

SPATIAL PROBABILITY LEARNING IN THE DOLPHIN (*TURSIOPS TRUNCATUS*)

D. Simons, Psychologisches Institut der WWU Münster and M. Huigen, Delphinarium Münster.

Introduction

For decades Probability Learning (in the following abbreviated as PL) in human beings and animals has been studied. A comparison between man and animals, however, is of secondary importance. PL in animals experienced an upswing after BITTERMAN and his co-workers (1958) had begun to interpret the results of PL on a comparative phylogenetic basis. However, hopes to achieve a better understanding of the phylogenetics of learning with the help of PL were hardly fulfilled (LEHR & PAVLIK, 1970). Experimentally worked-out facts which are not easy to explain remained (SUTHERLAND & MACKINTOSH, 1971; WOODWARD & BITTERMAN, 1973). In the meantime, the enthusiasm for PL in animals has slackened. On the other hand, in spite of contrary expectations, PL in man has become very popular. There are some basic terms which should be briefly explained for the reader who is not very familiar with PL. (An excellent and extensive review of PL can be found in BITTERMAN, 1965):

1) In simple Discrimination Learning one stimulus is reinforced while the other is not. In PL both stimuli (= visual PL) or both positions (= spatial PL) are reinforced to a certain ratio (e.g. 70:30).

2) If an animal chooses exclusively the more often baited position, it is called *maximizing*. This behaviour, which guarantees maximal reward, would correspond to a reinforcement approach. According to EDWARDS (1956) the more varied the probabilities of reward of both stimuli/positions are, the more the tendency to maximize is increased (e.g. 90:10). Further, EDWARDS (1956) assumed that the amount of reinforcement could speed up or slow down the tendency to maximize.

3) If an animal distributes its preferences according to the probabilities of reinforcement, we speak of *matching*. If no systematic sequences in behaviour are evident beyond this, it is called *random matching*. The procedure of *matching* can be derived from ESTES' statistical learning theory' (1950).

A model which should explain all of the results has been presented by SIEGEL (1959). According to this model, the empirical distribution of the preferences does not correspond to the given probabilities of reward when the "marginal utility of a correct prediction" is greater than the "marginal utility of variability" (SIEGEL, 1959, p. 337). As a rule, however, this model can only give post-hoc explanations at least for experiments with animals.

4) If, at first, the more often baited position (70:30) is baited less (30:70), that is, the probabilities of reward are interchanged, we speak of *reversal*. Spatial reversal in its simplest form (two-choice, $p=0$ vs. $p=1.0$) has already been carried out with dolphins (BEACH et al., 1974). We shall return to this later.

5) The results of PL are among other things dependent on the methods of training.

Two methods are used for the study in question: *correction* (abbr.: C) and *noncorrection* (abbr.: NC). With C the animal has the possibility to go directly to the correct side after making an unsuccessful choice. With NC this possibility is not given and the animal begins to make a new, independent attempt. C, therefore, always leads to reward. In "Phyletic Differences in Learning" BITTERMAN (1965) analyses the stage research had reached at that time. Two results are of importance:

I. In spatial PL only the goldfish (if not always), earthworm and cockroach (LONGO, 1964) demonstrated *random matching*. All the other animals which were tested (rhesus monkey, rat, pigeon, turtle) maximized or demonstrated *systematic matching*.

II. In visual PL more species demonstrated *random matching*: pigeon, turtle, decorticated rat, goldfish, earthworm.

With mammals, however, true *random matching* could be proven neither in spatial PL nor in visual PL (WOODWARD & BITTERMAN, 1973).

Maximizing has been interpreted as a higher development of adaptation. However, this theory includes the weak points of all functional explanations. Contrary to this, there are complex sequences (e.g. in systematic matching) which might be closer related to the higher development of CNS. Some results, which - with the exception of the study done by WILSON and ROLLIN (1959) on rhesus monkeys - are exclusively based on experiments with rats, should illustrate how difficult it is to make a general evaluation on PL: often it could be noted that the animals did not maximize 100% (LEHR & PAVLIK, 1970 ; SUTHERLAND & MACKINTOSH, 1971). HIRAOKA (1974) found that the results were dependent on the method. In the T-maze rats maximized after 240 trials, but in the Skinner-Box maximizing was not achieved even after 920 trials. 1) Still further, great inter- and intra-individual differences made a general evaluation of the results more difficult (LEHR & PAVLIK, 1970; CALFEE, 1968; WILSON & ROLLIN, 1959). It was generally agreed that non-correction led to maximizing faster than correction (LEHR & PAVLIK, 1970; SUTHERLAND & MACKINTOSH, 1971; UHL, 1963; PARDUCCI & POLT, 1958; CALFEE, 1968; BRUNSWICK, 1939; WILSON & ROLLIN, 1959). Beyond this, UHL (1963) found that when the differences in the probabilities of reward increased, the results of C and NC were more and more similar. According to HULL & SPENCE (1938) reversal was faster in C than in NC, but the results of PARDUCCI & POLT (1958) were exactly the opposite. In reversal it also occurred that animals still chose the side which from the beginning was more often baited, that is, in spite of a reduction in reinforcement, they continued to choose the same side (CALFEE, 1968; WILSON & ROLLIN, 1959). CALFEE (1968) also made clear that the preceding 3 or 4 trials were of importance to the actual decision made by the animal. The behaviour changes back and forth from perservering to alternating. Typical of almost all of the experiments is the absence of behavioural descriptions and analyses. From most of the studies one can hardly find anything which is of importance in answering the following questions:

- Is each decision the result of the analysis of quantitative reinforcement conditions?
- Is there perhaps a difference between the given probabilities of reward and the probabilities of reward which the animal perceives?
- Is VTE (vicarious trial and error) behaviour to be observed? If so, when? How often? etc.
- When and where are decisions to be made about?
- Under which conditions sequential behaviour develops, etc.?

After analyzing the literature we cannot say which quality the decisions in PL have. We hoped to find out more about the niveau of decision-making by choosing a more complicated method. Analogous to some experiments done with children (WEIR, 1964), the animal should be able to choose between three positions and not only two, whereby the middle position should constantly remain unbaited. By that means we hoped to eliminate a more 'incidental' change.

Subject

The only subject used was a male atlantic bottle-nosed dolphin (*Tursiops truncatus*) named George. George, captured in 1969, is about 17 years old and takes part in the show. George had already been subject in the experiment done on color vision (published in this journal in 1977).

The experiments were carried out instead of the first show, which usually does not take place in winter. As a result, the time available for the experiments and the quantity of reward (normally

- 1) Note: explanation of terms on p. 9.

1 kilogramm of sardines) were limited. The number of trials pro session fluctuated between 20 and 100. George took part in the experiment, showing great and lasting "enthusiasm".

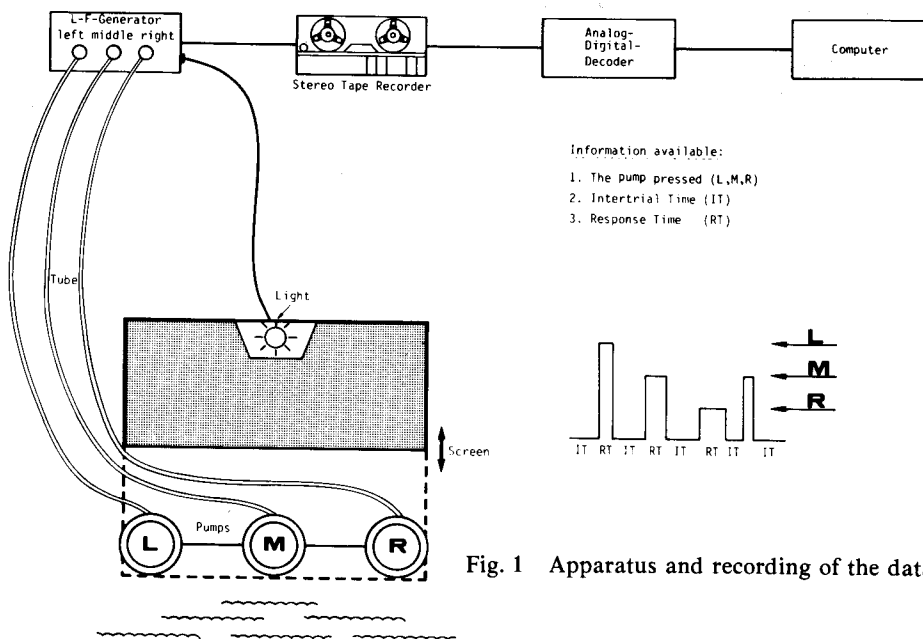


Fig. 1 Apparatus and recording of the data

Materials

Fig. 1 shows the experimental apparatus and the recording/transmission of the data. The apparatus was attached to the holding pool lengthwise (holding pool: 4.0 x 2.7 x 3.7 m). Optical and acoustic contact with other animals (dolphins, sea lions) could not be avoided. Due to the small dimensions of the pool, an "attention lever" (cf. NOORDENBOS & BOOGH, 1974) could not be used. This specific problem will be discussed at a later point. A maximum of three pumps (half spheres, made out of hard rubber, used to pump up air mattresses) could be fixed to the front cross-piece of the apparatus, directly above the water's surface. The distance between two adjacent pumps was 37.5 cm. The diameter of the pumps was 14 cm and the diameter of the outer mounting ring was 19 cm. (The illustration is true to scale.) When required (noncorrection procedure) the pumps could be manually concealed by a screen (92 x 40 cm in size). If George looked for clues from the person behind the screen, then mostly by projecting himself out of the water about 1 to 2 m in front of the apparatus or by swimming beside it. However, this type of behaviour occurred seldom on the whole. There was an automobile brake-light attached to the top side of the screen which lit up when sufficient pressure was put on the pump. An admissible trial - registrable by tape recorder - could, therefore, be shown by the lamp (feedback).

If the answer was correct, a whistle sounded (secondary reinforcement). After the whistle had sounded, the primary reinforcement (a sardine weighing about 25 grams) followed. The pressure wave caused by the pressing was transmitted through thin hoses to push-buttons (from washing-machines). A generator produced corresponding signals which were recorded by a stereo-tape recorder. With the help of an Analog-Digital-Decoder the impulses were transformed. Consequently it could be determined how long each pump had been pressed and how long the intertrial interval was.

The probabilities of reward (if not otherwise stated were 70:(00)30) were independent in non-correction and dependent in correction. A series consisted of 100 trials, presented in random order. In the series with $p = .7$ a maximum of 5 positive and 3 negative events followed one after the other. In the series with $p = .3$ it was just the opposite. The series was made up of blocks of 50 trials each.

No starting signal was given at the beginning of the experiments other than simply raising the screen. This proved to be an effective SD for the beginning of a new trial. The screen was lifted every 8 sec. In order to keep the extent of this paper as brief as possible, the respective procedure and the results will be reported jointly (arranged chronologically).

Methods and results

Pre-test:

Immediately after the initial training (approximately 25 trials), we started with the experiments (three-choice PL; right position (RP): $p = .3$; middle position (MP): $p = .0$; left position (LP): $p = .7$). We wanted to see how the behaviour would develop spontaneously. A possibility for correction was not given. Since George constantly tore the pumps out of the holder, the experiments had to be temporarily discontinued. After resuming the experiments an unintended preference for the more often baited position could be noted. During 200 trials George chose, without exception, the left position (0:0:200). In the following reversal George did not change, but continued to swim to the left position (1:5:415). During the last 100 trials he swam exclusively to the left, although the probability of reward of $p = .3$ was reduced to $p = .1$ there.

A short extinction phase was introduced (RP: $p = 1.0$; MP: 0.0 ; LP: 0.0) but it could hardly influence his left preference (2:3:46). Even the removal of the left pump had little influence on his choice (3:149). At this point we discontinued the pre-test. Contrary to our expectations - and contrary to the findings of BEACH et al. (1974) - spontaneous changes did not occur when the probability of reward was dramatically reduced. This reduction of reward had obviously more the effect of an intermittent reinforcement which, for its part, increased his resistance to extinction considerably. George's behaviour during the extinction phase (41:73:179): in the beginning George often swam back and forth between the middle and the left positions before he finally chose the right position. Later on he preferred the series left-middle-right. Often, several times after another, he swam to the left, but hardly ever to the middle. A comparison of the response times exposes the fine structure of George's behaviour: generally the whistle sounds after 1.0 to 1.4 seconds. However, the response times for the left position were significantly shorter - in the end between 0.2 and 0.6 seconds - so that during that period of time acoustic feedback was not possible at all. The response times for the middle position show a bimodal distribution. Within a series (e.g. left-middle-right) the response times were very short (0.2 to 0.6 seconds), but when isolated decisions were made in favour of the middle position, the times were between 1.0 and 1.4 seconds. The response times for the right position (the only rewarded position during the extinction phase) were, with only one exception, from 1.0 to 1.4 seconds. This one exception was his first choice for the right, where success obviously was not "expected". "Fixation" - a typical behaviour during extinction - was shown in his second attempt, made to the left: George pushed the left pump for 10.08 seconds. *Facit:* during the extinction phase George showed sequential behaviour, whereby his behaviour - expressed objectively by the response times - gives the impression that the "expectation" of reward sinks more rapidly than the frequency choosing the unrewarded position. In the following alternating training through which the animal should learn to choose both positions somewhat equally (from now on 37.5 cm apart), preceded two-choice PL. Both positions, left and right, were reinforced alternately and we set as a training goal 30 approximately equally distributed trials.

George completed 548 trials (non-correction), from which 347 were to the right and 201 to the left. Tendencies towards equal distribution could be noted again and again, but these were never kept up for a longer period of time. The longest period in which equal distribution was observed was between trials No. 270 - 319. During these 50 trials he swam to the right 26 times and to the left 24 times. Alternating from left to right (= following the rule) was seldom, five times being the limit. At the end of the training phase George clearly preferred the right. Behaviour during the training phase:

From time to time one of the two positions served as an "anchor". This "anchor" position was left only if there was no success over a long period of time (= win-stay-, lose-shift-strategy). The other position was immediately abandoned after one trial (= win-shift-strategy). Alternatively the left or the right took over the function of the "anchor". Here, too, as in the extinction phase, his behaviour is systematic.

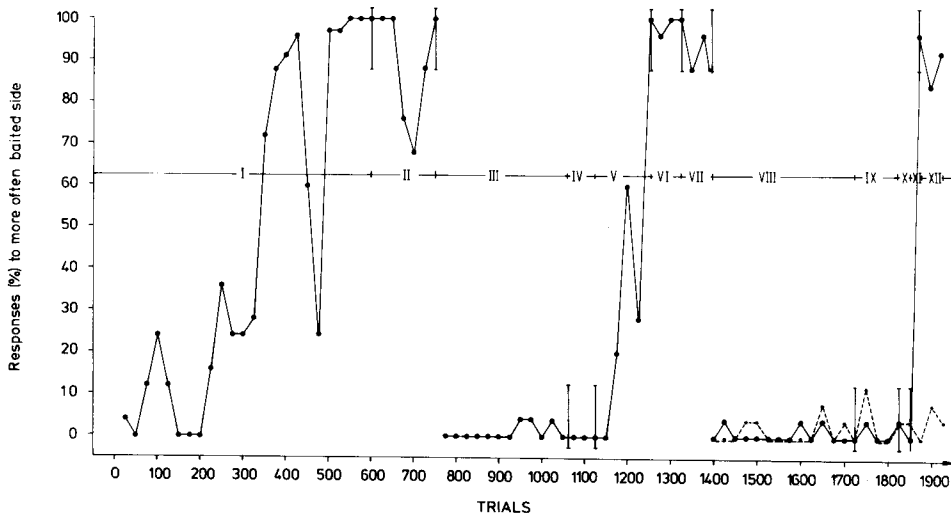


Fig. 2: Frequency of responses (%) to the more often baited side, in blocks of 25 trials. Roman numerals stand for the different methods which are described in the text. The dashed curve shows the frequency of responses (%) to the middle position.

Main tests (Fig. 2):

Two-choice probability learning

I. Method: RP: $p = .3$; LP: $p = .7$; non-correction. As George was clearly inclined to go to the right at the end of the training, the left position became the more often rewarded side.

Results: The change to the more often baited side took 500 trials. Then George clearly maximized. There was a noticeable "relapse" after 425 trials.

II. Method: RP: $p = .3$; LP: $p = .7$; correction.

Results: George's choices were temporarily distributed according to the probabilities of reward. There was no evidence of random matching whatsoever. After 125 trials George maximized again.

III. *Method*: RP: $p = .7$; LP: $p = .3$ (reversal), correction.

Results: Reversal with correction was unsuccessful in 300 trials. George might have followed a simple rule: "go to the left position, win-stay, lose-shift."

IV. *Method*: RP: $p = .7$; LP: $p = .3$ (reversal); correction; distance between the left and the right position was 75 cm.

Results: Reversal did not occur. However, George pressed the left pump several times again and again (38 times). This "increase in frequency" is characteristic of extinction situations.

V. According to EDWARDS (1956) the amount of reinforcement plays a role in decision making as well (see also SIEGEL, 1959). We wanted to test this assumption "en passant".

Method: RP: $p = .7$; LP: $p = .3$ (reversal); correction; distance between left and right 75 cm. From trials No. 1140 - 1200 rewards were given in the ratio of 2:1, i.e. a whole sardine for the right position and a half for the left. In the trials No. 1140 - 1143 a mackerel was offered as incentive. The mackerel was placed in the middle above the apparatus for George to see.

Results: After three further attempts, George changed to the right and received the mackerel. Afterwards George swam to the right three times without being successful. He looked very "astonished" at that moment. In his second positive attempt to the right he received another mackerel, which had not been shown to him before. Hereafter George swam more often to the right. His answer to the re-establishing of the former reward ratio (1:1) was a relapse to his original leftside preference. However, he then changed to the right and soon maximized.

Three-choice probability learning

VI. *Method*: RP: $p = .7$; MP: $p = .0$; LP: $p = .3$; correction; the distance between each of the positions was 37.5 cm.

Results: George continued to maximize. If the left position was correct, then he pressed, without exception, the middle pump too (24 times), that is, he swam from right by way of middle to left. His response times to the middle pump were very short.

VII. *Method*: RP: $p = .7$; MP: $p = .0$; LP: $p = .3$; non-correction.

Results: George continued to maximize.

VIII. *Method*: RP: $p = .3$; MP: $p = .0$; LP: $p = .7$ (reversal); non-correction.

Results: George did not change his preference for the right position (360 trials).

IX. *Method*: RP: $p = .1$; MP: $p = .0$; LP: $p = .7$ (reversal); non-correction.

Results: Although the probability of reward of $p = .3$ had been reduced to $p = .1$, George went on choosing the right position (100 trials).

X. *Method*: RP: $p = .1$; MP: $p = .0$; LP: $p = .9$ (reversal); correction.

Results: There was no change in George's positional habits. When forced to correct (22 times out of 25) he chose the left side only after having chosen the middle position before.

XI. *Method*: RP: $p = .3$; MP: $p = .0$; LP: $p = .7$ (reversal); non-correction; the whole apparatus was moved further to the right (90 cm).

Results: George swam to the left position almost automatically (25 trials).

XII. *Method*: RP: $p = .3$; MP: $p = .0$; LP: $p = .7$; non-correction; the apparatus was removed onto its former place.

Results: In the beginning George showed VTE-behaviour. Then he preferred more and more the left position. During the last 20 trials he chose the left side exclusively.

The qualitative analysis of George's *behaviour* is, inspite of its limited objectivity, necessary in order to be able to evaluate the data adequately. Therefore, we would like to point at two important aspects:

1. Vicarious trial and error learning (VTE-learning). VTE-learning with its intentional movements is often the only indication for the observer to know whether the choice is preceded by something like a "decision making process". Here we are concerned with a phase of attentiveness, the searching for information (KLIX, 1971), which is often preceded by a change in behaviour, an increase in learning.

During the pre-test (three-choice PL) no such phases could be observed. George's behaviour appeared to be quite stereotype. On the other hand in the alternating learning at least 20 VTE phases were observed. One-sidedness did not bring success and therefore a change in behaviour was necessary. In the following (two-choice/three-choice PL), each time the method was changed, VTE phases appeared. Facit: The search for cues to the correct solution took place in certain phases. The "decisions" made in the beginning of each new method determine the choice in further trials. Rather variable behaviour in the beginning (learning) was followed by rather stereotype behaviour (performance) in later trials. From this follows that probability learning is probably not a continuous decision making process.

2. The temporal relationship between the response itself and preceding decision making process.

In order to avoid the starting-place determining the decision, methods have been developed to guarantee a straightline approach to the middle of the apparatus.

NOORDENBOS & BOOGH (1974) used an "attention lever", BEACH et al. (1974) used a tire which was attached beneath the surface of the water. In this study none of the above-mentioned methods were used because the holding pool was too narrow (2.7 m).

Independent from the method used, the question still remains as to how the temporal relationship between the response itself and preceding decision making processes can be understood. In accordance with the occurrence of VTE phases it appeared that during the learning phases the observation how George approached the apparatus did not make possible to predict the choice definitely. On the contrary during the apparently stereotype phases of performance the choice could be predicted by observing how George approached the apparatus. During such a stereotype phase we could throw a reward in front of the other pump, which was not preferred, but still George swam to "his" position in the next trial. The last two experiments (No. 11 + 12) seem to support these arguments. George swam almost automatically to the right side. Then the whole apparatus was moved further to the right (90 cm). Now George approached the apparatus exactly the same way as before and chose consequently the left position. After the apparatus had been removed onto its former place George showed VTE behaviour. It seemed to be the result of a conflict between the changing reinforcement ratio and his habits approaching the apparatus. During this VTE phase George's choice could not be predicted at all. Facit: Sometimes the decisions appeared to be made shortly before pressing a pump, sometimes much earlier. George's behaviour gave the impression that, analogous to the so-called 'Wegelernen' (LORENZ), there were a limited number of ways of approaching the apparatus, which were linked to certain decisions. So the decisions seemed to be primary. The relationship between positional preferences, habits concerning motor behaviour, and decision making processes make it difficult to understand *spatial* probability learning.

Conclusions

The classification of the results into general theories is comparatively easy: There is every reason to believe that George maximizes and in the long run even under all conditions (non-correction, correction, reversal). Signs of random matching could not be found in any of the data. George behaved almost exactly like type "R" (BITTERMAN, 1965), i.e. he behaved in an adequate manner compared with other mammals. George showed marked positional preferences which hardly could be changed.

In three-choice PL this result may be explained - at least partially - by the introduction of the middle position. George never abandoned choosing the middle position although he did not get any reward there. As a matter of fact the middle position seemed to make positional changes more difficult. A comparative experiment with children may throw light upon that question.

Similar findings - failure of reversal - have been reported by CALFEE (1968) and WILSON & ROLLIN (1959). It could be simply as well as impressively (at least for the authors) demonstrated that the amount of reinforcement may influence the decision making process.

In order to understand the decision making process, the reward "expectation" seems to be important. In spite of a reduction of the probability of reward of $p = .3$ to $p = .1$ (main test No. 9) George saw no reason to change his behaviour. The "costs" were obviously too low to make him change his positional habits. George's reactions to the reduction of the probability of reward are equivalent to the behaviour found during an extinction procedure:

1. Increase of response frequency: George went to the same position several times in a row.
2. Change in topography: George pressed the pump in a different manner and more intensive than otherwise.

In one case, he is known to have pressed the pump for 10.08 seconds!

3. Change in positional habits: George swam to another position which had been avoided before.

Supplementary experiment with children

A supplementary experiment with children should lead to a better understanding of the spatial probability learning task itself. Furthermore, we were interested in behavioural similarities between George and the human subjects. By similarity we mean 'analogy', not 'identity'.

Our subjects were two boys: Nick (8;2 years old) and Lars (4;10 years old). The same apparatus was used, outside the water of course. Beside the instruction how to use the apparatus the children were given the following task:

1. "You will get one pfennig for every correct response. After each response I will say "correct" or "incorrect" (non-correction procedure).
2. "You are allowed to press one pump after another until I say "correct" (correction procedure).

The results are shown in Table I:

Table I.

Response frequencies in three-choice PL with children. N = Nick (8;2), L = Lars (4;10).

		Response frequencies (%)								
probability:		.7 : .0 : .3			.3 : .0 : .7			.7 : .0 : .3		
method:		non-correct.			non-correct.			correction		
					reversal			reversal		
trials										
1 - 25	N.	40	36	24	68	16	16	64	12	24
	L.	60	24	16	36	36	28	68	16	16
26 - 50	N.	84	12	4	40	16	44	64	0	36
	L.	60	20	20	28	40	32	80	0	20
51 - 75	N.	64	12	24	28	12	60	48	0	52
	L.	64	20	16	12	56	32	92	4	4
76 - 100	N.	72	16	12	20	32	48	56	0	44
	L.	68	16	16	20	32	48	92	0	8
101 - 125	N.	84	12	4	0	4	96			
	L.				20	44	36			
126 - 150	N.	96	4	0	0	8	92			
	L.				20	32	48			

Contrary to George both children showed a greater sensibility to reductions in the probability of reward. From the beginning (non-correction) Lars displayed sequential choice patterns (e.g. left-middle-right) which were followed by win-stay and lose-shift strategies (correction). Nick started with win-stay and lose-shift strategies and after about 100 trials he began to maximize (non-correction). In the following correction situation he developed a simple alternating strategy (R-L-R-L-...). This strategy prevented him from choosing the middle position. Not so Lars: like George he used to go to the left side. When the left position was wrong, he went to the middle position and only then to the right side. And he pressed the middle pump more and more faster than the other pumps. This behaviour looked very similar to that of George during the extinction trials.

It remains the fact that in non-correction situations there was no quantitative difference in response frequency between the middle position and the less often baited side. In accordance with George's behaviour Lars and Nick displayed both VTE learning and rather stereotype choice patterns (e.g. sequential behaviour). Like George both children showed all kinds of behaviour which are typical for extinction procedures: change in response frequency, change in topography, change in positional preferences. During the last correction trials Nick said that we had obviously stopped to give money for pressing the middle pump! At the end of the whole experiment he was then asked how often he had got money for pressing the middle pump. Nick was really convinced that the middle position had been rewarded too. It follows that the empirical frequency of reward pro position must not be identical with the individual's assumption about the reward frequencies.

Summary

An atlantic bottle-nosed dolphin (named George, 17 years old, captured in 1969) was trained in a two-choice (70:30) and three-choice (70:00:30) spatial probability learning task for more than 3.500 trials. There was a tendency to maximize under all conditions (correction, non-correction, reversal). Signs of random matching could not be found in any of the data. George developed distinct positional preferences which made reversals very difficult. The behaviour George displayed after reductions of the probability of reward (e.g. during reversal tasks) was equivalent to the behaviour found during extinction procedures: 1. increase of response frequency, 2. change in topography and 3. change in positional preferences. In the beginning of every new experimental procedure George showed VTE learning which was generally followed by rather stereotype behaviour. Comparative experiments with children showed 'similar' results and supported the presumption that at least the results in three-choice probability learning could hardly be explained by simple reinforcement models.

Acknowledgements

We owe the construction of the apparatus and the recording system to the proficiency of Mr. J. Driese, Mr. W. Hinsenhofen, and Mr. B. Wuth. We thank Mr. D. Langheinrich for his ingenious computations. We are grateful to Mr. and Mrs. Grünefeld who translated most parts of the paper into english. All grammatical mistakes, however, are at the first author's responsibility. We are also indebted to Dr. W.H. Dudok van Heel who supported this research. Last but not least we thank George, Nick and Lars for their friendly co-operation.

Notes

ad *T-maze*: "The maze can be defined as a problem that requires for its solution the selection of the shortest route to a goal." (KRECH et al., 1969, p. 306). In a T-maze the animal has to choose between either a right turn or a left turn.

ad *Skinner-box*: "An apparatus constructed by SKINNER, called a Skinner-box, has frequently been used in experiments to illustrate operant conditioning. The box is equipped with a tray and a spigot (for delivery of food and water), a lever and a light protected by a screen." (HABER & FRIEND, 1975, p. 152).

References

- BEACH, F.A. III, R.L. PEPPER, J.V. SIMMONS Jr., P.E. NACHTIGALL and P.A. SIRI, (1974). Spatial habit reversal in two species of marine mammals. *Psychol. Rec.* 24, 385 - 391.
- BITTERMAN, M.E., (1965). Phyletic differences in learning. *Amer. Psychol.* 20, 396 - 410.
- BITTERMAN, M.E., J. WODINSKY and D.K. CANDLAND, (1958). Some comparative psychology. *Amer. J. Psychol.* 71, 94 - 110.
- BRUNSWIK, E., (1939). Probability as a determiner of rat behaviour. *J. Exp. Psychol.* 25, 175 - 197.
- CALFEE, R.C., (1968). Choice behaviour during long-term probabilistic reinforcement schedules. *J. Comp. Physiol. Psychol.*, 65, 232 - 237.
- EDWARDS, W., (1956). Reward probability, amount, and information as determiners of sequential two-alternative decisions. *J. Exp. Psychol.* 52, 177 - 188.
- ESTES, W.K., (1950). Toward a statistical theory of learning. *Psychol. Rev.* 57, 94 - 107.
- HABER, R.N. and A.H. FRIED, (1975). *An introduction to Psychology*. New York: Holt, Rinehart & Winston.
- HIRAOKA, K., (1974). Probability learning by rats in T-maze and Skinner box. *Ann. Animal Psychol.* 23, 69 - 75.
- HULL, C.L. and K.W. SPENCE, (1938). "Correction" vs. "Non-Correction" method of trial-and-error learning in rats. *J. Comp. Psychol.* 25, 127 - 145.
- KLIX, F., (1971). *Information und Verhalten*. Bern: Huber
- KRECH, D., R.S. CRUTCHFIELD and N. LIVSON, (1969). *Elements of Psychology*. New York: A.A. Knopf.
- LEHR, R. and W.B. PAVLIK, (1970). Within-subjects procedural variations in two-choice probability learning. *Psychol. Rep.* 26, 651 - 657.
- LONGO, N., (1964). Probability-learning and habit-reversal in the cockroach. *Amer. J. Psychol.* 77, 29 - 41.
- NOORDENBOS, J.W. and C.J. BOOGH, (1974). Underwater visual acuity in the bottlenosed dolphin. *Aquatic Mammals* 2, 15 - 24.
- PARDUCCI, A. and J. POLT, (1958). Correction vs. Non-correction with changing reinforcement schedules. *J. Comp. Physiol. Psychol.* 51, 492 - 495.
- SIEGEL, S., (1959). Theoretical models of choice and strategy behaviour: stable state behaviour in the two-choice uncertain outcome situation. *Psychometr.* 24, 303 - 316.
- SUTHERLAND, N.S. and N.J. MACKINTOSH, (1971). *Mechanisms of animal discrimination learning*. New York: Academic Press.
- UHL, Ch., (1963). Two-choice probability learning in the rat as a function of incentive, probability of reinforcement, and training procedure. *J. Exp. Psychol.* 66, 443 - 449.
- WEIR, M.W., (1964). Developmental changes in problem-solving strategies. *Psychol. Rev.* 71, 473 - 490.
- WILSON, W.A. Jr. and A.R. ROLLIN, (1959). Two-choice behaviour of rhesus monkeys in a noncontingent situation. *J. Exp. Psychol.* 58, 174 - 179.
- WOODWARD, W.T. and M.E. BITTERMAN, (1973). Further experiments on probability learning in goldfish. *Animal Learning & Behaviour* 1, 25 - 28.