

NATURE DE L'OPERANT ET RENFORCEMENT DE DÉBITS DE RÉPONSES LENTS CHEZ LE PIGEON

[NATURE OF THE OPERANT AND DIFFERENTIAL REINFORCEMENT OF LOW RATES IN THE
PIGEON]

HUGUETTE MANTANUS*, CHRISTINE FAYT & BERTHE FERETTE

* *Aspirant F.N.R.S.*

Pigeons were conditioned in a multiple, two-operants DRL schedule of reinforcement. The response delay required was 10 seconds. The operants used were key-pecking and treadle-pressing, each being associated with a coloured stimulus, and reinforced by grain presentation. After a 12-session baseline the subjects were shifted to a chain schedule in which the key-pecks were reinforced by appearance of a conditioned stimulus, the other response remaining reinforced by grain. After a return to baseline conditions the subjects were submitted to the reverse chain schedule, i.e. with the treadle press being reinforced by a conditioned stimulus. Then the subjects were finally returned to baseline conditions. Important lengthening of interresponse times, together with increase of the percent reinforced responses were observed for each operant when conditioned stimulus instead of grain was the reinforcement. These effects were less marked and there was more day-to-day variability in the case of the treadle-press response. The results are discussed with regards to the problem of the nature of operants, that is, the more or less arbitrary relationship between the behavior unit and its reinforcing consequences.

Le coup de bec sur un disque de plexiglas, réponse opérante du pigeon, a pendant longtemps reçu le même statut que la pression de la patte sur un levier exécutée par le rat, le chat ou le singe. Des études critiques ont récemment mis en doute cette équivalence, en attirant l'attention sur certaines caractéristiques du coup de bec.

Deux chercheurs américains, Brown et Jenkins (1968) ont constaté que le becquetage du disque ou «clé-réponse» apparaît de manière spontanée en situation expérimentale. Il suffit de présenter des graines à intervalles réguliers en illuminant la clé-réponse quelques secondes avant l'apparition de la nourriture. Il semble alors que les sujets transfèrent vers la clé les coups de bec qu'ils donnent aux graines. Avec cette procédure, qui n'est pas sans évoquer les associations de type pavlovien, la réponse apparaît d'emblée dans sa forme définitive. L'expérimentateur n'intervient pas pour renforcer des approximations successives, d'où le nom d'«autoshaping» utilisé pour désigner cette méthode.

Une autre particularité de la réponse «coup de bec» est son adaptation médiocre au programme de renforcement de débits lents (DRL). L'impossibilité des pigeons à produire des délais supérieurs à 22 secondes (Staddon, 1965) a été attribuée au débit naturellement élevé du coup de bec. Cette fréquence naturelle entrerait en compétition avec l'exigence d'espacement des réponses du programme DRL (Staddon, 1972).

Ces observations ont été à l'origine d'une réflexion sur le problème de la nature de la réponse, que l'on peut résumer ainsi : des contraintes biologiques privilégient les relations entre certains types de comportements et certains renforcements. Ces contraintes modifieraient la probabilité d'apparition des réponses opérantes en fonction de la situation renforçante. Ainsi, les réponses apparentées à la prise de nourriture seraient plus facilement induites par des conditions de renforcement alimentaire.

Cette hypothèse a fait l'objet de quelques vérifications expérimentales ; elles consistent à comparer, chez les mêmes sujets, des réponses plus ou moins arbitraires par rapport au renforcement. Kramer & Rodriguez (1971) ont montré que des rats renforcés par de l'eau ont une moins bonne performance en DRL 10 sec. les jours où la réponse opérante consiste à lécher un tube, par rapport aux séances où ils doivent appuyer sur un levier. De même des pigeons, renforcés par des grains et utilisant en alternance une réponse de coup de bec et un appui de la patte sur un levier, montrent une adaptation plus rapide et atteignent des délais plus élevés en DRL avec la réponse d'appui (Mantanus et al., 1977).

A partir de cette dernière expérience, nous avons voulu pousser plus loin l'analyse des effets d'une liaison naturelle entre réponse et renforcement chez le pigeon. Notre hypothèse est la suivante : si la performance médiocre, c'est-à-dire le débit élevé de la réponse « coup de bec », est due à la présence d'un renforcement alimentaire, on devrait pouvoir améliorer les résultats en renforçant cette réponse par un événement à caractère positif, mais non alimentaire. La réponse d'appui, plus arbitraire, ne serait que peu affectée par une telle modification.

MÉTHODE

SUJETS

Deux pigeons voyageurs mâles, P1 et P2, âgés de 6 et 7 ans au début de l'expérience. Ces pigeons ont reçu un entraînement préalable au programme DRL à deux réponses. Pendant toute la durée de l'expérience, ils sont logés individuellement et maintenus à 90 % de leur poids.

APPAREILS

Une cage de Skinner pour pigeons, 40 × 40 × 40 cm, pourvue d'une clé-réponse et d'un levier, situés respectivement à 20 cm et 3 cm du plancher et distants de 20 cm. Un magasin à graines Gerbrands, rempli d'un mélange standard, une lampe d'ambiance de 25w et deux lampes 10w rouge et jaune complètent l'équipement. La programmation automatique est assurée par un ensemble de relais et de minuteries. Les données sont recueillies au moyen d'un enregistreur Gerbrands à 6 plumes et de 4 compteurs d'impulsions.

Phase I : Les sujets, déjà conditionnés au début de l'expérience, sont placés d'emblée en ligne de base. Pendant 12 séances de 60 minutes, ils sont soumis à un programme multiple DRL 10'' - DRL 10'' à deux réponses. Les composants ont un cycle de 30 minutes et sont signalés par des lampes de couleur : le stimulus rouge signifie que le coup de bec est la réponse renforcée, le jaune est associé à l'appui de la patte sur le levier. L'ordre de présentation des composants alterne chaque jour. Le renforcement consiste en un accès de 3 secondes au magasin de graines.

Phase II : Le programme est modifié de façon à ce que la réponse «bec» ne soit plus suivie d'un renforcement alimentaire. Au départ, le stimulus rouge est présent. Toute réponse survenant 10 secondes après la précédente est désormais renforcée par l'apparition du stimulus signalant la composante «patte». Au cours de celle-ci, l'appui sur le levier est renforcé par des graines comme dans la phase précédente et selon le même DRL 10 secondes. Après le renforcement d'un appui, le stimulus jaune est remplacé par le stimulus rouge et le cycle recommence. Ces conditions correspondent à un programme chaîne DRL-DRL (Ferster & Skinner, 1957).

Il faut remarquer que le changement de stimulus peut aussi marquer le début d'un intervalle : il suffit à l'animal de donner sa réponse 10 secondes après l'illumination d'une lampe pour être renforcé. L'alternance des composants n'est donc plus soumise à un cycle fixe : toute réponse correcte provoque un changement de stimulus. Le programme chaîne est en vigueur pendant 12 séances d'une heure.

Phase III : Cette phase constitue un contrôle : les sujets sont replacés en condition de ligne de base, c'est-à-dire 12 séances de programme multiple DRL 10'' - DRL 10'' avec renforcement alimentaire pour les deux réponses.

Phase IV : Le programme chaîne est de nouveau en vigueur, mais avec inversion des réponses. L'appui sur le levier commence la séquence. Il est renforcé par l'apparition du stimulus contrôlant la réponse de coup de bec. Cette dernière est renforcée par des graines, après quoi le stimulus jaune réapparaît. Le même délai de 10'' est exigé. Cette phase comporte également 12 séances.

Phase V : Cette phase est identique à la première et la troisième et vérifie le retour à la ligne de base.

RÉSULTATS

Le pourcentage de réponses renforcées, encore appelé «efficacité de comportement», a été calculé pour chaque séance au cours des différentes phases de l'expérience. L'effet de l'introduction du programme

chaîne est identique chez les deux sujets. La réponse qui constitue le premier terme de la chaîne voit son efficacité augmenter considérablement à partir de la deuxième ou troisième séance de mise en vigueur du programme. Le retour à la ligne de base est encore plus rapide, puisque le niveau antérieur réapparaît dès le changement de contingences.

Il faut également noter que l'efficacité de la réponse «patte» au cours de la phase IV n'est pas toujours aussi élevée que celle atteinte par la réponse «bec» au cours de la phase II. Elle peut même présenter

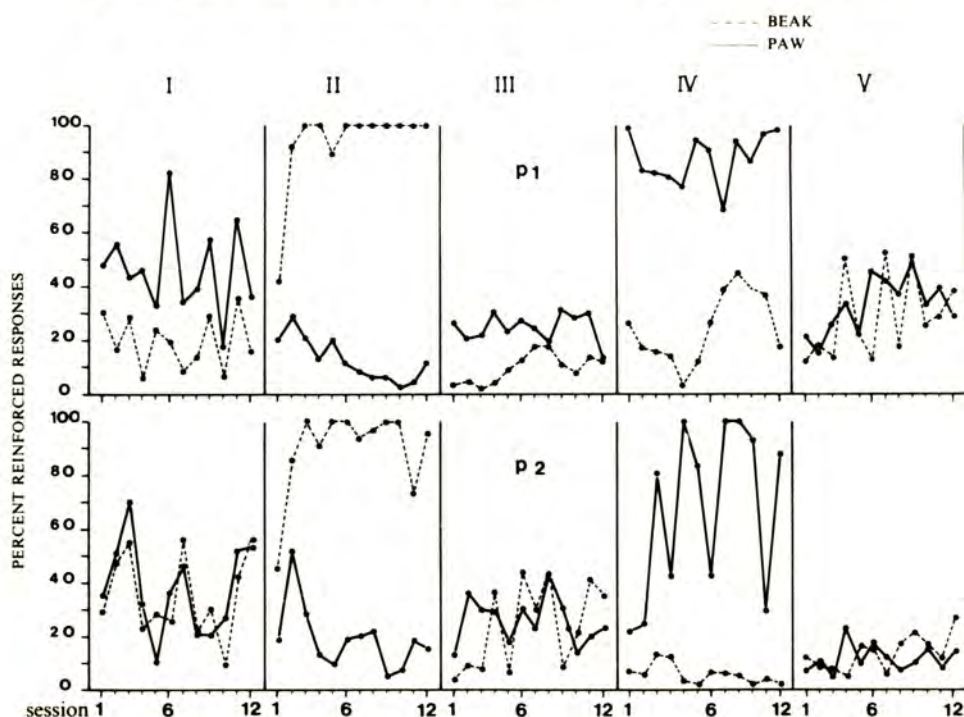


FIG. 1. PERCENT REINFORCED RESPONSES FOR EACH OPERANT, KEY-PECK («beak») AND TREADLE-PRESS («paw»), DURING THE DIFFERENT PHASES OF THE EXPERIMENT. Phases II and IV refer to chained schedule contingency. The other phases are baseline, multiple-schedule conditions.

une assez forte variabilité: celle-ci apparaît nettement sur les tracés de la Figure 1.

D'autre part, le programme chaîne semble aussi affecter légèrement l'efficacité de la réponse qui constitue le deuxième terme de la séquence, et qui est renforcée par des graines. On observe, surtout pour le sujet p2, une diminution d'efficacité par rapport à la ligne de base en programme multiple. Ceci correspond à une élévation du débit de la

deuxième réponse, parallèlement à la diminution du débit de la première. Ce phénomène a été appelé *contraste comportemental* (Reynolds, 1961).

Les Figures II et III présentent les résultats de l'analyse des intervalles inter-réponses (*interresponse times*, en abrégé *IRTS*). Il s'agit d'histogrammes moyens pour les 10 dernières séances de chaque phase. Les intervalles sont répartis en classes de deux secondes et convertis en pourcentages. Sont également considérés comme *IRTS* les intervalles séparant l'apparition d'un stimulus et l'émission de la première réponse associée à ce stimulus.

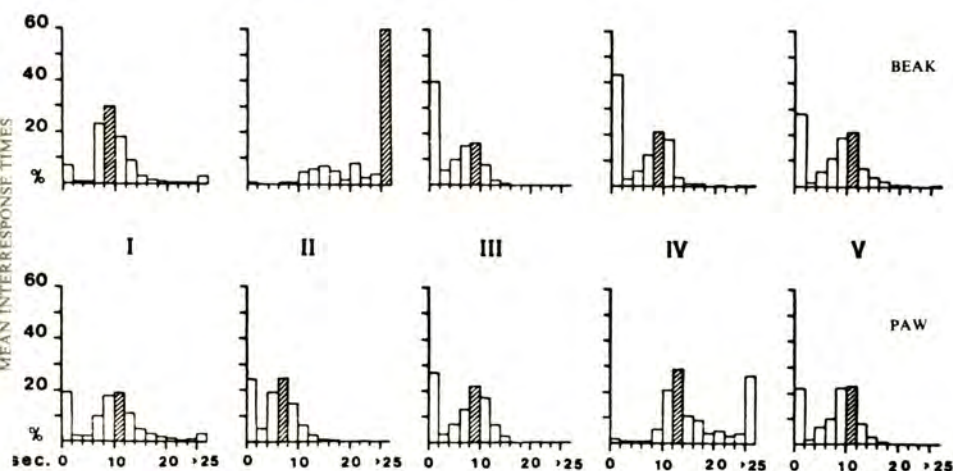


FIG. 2. PERCENT INTERRESPONSE TIMES IN 2-SECONDS TIME BINS, FOR BOTH OPERANTS OF SUBJECT P1. Each histogram is the mean of the last 10 sessions of the phase. Striped area corresponds to the mode of the distribution without regards to the importance of the first class (bursts).

Au changement d'efficacité constaté lors du passage au programme chaîne correspond une profonde modification de l'allure des histogrammes. Les réponses qui constituent le premier membre de la chaîne sont nettement plus espacées qu'en programme multiple. Dans la plupart des séquences, une seule réponse est émise après l'apparition du premier stimulus, avec une latence si longue (25 secondes et plus) qu'elle est aussitôt renforcée par l'apparition du deuxième stimulus.

Si l'on considère les *IRTS* «bec» de la phase II et les *IRTS* «patte» de la phase IV, on remarque certaines différences. Dans le cas de la réponse «patte», il y a toujours une plus forte proportion d'*IRTS* inférieurs à 10 secondes. De même la tendance centrale, représentée par le mode, est plus accusée. On observe même une distribution bimodale chez le sujet p1.

Les phénomènes de contraste comportemental mentionnés plus haut sont également traduits au niveau des histogrammes. En considérant les IRTs «patte» de la phase II et les IRTs «bec» de la phase IV, on

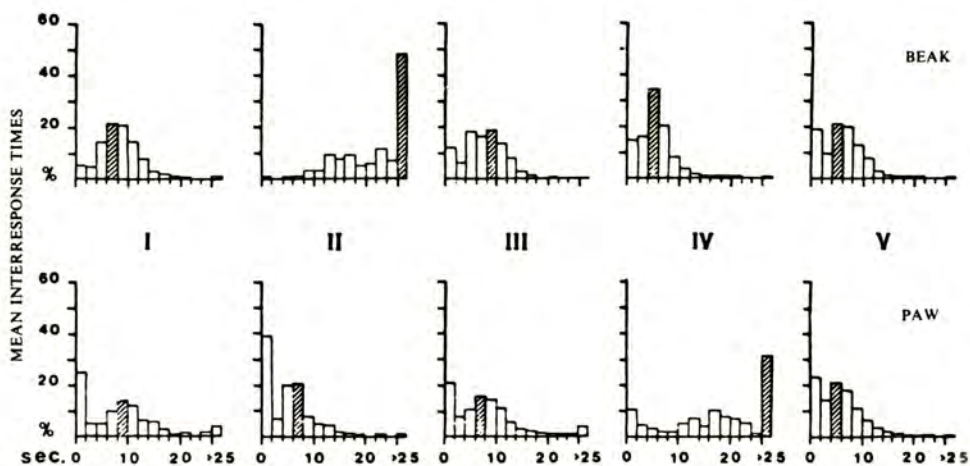


FIG. 3. EQUIVALENT OF FIGURE 2 FOR BIRD P 2.

voit que les réponses qui terminent la chaîne présentent soit une plus grande proportion d'IRTs inférieurs à 2 secondes, soit un recul du mode par rapport aux phases «multiple» précédentes.

DISCUSSION

Au vu des pourcentages de réponses renforcées et des distributions d'intervalles inter-réponses, on peut déclarer que la performance des pigeons est nettement affectée par le passage de programme multiple à programme chaîne. Qu'il s'agisse de l'appui sur le levier ou du coup de bec sur la clé, la réponse constituant le premier terme de la chaîne réagit très vite au changement de contingences. Les intervalles s'allongent, entraînant une augmentation de l'efficacité.

On peut toutefois se demander si les modifications résultent bien du fait que les réponses sont contrôlées par un autre type de renforcement, ou si d'autres facteurs peuvent expliquer ces résultats. En premier lieu, la succession des événements au cours des phases II et IV fait que l'IRT du premier terme de la chaîne vient toujours à la suite d'un renforcement alimentaire. Or, il est connu que la probabilité d'un IRT long est plus élevée après un IRT renforcé (Weiss et al., 1966): on pourrait donc attribuer l'allongement des IRTs du premier terme de la chaîne au renforcement qui les précède. Pour vérifier ceci, nous avons comparé les IRTs «bec» et «patte» des phases II et IV,

respectivement, aux IRTS post-renforcement de ces mêmes réponses au cours des phases de programme multiple. Il est apparu que les distributions d'IRTS post-renforcement possèdent les mêmes modes que les histogrammes généraux présentés dans les figures II et III. La seule différence concerne la première classe, qui ne comporte qu'une faible proportion (moins de 10%) d'intervalles. Il semble donc qu'on ne puisse rendre compte de la distribution des IRTS en programme chaîne à partir des IRTS post-renforcement.

Un deuxième facteur, l'influence des stimuli discriminatifs, a pu être écarté également. En effet, si les latences très longues observées en programme chaîne étaient dues à une réaction émotionnelle provoquée par le changement de stimulus, on retrouverait la même réaction dans n'importe quel programme, chaque fois que le changement a lieu. Or, les latences observées en programme multiple, lorsque l'on passe d'une composante à l'autre, sont de l'ordre de 2 à 5 secondes dans la majorité des cas.

Par ailleurs, les travaux d'autres auteurs semblent apporter une confirmation à nos observations : Thomas & Stubbs (1967), utilisant un programme chaîne DRL à 3 composants, remarquent un allongement des IRTS au niveau du premier membre de la chaîne. Un effet de contraste comportemental comparable à celui que nous observons a été mis en évidence par Wilton & Gay (1969) : le débit de réponses engendré par un programme de renforcement à intervalle variable s'élève lorsque ce programme devient le dernier terme d'un programme chaîne.

A propos de cet effet de contraste, il faut souligner que tout programme chaîne réunit les conditions nécessaires à l'apparition du phénomène. On y trouve une alternance de composants associés à des fréquences de renforcement alimentaire très différentes, celle du premier terme étant toujours nulle. Dans notre cas, l'augmentation de la probabilité d'IRTS courts au niveau du deuxième terme reflète la discrimination par l'animal de ces conditions de renforcements différents. On peut faire également l'hypothèse que l'accélération du débit observée est une réaction apparentée à l'effet d'omission-frustration (Staddon, 1974). En effet, le programme chaîne se différencie du programme multiple de ligne de base par le fait que le renforcement alimentaire est « omis » à la fin du premier composant, d'où une réaction compensatoire, sous forme de réponses plus rapprochées, au niveau du second composant.

Dans quelle mesure les résultats décrits sont-ils favorables à l'hypothèse de départ ? D'un point de vue très global, on pourrait dire que cette hypothèse se trouve confirmée. Notre manipulation fournit à la réponse « coup de bec » un gain d'efficacité supérieur à celui de l'autre operant. Cependant, il faut remarquer que l'augmentation d'efficacité, mesurée en termes de pourcentage de réponses renforcées, ne correspond pas à une meilleure adaptation aux exigences du programme. Les intervalles inter-réponses subissent un allongement beaucoup trop important par rapport au délai exigé : il n'y a pas de régulation

temporelle. Il est possible que la réponse du premier terme de la chaîne ne soit pas sous contrôle d'un événement aussi positif que le renforcement alimentaire. A la limite, le programme chaîne placerait la réponse qui vient en premier dans une situation proche de l'extinction du fait que le renforcement alimentaire n'est plus délivré à ce niveau. Ceci expliquerait l'allongement des IRTS, ainsi que la perte de régulation temporelle. Le programme chaîne ne satisfait donc que partiellement à mise à l'épreuve de notre hypothèse. Des contrôles seraient nécessaires où l'on utiliserait un événement plus renforçant que l'apparition d'un stimulus conditionné.

Il apparaît néanmoins dans cette expérience que, placés dans les mêmes conditions, le coup de bec et l'appui de la patte ne se comportent pas de façon identique. Si la même tendance se retrouve chez les deux operants, elle est toujours moins marquée et présente plus de variabilité au niveau de la réponse d'appui.

En conclusion, l'existence d'un lien plus ou moins arbitraire entre la réponse et le renforcement résulterait en une différence de sensibilité à une modification des conditions renforçantes. La réponse naturellement en relation avec le renforcement alimentaire serait plus affectée par le remplacement de celui-ci par un stimulus conditionné.

RÉFÉRENCES

- BROWN, P.L., & JENKINS, H.M. Autoshaping the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior*, 1968, 11, 1-8.
- FERSTER, C.B., & SKINNER, B.F. *Schedules of reinforcement*. New York : Appleton, 1957.
- KRAMER, T.J., & RODRIGUEZ, M. The effect of different operants on spaced responding. *Psychonomic Science*, 1971, 25, 177-178.
- MANTANUS, H., JULIEN, I., & PITZ, I. Comparaison de deux réponses operantes chez le pigeon. *Psychologie Française* 1977, 22, 61-68.
- REYNOLDS, G.S. Behavioral contrast. *Journal of the Experimental Analysis of Behavior* 1961, 4, 57-71.
- STADDON, J.E.R. Some properties of spaced responding in pigeons. *Journal of the Experimental Analysis of Behavior* 1965, 8, 19-27.
- STADDON, J.E.R. Temporal control and the theory of reinforcement schedules. In : R.M. GILBERT & J.R. MILLENSON (Eds.), *Reinforcement : behavioral analyses*. New York and London : Academic Press 1972.
- STADDON, J.E.R. A note on behavioral contrast and frustration. *Quarterly Journal of Experimental Psychology*, 1974, 26, 285-292.
- THOMAS, J.R., & STUBBS, A. Stimulus control of spaced responding in second-order schedules. *Journal of the Experimental Analysis of Behavior* 1967, 10, 175-183.
- WEISS, B., LATIES, V.G., SIEGEL, L., GOLDSTEIN, D. A computer analysis of serial interactions in spaced responding. *Journal of the Experimental Analysis of Behavior*, 1966, 9, 619-626.
- WILTON, R.N., & GAY, R.A. Behavioral contrast in chained schedules. *Journal of the Experimental Analysis of Behavior* 1969, 12, 905-910.