eman ta zabal zazu
UPV EHU

PSIKOLOGIA FAKULTATEA
FACULTAD DE PSICOLOGIA

Assessing the contribution of the top-down processing on the learned predictiveness effect

A strategy based on shortening the duration of the test trials

---

TRABAJO DE FIN DE GRADO

Alumna: Ariadna García Vergara

Tutor: Gabriel Rodríguez San Juan

Table of contents

1. Abstract

The *learned predictiveness* (LP) procedure has been interpreted as useful to discuss the mechanisms that regulate a form of attention related to the limited learning resources of the organisms. In this procedure, in Stage 1, half of the cues are established as accurate predictors of their outcomes, and the other half are established as poorer predictors. In Stage 2, all cues are equally predictive of a new outcome. On test, participants rate the likelihood that the cues would produce Stage 2 outcomes. It has been consistently found that participants rate the accurate predictors higher than poorer predictors. This pattern of results has been interpreted as supporting the notion that stimulus associability is directly related to predictive accuracy. We report an experiment investigating this LP effect in which we manipulated the maximum time to respond on each trial: 20 s vs. 4 s. Participants in the long-time period condition showed both a clear discrimination in Stage 1 and a LP effect on the test. Participants in the short-time period condition, however, showed a clear discrimination in Stage 1, but did not show the LP effect on the test. This suggests that the mechanism involved in the LP effect requires more time than the automatic associative mechanisms involved in discrimination learning. These results therefore invite to reconsider the importance of the LP effect in discussing the mechanisms regulating stimulus associability.

Key words: learned predictiveness; associative human learning; associability; inference

2. Introduction

Associative learning allows human and nonhuman animals to encode the relationship between environmental stimuli and their consequences. Possession and use of such information bestows advantages. Priority processing of stimuli that are good predictors of important consequences allows the efficient generation of suitable ways of responding to the predicted events. However, the attentional mechanisms responsible for this preparatory and response-generating function appear to coexist with other selective mechanisms that serve to optimize our learning resources. In particular, there is a need to learn about only stimuli that have unknown or uncertain consequences. Once the outcome of a stimulus has been fully established (in associative terms, once associative strength has reached an asymptote) it no longer needs to command attention, at least as far as further learning is concerned. This notion formed the basis of the theory proposed by Pearce and Hall (1980) which describes how the associability of a stimulus is determined. According to this theory, a stimulus with uncertain consequences (for example, a novel stimulus) has a high associability, (commands the resources necessary for learning) and is learned about readily. As a result, uncertainty about the outcome of the stimulus is reduced and, with it, stimulus associability (i.e., the rate to which the stimulus is learned about). Pearce and Hall adopted the terms used by Shiffrin and Schneider (1977), suggesting that initially a stimulus undergoes controlled processing but that, as its consequences become known, processing shifts to an automatic mode. Automatic processing is enough to ensure exploitation of the predictive information supplied by the stimulus but does not allow further learning about the stimulus.

The notion that the associability of a stimulus declines as it comes to predict its consequences has received substantial support from experiments using conditioning procedures with nonhuman animals as the subjects (e.g., Hall & Pearce, 1979, 1982; Pearce, Kaye, & Hall, 1982). But, as Le Pelley, Mitchell, Beesley, George, and Wills (2016) have pointed out, there is scant evidence for this proposal from studies of associative learning in humans (e.g., see Griffiths, Johnson, & Mitchell, 2011). Indeed, one procedure that has been widely used for the study of transfer effects in human learning, the *learned predictiveness procedure*, developed by Le Pelley and McLaren (2003; see also Lochmann & Wills, 2003, and Le Pelley et al., 2016, for an extensive review), has yielded quite contrary results to the notion of associability proposed by Pearce and Hall (1980). Next, this procedure is explained in more detail.

Although some of the studies that have exploited this procedure have used different cover histories, a similar general design has always been employed (see Table 1). In this general design, the letters represent cues (e.g., different food meals) and the numbers represent the outcomes than can occur after the presentation of the cues (e.g, the allergies that can suffer a fictitious patient after eating the food meals). So, for example, AW->1 indicates a sort of trial in which the cues A and W, are simultaneously presented, and followed by the occurrence of the outcome 1. The cues are presented in pairs across all the experiment. Participants receive first a sort of training in which one of two outcomes (1 or 2) can occur in each trial, and while one cue of each pair is consistently paired with an outcome (A and D are consistently paired with 1, and B and C are consistently paired with 2), the other cue is equally paired with both 1 and 2. So, participants are exposed to conditions in which they can learn that there are differences in the predictive accuracy of the cues (hence, we can refer to this sort of training as *differential training*).

Once that the initial training is finished, the background history presents a similar subsequent task, but mixing the same cues in different pairs and also substituting the occurrence of the outcomes 1 and 2 by two other different outcomes, 3 and 4. For example, in the study by Le Pelley and McLaren (2003), participants were told that they were attending different patients that suffered different sorts of allergic reactions (Mr. X in the first stage, and Mr. Y in the second stage). During this second stage of training, each pair of cues is formed by a previously-accurate predictor and a previously-inaccurate predictor. But, critically, in this stage the two cues of each pair are equally paired with the corresponding novel outcome (hence, we can refer to this sort of training as *non-differential training*).

The question of interest is whether the differences in predictive accuracy established during the initial differential training will influence the rate of learning about the cues in the non-differential training. This is assessed in a final test in which the same cues are again mixed in four different novel pairs, in such a way that two pairs are formed by cues paired with outcome 3 (AC and VX), and the two other pairs are formed by cues paired with outcome 4 (BD and WY). Two of these pairs (AC and BD) are formed by cues that were accurate predictors during the initial differential training, and the other two pairs (VX and WY) are formed by inaccurate predictors in that stage. It has been routinely found that in this final test, participants rate more likely the occurrence of the correct outcome in the presence of the pairs formed by the cues that were accurate predictors during the initial differential training (e.g., Le Pelley & McLaren,

2003; Lochmann & Wills, 2003; Rodríguez & Hall, 2019; Experiment 1). Le Pelley and McLaren (2003) interpreted this effect as indicating that the learning-rate (i.e., the associabilty) of the predictive cues had been enhanced by the first stage of training. An interpretation that is directly contrary to the account of associability change offered by the Pearce-Hall (1980) model, but supports an alternative view of stimulus associability in terms of the theory proposed by Mackintosh (1975), according to which the predictive accuracy is directly related to associability.

The existence of these two sorts of results (some giving support to Pearce and Hall and some others supporting the contrary notion proposed by Mackintosh) has encouraged theoretical attempts in order to integrate these two views of associability. These hybrid models conceive the Pearce-Hall's and Mackintosh's associability factors as interactive elements of a more complex attentional mechanism able to regulate the learning-rate (e.g., Esber & Haselgrove, 2010; Hall & Rodríguez, 2010; Le Pelley, 2004; Pearce & Mackintosh, 2010). However, it is possible that these theoretical steps have been made before to confirm that all the discussed results actually reflect associability effects. This could be the case of the learned predictiveness effect. In a recent study, Rodríguez and Hall (2020; Experiment 2) provided a demonstration of the learned predictiveness effect using a modified procedure (see Table 1) that precluded the explanation in terms of the account offered by Le Pelley and McLaren (2003) based on the notion of associability derived from Mackintosh (1975).

Standard learned predictiveness procedure (e.g., Rodríguez & Hall, 2019; Experiment 1 Le Pelley & McLaren, 2003)			
Differential training	*Test: O1 or O2?*	*Nondifferential training*	*Test: O3 or O4?*
AV→1	AV	AX→3	AC
BV→2	BV	BY→4	BD
AW→1	AW	CV→3	VX
BW→2	BW	DW→4	WY
CX→2	CX		
DX→1	DX		
CY→2	CY		
DY→1	DY		

Rodríguez & Hall (2020; Experiment 2)				
Nondifferential training	*Test: O3 or O4?*	*Differential training*	*Retest: O3 or O4?*	*Test: O1 or O2?*
AX→3	AC	AV→1	AC	AV
BY→4	BD	BV→2	BD	BV
CV→3	VX	AW→1	VX	AW
DW→4	WY	BW→2	WY	BW
		CX→2		CX
		DX→1		DX
		CY→2		CY
		DY→1		DY

Present experiment

Differential training	*Test: O1 or O2?*	*Nondifferential training*	*Test: O1? O2?*
AV→1	AV	AX→3	**Group 20:** test trials of 20 seconds
BV→2	BV	BY→4	AC
AW→1	AW	CV→3	BD
BW→2	BW	DW→4	VX
CX→2	CX		WY
DX→1	DX		**Group 4:** test trials of 4 seconds
CY→2	CY		AC
DY→1	DY		BD
			VX
			WY

Table 1. Each letter represents a cue (a foodstuff) and numbers represent outcomes (Os: allergic reactions).

The procedure used in this experiment (Rodríguez and Hall, 2020; Experiment 2) was essentially the same as that used by Le Pelley and McLaren (2003), except that the order of the training phases was reversed, with differential training being followed by non-differential training. With this arrangement, any difference in associability between accurate and inaccurate predictors produced by differential training with Outcomes 1 and 2 would be irrelevant given that learning about the relationship between the cues and Outcomes 3 and 4 had already occurred. The critical test was that given after differential training ("retest" in Table 1) in which it was observed, as in the case of the standard learned predictiveness effect, superior performance with AC and BD than with VX and WY.

This pattern of results shows that the learned predictiveness effect survives a manipulation in which the order of the differential and non-differential training phases is reversed. To the light of these results, Rodríguez and Hall (2020) concluded that learned predictiveness can be obtained when effects depending on changes in associability (of the sort envisaged by the theory of Mackintosh, 1975) cannot be responsible. However, these results do not directly falsify that, under the standard conditions, the effect is mediated by associability changes. The goal of the present study was to get more direct evidence against this notion.

Experiment

A relevant feature of the learned predictiveness effect is that, as in many other human learning procedures, the researcher does not estimate the strength of the target associations registering their assumed effects on the magnitude of an automatic and involuntary response. Instead, participants are asked to judge how likely is the occurrence of some given events, and they are often given a relatively long time to emit these judgments. Taking into account this feature of the procedure, Mitchel et al. (2012) have argued that the learned predictiveness effect might not be determined by an automatic, bottom-up, attentional mechanism of the sort

described previously for associability changes, but rather that it might be the consequence of an "inference based controlled attentional process." When the participants are faced with new outcomes in the non-differential phase of a learned predictiveness study (i.e., the second stage), they infer that the cues that were predictive in the prior differential training are likely to be predictive in the second.

The hypothesis that we want to test in this experiment is derived from this notion. If the learned predictiveness effect depends on a top-down inferential mechanism, rather than a bottom-up automatic associability mechanism, it should not survive to a severe limitation of the time available to respond on test. Given the cognitive load of the task, the inferential process alluded by Mitchel et. al (2012) may not be enough to produce the expected effect. Participants will need to keep on doing this process of inference during the test and this will take some time.

In order to test this notion, we designed an experiment with two groups (see Table 1). One of these groups, Group20, received an almost identical procedure to that used in the study by Rodríguez and Hall (2020; Experiment 1) in their demonstration of the standard learned predictiveness effect. In this group, the participants were given 20 seconds on each trial of the two different test stages. Participants from Group4 received an identical training to that received by Group20, except that their participants were given just 4 seconds to respond on each test trial. This limitation of time might be conflicting with the test procedure usually employed in previous studies. On these studies, participants were required to give a numerical score (from zero to 10) to each of the two possible outcomes. In the present experiment, we attempted to make simpler the test responding in order to avoid non-interesting explanations to any precluding effect obtained on Group 4 (e.g., participants in this group did not show a good performance because they had not enough time to finish their rating). Accordingly, on the test trials of the present experiment, all the participants were required just to choose what allergy, between the two possible outcomes, was going to happen.

If the learned predictiveness effect is not due to any mechanism related to changes in stimulus associability, we expected to observe the effect in Group20 but not in Group4. If, on the contrary, the learned predictiveness effect is caused by stimulus associability, we expected to observe the effect in both Group20 and Group4.

3. Method

Participants

Sixty-eight students (48 female; M_{age}= 21.12 in a range from 18 to 34) from the University of the Basque Country who, after being informed, agreed to participate in an experiment that involved cognitive tasks. All of the students had normal or corrected-to-normal vision. *The Research Ethics Committee of the University of the Basque Country* (CEISH) approved the experimental protocol.

Apparatus and stimuli

The participants were tested individually, sitting at approximately 50cm from the 17-in screen of a standard PC. The eight cues, A to D for predictors and W to Z for non-predictors, were images of fruits (apple, orange, melon, grapes, banana, strawberry, pear and cherry) and were presented on a white background.

The assignment of the images to a specific cue was randomized for each participant. The outcomes were pictures of the aversive reactions presented on the same background and accompanied by a text specifying the name of the reaction in Spanish underneath. The outcomes in the first stage (non-differential training) were ‘’headache’’ and ‘’conjunctivitis’’ and in the second one, ‘’stomach ache’’ and ‘’rash’’. The assignment of these negative reactions to Outcomes 1 or 2, and outcomes 3 or 4 was also randomized.

Procedure

Subjects were informed that they would play the role of an allergist who had to learn to predict the negative reactions that a patient, Mr. X, would suffer after having ingested certain pairs of fruits.

Differential training. This stage comprised 14 blocks of trials, with each of the eight trial types shown in Table 1 occurring once per block. Each trial began with the simultaneous presentation of the pictures of two cues (arranged horizontally at the center of the screen, 3 cm apart) and a question (arranged horizontally at the center of the screen, 3 cm underneath the pictures of the cues): Stomachache or rash? (for participants for whom stomachache was Outcome 1), or Rash or stomachache? (for participants for whom rash was Outcome 1). After 6 s, a picture illustrating the appropriate reaction, along with the name of that reaction, was presented for 3 s. In each block, the trial order was randomized apart from the restriction that

the same pair of cues could not occur on consecutive trials (i.e., it was not permitted for the same pair of fruits to be presented on the last trial of a block and the first trial of the next block). For each trial type the position (left/ right) of the fruits on the screen was counterbalanced across blocks, and the order in which these positions varied was randomized across the experiment.

Test of differential training. Following the initial stage of training, participants were told that they would be tested on what they had learned. The test consisted of eight trials, one for each type of trial of Stage 1 (see Table 1). Each trial consisted of a pair of fruits. The left/right position of each fruit and the order of presentation of each type of trial were randomized across participants. Participants were informed that Mr. X would again eat pairs of fruits and that they should rate, for each pair, what outcome, Outcome 1 or 2, would occur subsequently. Responses were made on paper sheets showing the name of the two outcomes for each trial. Participants were asked to draw a cross on the name of the outcome that they thought that was to happen. Participants from Group20 were allowed 20 s to make their response, and participants from Group4 were given just 4 s. No feedback was provided in this test.

Non-differential training. Immediately after differential training test, participants were told that they were now to deal with a different patient, Mr. Y, who suffered different reactions to the same cues. There were four blocks in this stage in each of which the four non-differential trial types shown in Table 1 appeared once per block in random order. The structure of trials in this stage of training was identical to that of trials in the previous differential training stage.

Test of non-differential training. Following the second stage of training, participants received instructions for a new test. This consisted of four trials with new combinations of cues, AC and BD being composed of cues that had been predictive in the first stage of differential training, VX and WY composed of cues that had been poor predictors in that first stage (see Table 1). As in the previous test stage, responses were made on paper sheets showing the name of the two outcomes for each trial (in this case Outcome 3 or 4). Participants were asked to draw a cross on the name of the outcome that they thought that was to happen. Again, participants from Group20 were allowed 20 s to make their response, and participants from Group4 were given just 4 s. No feedback was provided in this test.

Data treatment and analysis.

For the differential training test, percentages of right answers were calculated for each group (Group 20/Group4). Participants that over the total of eight trials of this test failed on four or more trials (i.e., marked the occurrence of the incorrect outcome) were excluded from the analysis, on the grounds that they had failed to learn the contingencies in effect during differential training. For the non-differential training test, percentages of right answers were calculated for each cue condition (Accurate vs. Innacurate) and group (Group 20 vs. Group4).

Data were analyzed by the analysis of variance (ANOVA) and a criterion of statistical significance of *p* less than .05 was adopted. Effect sizes for ANOVAs are reported as partial eta squared and those for pairwise comparisons are reported using Cohen's d. The 95% confidence intervals (CIs) around the effect sizes are also reported in parentheses following the effect size.

4. Results

Figure 1 shows the percentage of correct answers of the groups 20 and 4 across the eight trials of the test after the differential training. Both groups showed a good performance. As it can be seen, shortening the time available to respond did not preclude the manifestation of a good discrimination in Group4. A series of analysis of variance (ANOVA) performed with these data did not show any significant difference in any of the trials, *F*s < 1.

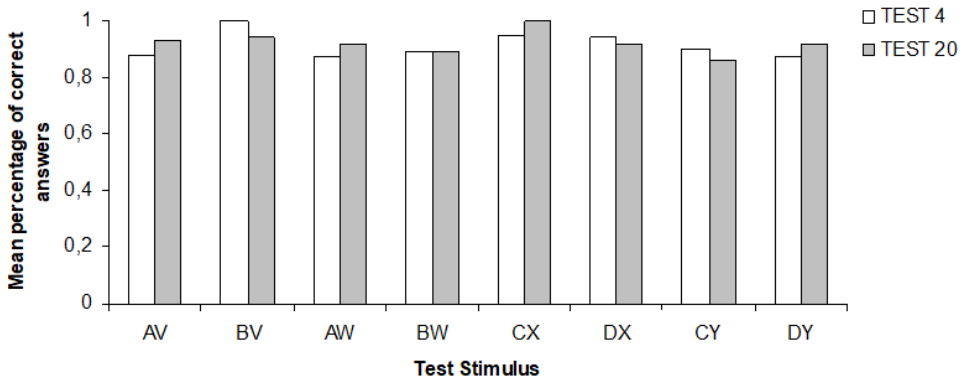

Figure 1. Percentage of correct answers given in the differential training test by Group4 (white boxes) and Group 20 (gray boxes).

Figure 2 shows the mean percentage of correct answers showed by Groups 20 and 4 to accurate and inaccurate cues, on the test stage after the non-differential training. Group 20 exhibited a pattern of results indicating a learned predictiveness effect: participants from this

groups showed better performance to the accurate than to the inaccurate cues. This pattern was not observed, however, in Group 4. Participants from this group showed a, non-specially high, similar level of performance to both accurate and inaccurate cues. The overall pattern of results indicates that a shortening of the time available to respond on test precludes the arising of a learned predictiveness effect. Statistical analysis confirmed this description of the data. A mixed 2 (Group: G20 vs. G4) x 2 (Cue: Accurate vs. Inaccurate) ANOVA revealed a non-significant effect of group, F<1, a significant main effect of cue, $F(1, 62) = 6.33$, $p = 0.014$, and a significant interaction group x cue, $F(1, 62) = 10.1$, $p = 0.0024$. Further analyses performed in order to clarify the source of this interaction revealed that there was an effect of cue in G20, $t(33) = 3.2$, $p = 0.0031$, but not in G4, $t(30) < 1$, and that there were not differences between groups either in their performance to the accurate cues, $t(31) = 1.24$, $p = 0.21$, neither in the performance to the inaccurate cues, $t(31) < 1$.

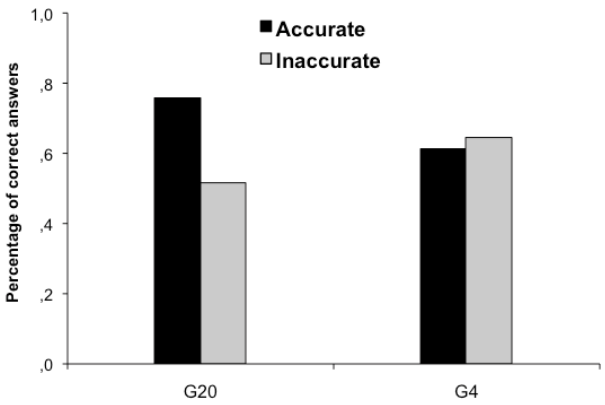

Figure 2. Percentage of correct answers given to accurate and inaccurate cues, by Group4 and Group 20, in the non-differential training test

5. Discussion

The experiment reported here was designed in order to test the hypothesis that the learned predictiveness effect is mediated by other mechanisms different to changes in associability. We found that a shortening in the time (from 20 to 4 seconds) available for test responding precluded the appearance of the effect. It could be argued that the absence of the effect in Group 4 might be reflecting that the limitation of time in this group impeded to express the respond of the participants. However, this account is not supported by the fact that such a short time was enough for the participants to show a very good performance (so good as that showed by participants in group 20). The pattern of results is therefore better explained in the terms of our hypothesis: the learned predictiveness effect seems to be due to the expression of an inference

process from all the knowledge that the participants have learned during the task. For this inference to be processed and expressed successfully, it seems necessary to dispose of a relatively long time during the test trials. If learned predictiveness were the result of the operation of bottom-up, automatic, attentional processes, a reduction in the time available for responding might have even enhance the size of the effect.

These results thus come to add to previous evidence (e.g., Hall & Rodríguez, 2010; Rodríguez & Hall, 2019) indicating that the learned predictiveness effect might not be reflecting the operation of attentional mechanisms related to optimize the limited learning resources of the organisms. In other words, although the logic of the design is open to interpretations of its results in terms of effects on associability, it seems that actually it does not capture the essence of other similar designs applied in the non-human literature (e.g., Hall & Pearce, 1979). It seems convenient that the hybrid models mentioned in the Introduction of this work, will reconsider the (perhaps scarce) importance of the learned predictiveness effect for their proposals about the mechanisms regulating the learning-rate.

6. References

Haselgrove, M., Esber, G. R., Pearce, J. M., & Jones, P. M. (2010). Two kinds of attention in Pavlovian conditioning: evidence for a hybrid model of learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *36*(4), 456.

Griffiths, O., Johnson, A. M., & Mitchell, C. J. (2011). Negative transfer in human associative learning. *Psychological science*, *22*(9), 1198-1204.

Hall, G., & Pearce, J. M. (1979). Latent inhibition of a CS during CS–US pairings. *Journal of Experimental Psychology: Animal Behavior Processes*, *5*(1), 31.

Hall, G., & Pearce, J. M. (1982). Restoring the associability of a pre-exposed CS by a surprising event. *The Quarterly Journal of Experimental Psychology Section B*, *34*(3b), 127-140.

Hall, G., & Rodriguez, G. (2010). Associative and nonassociative processes in latent inhibition: An elaboration of the Pearce-Hall model. *Latent inhibition: Cognition, neuroscience and applications to schizophrenia*, 114-136.

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*(3b), 193-243.

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.

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.

Lochmann, T., & Wills, A. J. (2003). Predictive history in an allergy prediction task. In *Proceedings of EuroCogSci* (Vol. 3, pp. 217-222).

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological review*, *82*(4), 276.

Mitchell, C. J., Griffiths, O., Seetoo, J., & Lovibond, P. F. (2012). Attentional mechanisms in learned predictiveness. *Journal of Experimental Psychology: Animal Behavior Processes*, *38*(2), 191

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.

Pearce, J. M., Kaye, H., & Hall, G. (1982). Predictive accuracy and stimulus associability: Development of a model for Pavlovian learning. *Quantitative analyses of behavior*, *3*, 241-256.

Pearce, J. M., & Mackintosh, N. J. (2010). Two theories of attention: A review and a possible integration. *Attention and associative learning: From brain to behaviour*, 11-39.

Hall, G., & Rodríguez, G. (2019). Attention to perceive, to learn and to respond. *Quarterly Journal of Experimental Psychology*, *72*(2), 335-345.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological review*, *84*(2), 127.