Learning with Delayed Rewards in Octopus

M. J. WELLS and J. Z. YOUNG
Department of Zoology, Cambridge, and
Department of Anatomy, University College, London

Received July 26, 1968

Summary. 1. Octopuses can be trained to discriminate between white and black discs given at successive trials when the food rewards and shocks are delayed for as much as 30". There are slight signs of learning even with delays of 60".

2. The amount of information acquired from each attack (at white or black) decreased with delay in reward.

3. The tendency to attack the positive figure declined with delay in reward, especially beyond 30".

4. If rewards were then given without delay attacks increased, showing that there is a positive process of learning to attack.

5. Octopuses also show some memory storage allowing distinct reactions by touch to rough and smooth spheres with delay of up to 30" in reward.

6. The amount of information gained from each occasion of taking a sphere decreased progressively with delay.

7. There is evidence from both visual and tactile fields that discrimination learning under these conditions of successive presentation involves the formation of distinct representations ensuring take (attack) with one object and rejection (retreat) with another. The two may show different rates of information storage, which do not alter in the same way when changes are made in the delay before reward or punishment.

8. With longer delays (60" and 120") some individuals made "discriminant" scores, but there are reasons for believing that this can be attributed to extraneous circumstances, including alternate presentation, and a fall in the probability of taking within sessions.

9. After removal of the vertical lobe octopuses were unable with rewards delayed 15" to learn a black-white discrimination against the preference.

10. Animals without vertical lobes trained with tactile discriminations showed less than normal capacity to learn with 10" delay. With longer delays the individuals were characteristically variable. Some quickly came to take the positive sphere on nearly all occasions, others performed randomly, taking both spheres very often.

11. Two characteristics of animals without vertical lobes are thus to swing to extreme preferences and to be unable to learn not to take objects that yield shocks.

12. Animals with the supraoesophageal lobes bisected learn rather less well than normals when rewards are delayed, but do not show the aberrations characteristic of those without vertical lobes.

I. Introduction

In any animal that uses distance receptors there is liable to be a delay between the receipt of signals indicating possible food or enemies and the actual arrival of the result of any action that the animal may
take. It was long ago indicated by Sherrington (1906) that the bridging of this gap or delay may well be one of the major functions of higher parts of the nervous system and in particular of the cerebral cortex. Nevertheless it is often supposed that learning is difficult unless the reward follows very quickly after a stimulus (if conditions are such as to prevent secondary reinforcement) (Grice, 1948). This is curious since there is much evidence for animals of various species that if two figures are shown, one with a food object, choice will be correct even minutes or hours after removal of the food (Hunter, 1912; Konorski, 1959). It has been shown that octopuses are capable of making such delayed responses for up to 30" (Dilly, 1963). Moreover they can make successful detours in mazes when the goal is obscured for as long as one or two minutes (Wells, 1964, 1967). We were therefore interested to determine whether they could learn in successive discrimination situations if the rewards were delayed. It proved that they can do so both for visual and tactile discriminations with delays up to 30" and possibly with even longer delays.

II. Visual Learning with Delayed Rewards

1. Methods

O. vulgaris were trained in white plastic tanks 27 x 100 x 50 cm with circulating water. A home was provided at one end of the tank in the form of a small box 15 x 10 x 7 cm let into the back of the tank. The animals mostly remained within the home when not being tested. If there was a tendency to wander towards the front end of the tank it was checked by inserting a sheet of transparent plastic covering the full width of the tank at some 15 cm in front of the home. This was raised a few seconds before each trial. With most animals the partition could be dispensed with after it had been used for a while. No complete solution has yet been found to the problem of ensuring that an octopus is always in a given position at the beginning of a test.

The animals used were of 250—500 g, of either sex. They were kept for some days in the laboratory before experimentation and tested repeatedly by presentation of crabs. The experiments were begun only when they had made several attacks at crabs.

The octopuses were trained to discriminate between white and black plastic discs 4.6 cm diameter presented successively on the end of transparent plastic rods at the end of the tank remote from the home. Attacks at the "positive" figure were rewarded with a small piece of sardine presented on the end of a wire. After attacks at the negative figure the octopus was given a 10 v. A.C. shock by touching it with a pair of electrodes plunged into the tank.