z
EF+ EG- HI+ JK-	LM- NO-	

Figure 1

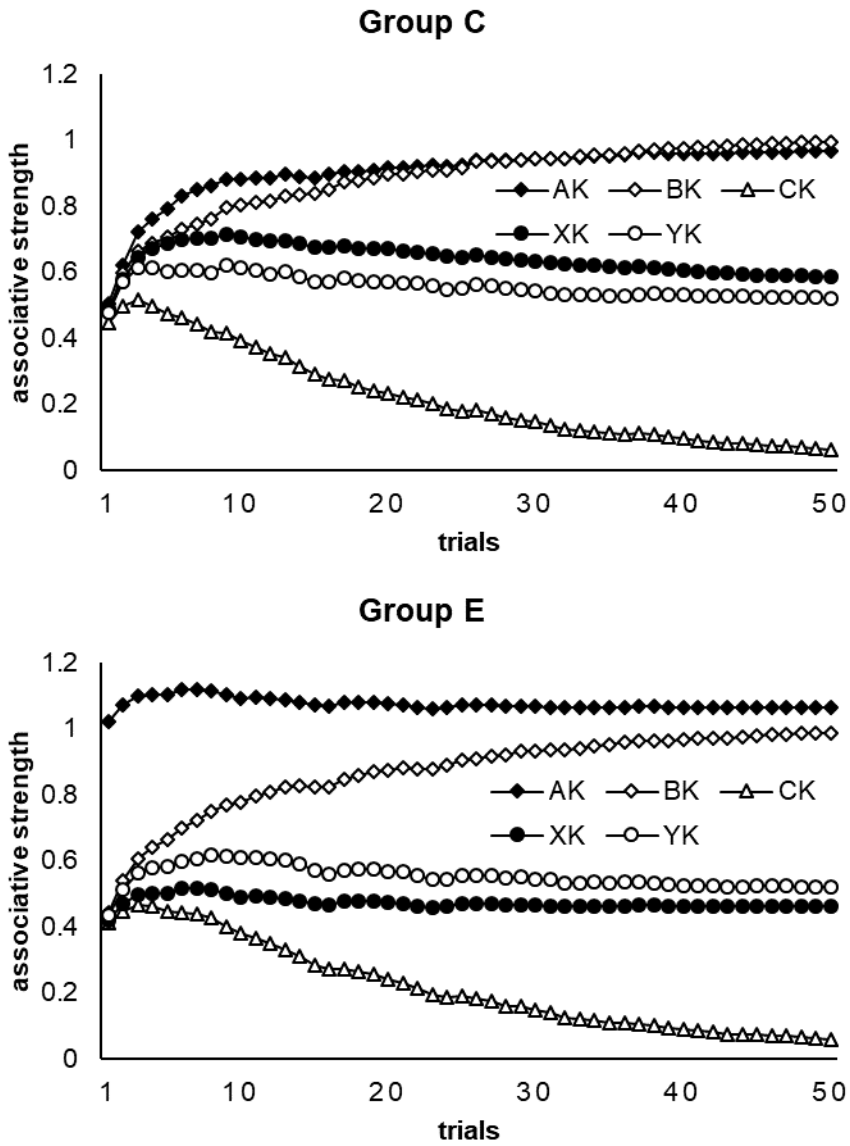

Figure 1. Computer simulation based on the Rescorla-Wagner (1972) equation under the common-cue assumption for the predicted course of associative strengths during Stage 2 in Group C (upper-panel) and Group E (lower-panel) of Experiment 1. Simulations were conducted using ALTSim (Thorwart, Schultheis, König, & Lachnit, 2009).

Figure 2

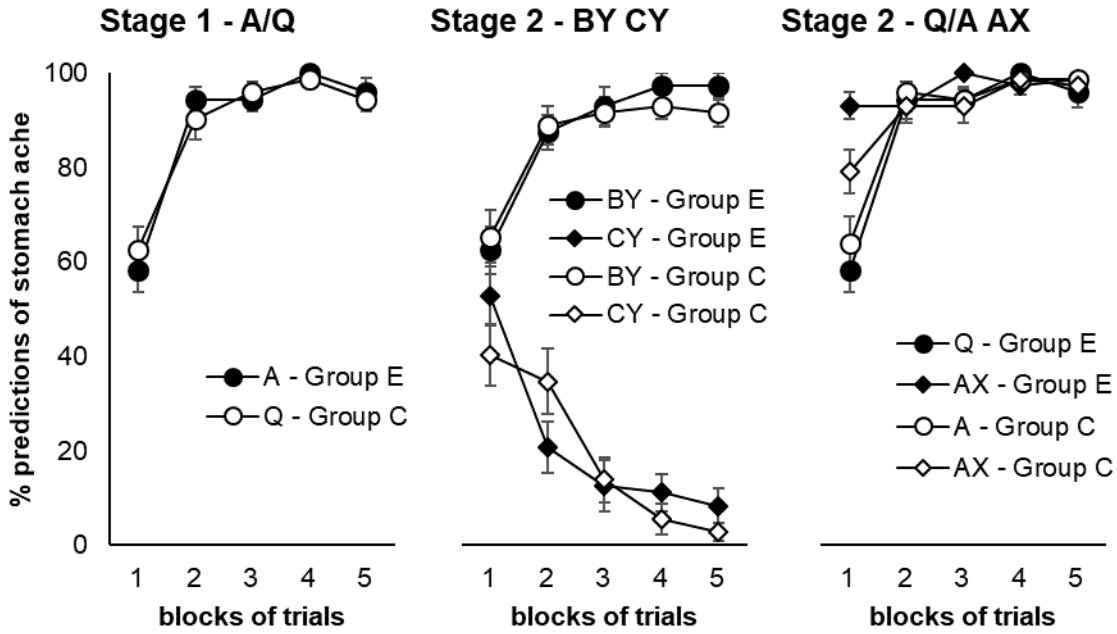

Figure 2. Mean percentages of outcome predictions across the five blocks in Stage 1 (left-hand panel) and the five blocks in Stage 2 (centre and right-hand panels) of Experiment 1. Error bars denote stand errors of the means.

Figure 3

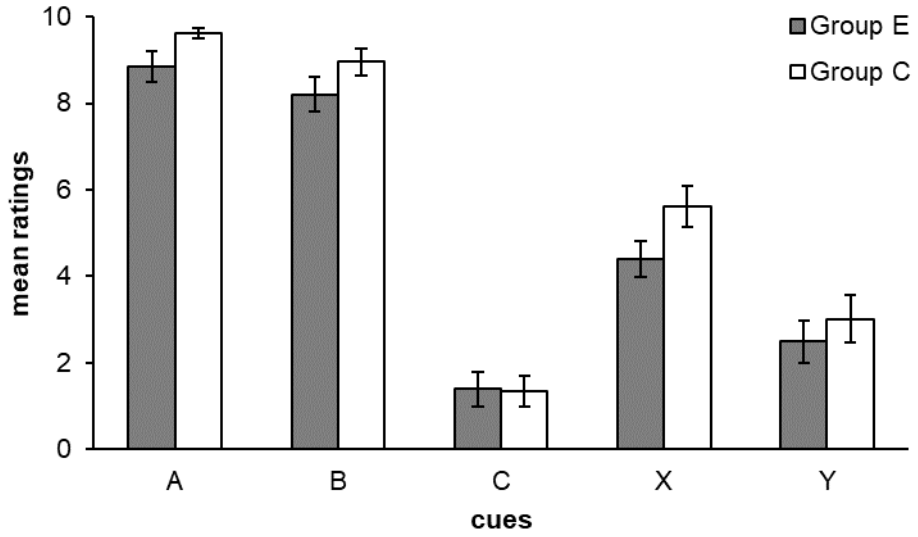

Figure 3. Mean ratings of the likelihood that individual cues would be followed by the outcome during the test in Experiment 1. Error bars denote stand errors of the means.

Figure 4

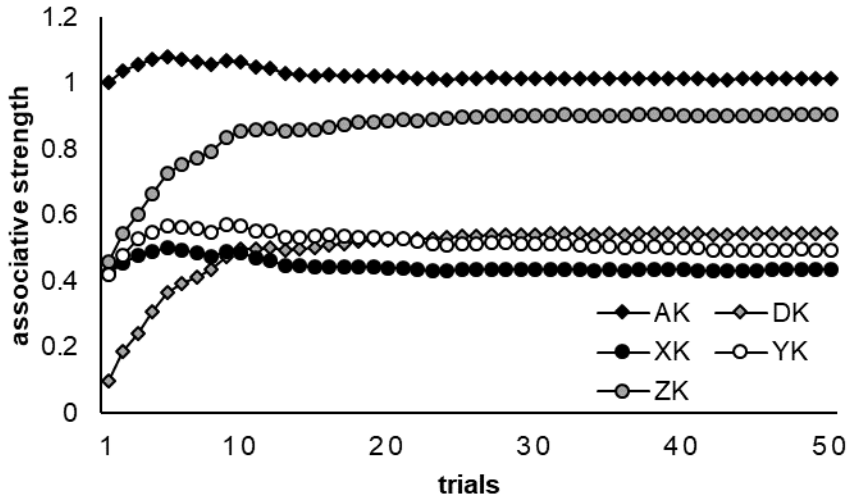

Figure 4. Computer simulation based on the Rescorla-Wagner (1972) equation under the common-cue assumption for the predicted course of associative strengths during Stage 2 of Experiment 2. Simulations were conducted using ALTSim (Thorwart et al., 2009).

Figure 5

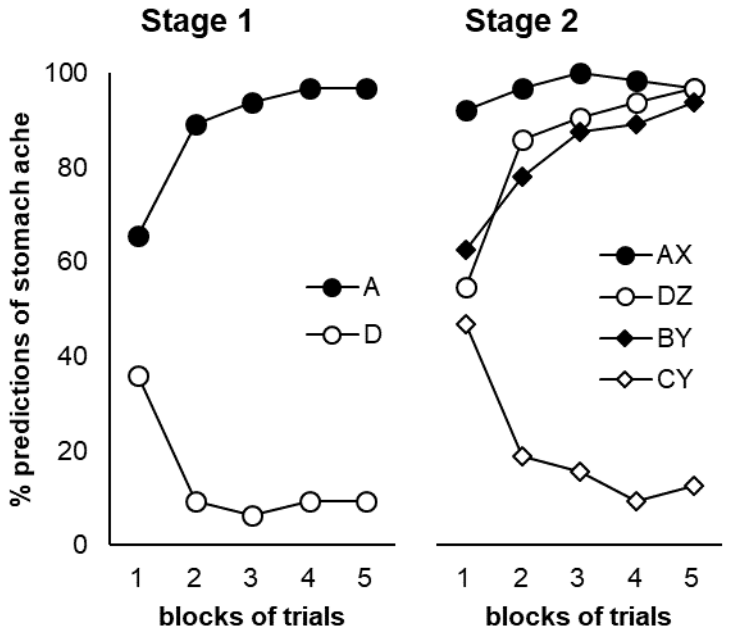

Figure 5. Mean percentages of outcome predictions across the five blocks in Stage 1 (left-hand panel) and the five blocks in Stage 2 (right-hand panel) of Experiment 2.

Figure 6

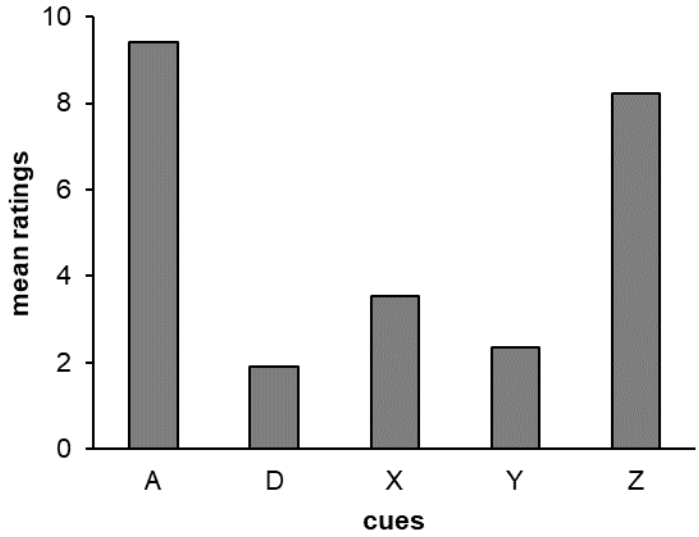

Figure 6. Mean ratings of the likelihood that individual cues would be followed by the outcome during the test in Experiment 2.