



En el mundo real la contingencia perfecta entre la respuesta y el reforzamiento es rara. **No siempre que pasemos muchas horas estudiando obtendremos una buena nota.**

Para estudiar esta relación tan compleja se crean los programas de reforzamiento:

Programa de reforzamiento: Programa o regla que determina la forma en que se relacionan las presentaciones del reforzador con las ocurrencias de la respuesta instrumental. La entrega del reforzador puede depender de varios factores, número de respuestas, paso del tiempo, etc.

Los programas de reforzamiento se investigan normalmente en cajas de Skinner que permiten una observación continua de la conducta, por lo que pueden observarse y analizarse los cambios en la tasa de respuesta. Proporcionan más información sobre los factores que controlan el mantenimiento de la conducta.

A) PROGRAMAS SIMPLES DE REFORZAMIENTO INTERMITENTE

Son programas en los que un solo factor determina qué ocurrencia de la respuesta instrumental se refuerza.

Programas de razón

El reforzamiento depende del número de respuestas.

Reforzamiento continuo: Una sólo respuesta provoca la entrega del reforzador.

Reforzamiento parcial o intermitente: La respuesta se refuerza sólo alguna vez.

Razón fija: Entrega del reforzador cada X respuestas, (5, 10 etc). Se representa mediante RF5, RF10, etc.

Otros conceptos importantes

Registro acumulativo de la conducta: Es una forma especial de representar la forma de repetirse una respuesta a lo largo del tiempo. Se utilizaba sobre todo antes de la existencia de los ordenadores. (gráfico pág 165). La pendiente de la línea trazada representa la tasa de respuesta.

Pausa posreforzamiento: Es la tasa de respuesta 0 que se observa justo después del reforzamiento.

Carrera de la razón: La tasa de respuesta alta y estable que completa cada requerimiento de la razón.

¿Qué ocurre cuando el requerimiento de la razón se incrementa un poco? Es decir, pasar por ejemplo de RF 120 a RF 150

- Con mayores requerimientos de la tasa, tienen a ocurrir pausas posreforzamiento más largas.



CAPITULO 6: PROGRAMAS DE REZORZAMIENTO Y CONDUCTA DE ELECCIÓN

- Si el incremento es muy grande (de RF120 a RF500), es probable que el animal realice una pausa periódica cada vez que complete la razón requerida. Es lo que se conoce como tensión de razón.

Razón variable: Número de respuestas distinto para la entrega del reforzador. Se representa por RV y en el caso de RV10 quiere decir que su media es 10 (imaginemos que en un primer ensayo se necesitaban 13 respuestas, en un segundo ensayo 7 respuestas y en el ensayo final 10 respuestas. La media total de los 3 ensayos es RV 10.

Las pausas predecibles en la tasa de respuesta son menos probables con los programas de razón variable que con los programas de razón fija. Los organismos suelen responder a una tasa claramente estable en los programas de RV.

Programas de intervalo

Cuando ocurre cierta cantidad de tiempo desde la última entrega del reforzador.

Intervalo fijo (IF): El tiempo no varía de una ocasión a la siguiente.

Festón del intervalo fijo: A medida que el tiempo para la disponibilidad del próximo reforzador se acerca, la tasa de respuesta aumenta. Este aumento en la tasa de respuesta se manifiesta como una aceleración en el registro acumulativo hacia el final del IF. **Esto demuestra una habilidad del sujeto para contar el tiempo.**

OJO!!! : Un programa de IF no garantiza que el reforzador se proporcionará a en algún momento a lo largo del tiempo, el intervalo determina sólo cuándo el reforzador estará disponible, no cuando se entrega. **Es decir, además del paso del tiempo, sigue siendo necesaria algún tipo de respuesta por parte del sujeto para acceder al reforzador. Por ejemplo, cada 4 minutos estará disponible comida para la paloma siempre y cuando haya dado 20 picotazos previamente.**

Intervalo variable (IV): El tiempo varía de una ocasión a la siguiente. **Se representa por (IV 5), (IV 10), etc. Entendiendo este número como la media de todos los intervalos, de la misma manera que hablábamos más arriba de la razón variable.**

Los programas de IV mantienen tasas de respuesta constantes y estables sin pausas regulares. **(al igual que los programas de razón variable)**

Programas de intervalo y espera limitada: **Esto tiene relación con el OJO!! que marcaba antes.** En la espera limitada el reforzador está disponible durante un tiempo X, si pasado ese tiempo el sujeto no ha dado el número de respuestas necesario, el reforzador deja de estar disponible. **Siguiente el ejemplo anterior, cada 4 minutos estará la comida disponible durante 1 minuto. Si durante ese minuto de disponibilidad de la comida (+ los 4 minutos anteriores de cada intervalo) la paloma no ha llegado a 20 picotazos no podrá acceder a ella y dejará de estar disponible, comenzando a contar de 0.**



Comparación de los programas de razón y de intervalo

Similitudes

Razón fija e intervalo fijo:

- Producen una pausa posreforzamiento después de la administración de cada reforzador.
- Producen altas tasas de respuesta antes de la administración del siguiente reforzador.

Razón variable e intervalo variable:

- Mantienen tasas estables de respuesta, sin pausas predecibles.

Diferencias

- Activan diferentes cambios neuroquímicos en el cerebro
- Los programas de intervalo están relacionados con el sentido temporal del organismo.

En un experimento con palomas, una reforzada con RV y la otra con IV, se comprobó que la paloma reforzada con el programa de RV respondió a una tasa mucho más alta que la paloma reforzada con el programa de IV. El programa de RV motivó una conducta instrumental mucho más vigorosa.

Reforzamiento diferencial del tiempo interrespuesta: El intervalo entre una respuesta y la siguiente se denomina tiempo interrespuesta o TIR. Un sujeto que produce TIR cortos responderá a una tasa alta y si produce TIR largos, su tasa será baja.

Un programa de razón favorece los TIR cortos (es la respuesta del sujeto la que determina cuando accederá a la comida, por lo tanto cuanto más rápida se produzca esa respuesta (TIR corto), antes accederá a la comida). Por el contrario, un programa de intervalo favorece los TIR largos.

Programas de reforzamiento de tasas de respuesta

Los programas de tasas de respuesta requieren específicamente que el organismo responda a una tasa particular para obtener el reforzador. Es como una mezcla de los dos anteriores (tiempo y número). Por ejemplo un programa de reforzamiento en el que una respuesta se refuerce sólo si ocurre dentro de los 5 segundos siguientes a la respuesta precedente. Si el animal realiza una respuesta cada 5 segundos, su tasa de respuesta será de 12 por minuto. Por lo tanto habrá reforzamiento si la tasa de respuesta es de 12 por minuto o mayor. Si la tasa de respuesta es menor, el organismo no será reforzado. Esto recibe el nombre de **Reforzamiento diferencial de tasas altas, o RDA.**

Reforzamiento diferencial de tasas altas (RDA): Una respuesta se refuerza sólo si ocurre antes de que haya transcurrido cierta cantidad de tiempo tras la respuesta precedente. Estos programas fomentan que los sujetos respondan rápidamente.

Reforzamiento diferencial de tasas bajas (RDB): Una respuesta se refuerza sólo si ocurre después de que haya transcurrido cierta cantidad de tiempo tras la respuesta precedente. Estos programas fomentan que los sujetos respondan lentamente.



B) CONDUCTA DE ELECCIÓN: PROGRAMAS CONCURRENTES

Los experimentos en los que sólo se mide una respuesta (los vistos hasta ahora) no proporcionan una visión completa de la conducta. En la vida real los organismos ocupan su tiempo en muchas actividades y están continuamente estableciendo elecciones entre varias conductas.

Las situaciones de elección pueden ser bastante complicadas y muy diferentes. Por esta razón los psicólogos han desarrollado el estudio de situaciones más simples. La situación de elección más simple consta de dos respuestas alternativas, cada una de estas seguida por un reforzador de acuerdo con su propio programa de reforzamiento.

Históricamente se utilizaban **laberintos T** para estudiar la conducta de elección. Hoy se utilizan **cajas de Skinner** equipadas con dos teclas. Responder en cada tecla es reforzado según algún programa de reforzamiento. Los dos programas funcionan concurrentemente, por eso reciben el nombre de **programa concurrente**.

Ejemplo página 173 (gráfico). Programa concurrente. Los picotazos sobre la tecla A siguen un programa IV 60 seg. Los de la tecla B siguen un programa RF 10. El objetivo del experimento es ver cómo la paloma distribuye sus picotazos en las dos teclas y cómo influye el programa de reforzamiento propio de cada tecla en las elecciones del animal.

Medidas de la conducta de elección

Se puede medir de varias formas:

Tasa relativa de respuesta de cada alternativa: La tasa relativa de respuesta en la tecla A se calcula dividiendo la tasa de respuesta en la tecla A por la tasa total de respuesta (suma de la tasa A + la tasa B)

$$\text{Tasa relativa de respuesta (A)} = R_A / (R_A + R_B)$$

- Si la paloma picotea igualmente en las dos teclas de respuesta, la razón será 0,5.
- Si la tasa de respuesta es mayor en A que en B, la razón será mayor que 0,5.
- Si la tasa de respuesta es menor en A que en B, la razón será menor que 0,5.

La forma en que un organismo distribuye su conducta entre las dos alternativas de respuesta está muy influida por el programa de reforzamiento que funciona para cada respuesta.

Tasa relativa de reforzamiento: Es la frecuencia con la que el individuo obtiene recompensas en cada lado. Se calcula de igual manera a la tasa relativa de respuesta. (pero con una "r" minúscula)

$$\text{Tasa relativa de reforzamiento (A)} = r_A / (r_A + r_B)$$



La ley de igualación (de Herrnstein)

En el punto anterior (medidas de la conducta de elección) en los ejemplos se utilizaba el mismo programa de reforzamiento (IV 60 seg) para las dos teclas. Por lo tanto la tasa relativa de respuesta y la tasa relativa de reforzamiento daba 0,5 ya que los programas eran iguales... ¿Pero qué ocurre si los programas son diferentes?

Experimento en palomas:

Tecla A: IV 6 min = Pulsando siempre a esta tecla obtendría un máximo de 10 reforzadores por hora.

Tecla B: IV 2 min= Pulsando siempre a esta tecla obtendría un máximo de 30 reforzadores por hora.

¿Cómo fue la respuesta de las palomas?

La tasa relativa de respuesta en una alternativa dada estaba siempre muy cercana a igualar la tasa relativa de reforzamiento logrado en esa alternativa. Si las palomas obtenían una proporción mayor de reforzadores en A, picoteaban más en A. Por lo tanto, la tasa relativa de respuesta en una alternativa se iguala con la tasa relativa de reforzamiento en esa alternativa. **Esto es la ley de igualación.**

$$R_A / (R_A + R_B) = r_A / (r_A + r_B)$$

Otra forma de escribir esta fórmula es: $R_A / R_B = r_A / r_B$

Infraigualación, supraigualación y sesgo de respuesta

A pesar de lo que hablábamos de la ley de igualación, las tasas relativas de respuesta no siempre se igualan con exactamente con las tasas relativas de reforzamiento.

Esto lo vamos a estudiar añadiendo dos parámetros “b” y “s” a la ecuación.

$$R_A / R_B = b (r_A / r_B)^s$$

s= sensibilidad de la conducta de elección a las tasas relativas de reforzamiento para las alternativas de respuesta.

Si s es menor que 1,0 = **infraigualación**. Sensibilidad reducida a la tasa relativa de reforzamiento.

Si s es mayor que 1,0 = **supraigualación**. Sensibilidad aumentada a la tasa relativa de reforzamiento.

La infraigualación se encuentra más a menudo que la supraigualación.

b = sesgo de respuesta (bias en inglés, por eso la b)

Los sesgos de respuesta influyen en la elección cuando las alternativas de respuesta son diferentes (picotear y presionar con la pata). El sesgo también adquiere importancia cuando el reforzador es diferente (grano de trigo y semillas).

En ausencia de sesgos, b=1. Dependiendo de la naturaleza del sesgo o de la preferencia, b será mayor o menor que 1.



CAPITULO 6: PROGRAMAS DE REZORZAMIENTO Y CONDUCTA DE ELECCIÓN

La ley de la igualación y el valor del reforzador: La tasa relativa de respuesta es una función de la cantidad relativa de cada reforzador, de la demora y de la palatabilidad. Cuanto mayores, más sabrosos y más inmediatos sean los reforzadores, serán de mayor valor.

La ley de la igualación y los programas simples de reforzamiento: Incluso las situaciones de respuesta única pueden involucrar una elección entre realizar la respuesta especificada (picotear o presionar una palanca) u ocuparse de otras actividades (caminar, acicalarse, etc)

Según esto, la ley de la igualación quedaría así:

$$R_A / (R_A + R_O) = r_A / (r_A + r_O)$$

Despejando $R_A =$ $R_A = (R_A + R_O) r_A / (r_A + r_O)$

- R_A = Tasa de respuesta del programa
- R_O = Tasa de respuesta de las otras actividades
- r_A = Reforzamiento explícito del programa
- r_O = Reforzamiento de las otras actividades

Y para liar un poco más las cosas...

Si consideramos que $(R_A + R_O)$ es una constante que no guarda relación con el reforzador, podemos llamarla constante k. La fórmula quedará de la siguiente manera:

$$R_A = k r_A / (r_A + r_O)$$

Esta última ecuación proporciona dos maneras de modificar la tasa de una respuesta:

- Mediante el cambio de su tasa de reforzamiento
- Mediante el cambio de la tasa de las otras fuentes de reforzamiento

Mecanismos de la ley de la igualación

3 mecanismos:

1. Maximización molecular
2. Maximización molar
3. Mejoramiento

Antes de entrar en detalle con cada uno de ellos. Vamos a explicar un concepto importante para entender el 1 y el 2:

Igualación y maximización de las tasas de reforzamiento: Los organismos distribuyen sus acciones entre las alternativas de respuesta para recibir la máxima cantidad posible de reforzadores en cada situación.

Y ahora sí, pasamos a explicar cada uno de los mecanismos:



CAPITULO 6: PROGRAMAS DE REZORZAMIENTO Y CONDUCTA DE ELECCIÓN

1. Maximización molecular: Los organismos siempre eligen cualquier alternativa de respuesta con mejores probabilidades de que sea reforzada en ese momento.

Experimento en palomas con dos teclas (A y B) y programa IV-IV. Cuanto más tiempo permanezca la paloma en la tecla A, más probabilidad habrá de que transcurra el intervalo para la tecla B y que por lo tanto el reforzamiento esté disponible para B.

2. Maximización molar: El organismo distribuye sus respuestas entre las varias alternativas a fin de maximizar la cantidad de reforzamiento que consiguen a la larga. Se utilizó para explicar los programas de razón en los que los organismos rara vez cambian de alternativa, simplemente responden a la alternativa que requiera menor número de respuestas (RF10 frente a RF20)

Sin embargo la maximización molar encuentra problemas a la hora de explicar la conducta de elección en programas de intervalo (IV-IV) y para programas razón-intervalo (RF-IV).

3. Mejoramiento: Los mecanismos de mejoramiento operan a una escala temporal de las teorías molas y molecular. El término “mejoramiento” se refiere a hacer algo mejor que el pasado reciente. No se refiere a elegir la mejor alternativa en ese momento.

Término importante sobre el mejoramiento:

Tasa local de respuesta y reforzamiento: Las tasas locales están calculadas sólo sobre el período de tiempo que un sujeto dedica a una alternativa particular de respuesta. Si la situación incluye dos opciones (A y B), la tasa local de respuesta en A se calcula dividiendo la frecuencia de respuesta en A por el tiempo que el sujeto dedica a responder en A.

Ejemplo numérico (que siempre ayudan)

- Un organismo da 75 respuestas en 20 minutos en la alternativa A
- La tasa global en A será de 75 respuestas/hora (puede que el resto del tiempo se haya dedicado a responder en B)
- Sin embargo la tasa local en A sera de 225 respuestas/hora (75 cada 20 minutos)

La teoría del mejoramiento asume que los organismos cambian de una alternativa de respuesta a otra para mejorar la tasa local de reforzamiento que están recibiendo. El mecanismo de mejoramiento da como resultado la igualación.

Esta teoría sostiene que un organismo cambiará su preferencia en favor de la alternativa de respuesta que arroje la tasa local de reforzamiento más alta.

C) ELECCIÓN CON COMPROMISO

En nuestra vida diaria, en ocasiones elegir una alternativa hace que otras no estén disponibles. De igual forma, las elecciones importantes en la vida a menudo implican un beneficio pequeño a corto plazo frente a un beneficio más demorado pero mayor.

Es complicado extrapolar este tipo de decisiones al laboratorio para poder estudiarlas, por ello los científicos han desarrollado el **programa concurrente encadenado** de reforzamiento.



Programa concurrente encadenado (gráfico pág 183)

Este programa involucra al menos dos fases o eslabones:

1ª fase: Eslabón de elección: Se permite al participante elegir entre dos programas alternativos realizando una de dos respuestas.

2ª fase: Eslabón terminal: Una vez que el participante ha elegido uno de los dos eslabones de elección, irá al eslabón terminal que está formado por dos programas de reforzamiento dependiendo de qué opción haya escogido en la 1ª fase. (esto es como lo que comentábamos antes de la vida real, en ocasiones elegir una opción hace que otras opciones no estén disponibles)

Los estudios de este tipo han demostrado que los sujetos prefieren la alternativa de razón variable (RV) ya que proporciona reforzamiento por un número relativamente pequeño de respuestas.

El estímulo del eslabón terminal se convierte en un reforzador condicionado. Por lo tanto un programa concurrente se caracteriza en que las respuestas al eslabón inicial se refuerzan por la presentación de un reforzador condicionado. Las diferencias en el valor del reforzador condicionado determinarán entonces la tasa relativa de cada respuesta de elección en el eslabón inicial. (el programa (eslabón terminal) que suscite más valor a la paloma será el que elija al principio en el eslabón de elección)

Estudios de autocontrol

El autocontrol a menudo supone elegir una recompensa grande demorada en lugar de una recompensa pequeña inmediata.

Experimento en palomas: (gráfico pág 185)

Utilizando el programa concurrente encadenado. En cada uno de los programas terminales responder era recompensado o bien mediante el acceso a una pequeña cantidad de grano (alternativa A) o el acceso a una mayor cantidad de grano que estaba demorada 4 segundos (alternativa B).

Estudiaron la conducta de elección en dos circunstancias diferentes:

- 1. Procedimiento de elección directa:** (dibujo superior de la 185). No existe eslabón de elección, directamente se accede a la recompensa. En este ejemplo las palomas no mostraban autocontrol y directamente seleccionaban la recompensa pequeña inmediata.
- 2. Procedimiento concurrente encadenado** (dibujo inferior de la página 185, la foto de la tía no, sino el de la paloma). En este ejemplo, los componentes terminales estaban demorados después de que las palomas realizaran su elección inicial. Si la demora era suficiente antes de los componentes terminales, las palomas mostraban autocontrol y elegían la recompensa grande demorada.

Conclusiones: (este párrafo lo copio tal cual del libro sin resumir porque me parece interesante e importante): Las preferencias cambian a favor de cualquier recompensa



CAPITULO 6: PROGRAMAS DE REZORZAMIENTO Y CONDUCTA DE ELECCIÓN

grande demorada si se requiere que los participantes esperen más tiempo para recibir cualquier recompensa tras realizar su elección. Si las recompensas se entregan rápidamente tras una respuesta de elección, los sujetos generalmente prefieren una recompensa pequeña inmediata más que una recompensa grande demorada. Sin embargo, si se añade una demora constante a la entrega de ambas recompensas, es más probable que los individuos muestren autocontrol y prefieran la recompensa grande demorada.

Explicaciones del autocontrol:

El valor de un reforzador se reduce en función del tiempo que haya que esperar para obtenerlo. La función matemática que representa esto se llama función descontadora del valor.

$V = M / (1+KD)$ (esta función recibe el nombre de función del decaimiento hiperbólico)

V= valor del reforzador

M= Magnitud de la recompensa

D= Demora de la recompensa

K= Parámetro de la tasa descontadora

Si el reforzador es entregado sin demora ($D=0$), el valor del reforzador estará directamente relacionado con su magnitud. Cuanto más se demora el reforzador menor es su valor.

La función descontadora del valor explica el problema del autocontrol en el gráfico de la página 186. Me ha costado un poco entenderlo pero es bastante sencillo. Tenemos una línea horizontal que es el tiempo y una vertical que es el valor de la recompensa. También tenemos 2 momentos en el tiempo ($T1$ y $T2$). Dependiendo de qué momento escojamos elegiremos una recompensa u otra. Así por ejemplo en $T1$ vemos que la línea que lleva a la recompensa pequeña está por encima (tiene más valor) que la línea que lleva a la recompensa grande. Por lo tanto en $T1$ escogeremos la recompensa pequeña. Por el contrario en $T2$ la línea que lleva a la recompensa grande está por encima y escogeremos esta opción. La opción $T2$ es la opción de autocontrol

Las funciones descontadas y los problemas de autocontrol en la adicción a la droga:

Cuando más acusada sea la función descontadora de la demora de una persona, más dificultad tendrá para mostrar autocontrol (seleccionando una recompensa mayor demorada) en lugar de impulsividad (seleccionando una recompensa más pequeña pero más rápida).

Experimentos para medir el autocontrol en personas dependientes de la heroína demostró que tenían más impulsividad que el resto de las personas. (gráfico de la izquierda en 187). Para los dependientes de la heroína, el valor del dinero (con el que se realizaba el experimento) disminuyó muy rápido si iban a tardar en dárselo.



¿Puede entrenarse el autocontrol?

La respuesta es sí.

Ejemplo con niños. 3 fases.

Fase 1- Pretest: Se les pregunta a los niños si quieren 2 centimos ahora o 3 centimos al final del día. A los que eligen los 2 céntimos se les dan enseguida, a los que eligen los 3 centimos se les dan al final del día.

Fase 2- Sesión de entrenamiento: Dependiendo de la elección en la fase 1, se hacen 2 grupos y se comienza con la sesión de entrenamiento. Deben resolver problemas sencillos para obtener la recompensa, las respuestas correctas del grupo de los 2 céntimos recibirán el dinero al instante. Las respuestas correctas del grupo de los 3 céntimos recibirán el dinero al final del día.

Fase 3 – Postest: se les vuelve a preguntar a los niños lo mismo que en la fase 1.

El hallazgo fue que había mayor cantidad de niños escogiendo la recompensa más grande demorada (3 céntimos) en el postest. Por lo tanto, siempre y cuando las tareas de entrenamiento (fase 2) impliquen bajo esfuerzo, el entrenamiento con la recompensa demorada incrementará la preferencia por la recompensa más grande demorada.

Todo esto del autocontrol lo encuentro bastante relacionado con alguno de mis artículos:

- **“Voluntad vs Hábito: el desafío de los 30 días”** (21 Octubre 2009)
- **“Cómo aumentar tu autodisciplina”** (26 y 30 Octubre 2009)