

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
FACULTAD DE PSICOLOGIA



APRENDIZAJE Y DESARROLLO DE PREFERENCIAS
EN PROGRAMAS INTERTEMPORALES.

T E S I S
QUE PARA OBTENER EL GRADO
DE MAESTRIA EN ANALISIS EXPERIMENTAL
DE LA CONDUCTA

PRESENTA:

OSCAR ZAMORA AREVALO

DIRECTOR DE TESIS:
DR. ARTURO BOUZAS RIAÑO

Tesis apoyada por los proyectos INC307396 DE DGAPA y 5-4676H de CONACYT



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

DE LAS DEDICATORIAS ...

A Renata con todo el corazón a 7 años de estar juntos gracias a un prodigio de casualidades.

A Arantxa y Amos Sebastián por ser chispa y razón para seguir y porque con ellos nacen y crecen los sueños.

A los jefes, carnales y familia en general ... a pesar de su desconfianza.

A Francisco Javier Coronado Moreno (Pachís) q.e.p.d. por haber sabido ser amigo y guía, pero sobretodo por haber sabido ser familia.

DE LOS AGRADECIMIENTOS Y JUSTIFICACIONES...

Hay tesis hechas de una sola pieza y tesis que como ésta, están hechas por piezas. Hay tesis pensadas como tales desde el principio y hay tesis que como ésta se conciben así después de haberlas escrito. Si en todos los casos se espera que el autor justifique su tesis y muestre cual es su hilo conductor, tal demanda parece ser mayor en casos como el presente. Por otro lado la vida académica le impone a uno, con frecuencia, toda una serie de restricciones, que unidas a la propia capacidad o incapacidad, según se vea, hacen difícil escribir una tesis de principio a fin de manera satisfactoria. Por otra parte estas mismas restricciones, suelen estar asociadas a la necesidad de escribir un buen número de piezas sueltas que no siempre son utilizadas en el trabajo final. A parte de lo que uno dedica a tomar clases y a la investigación como tal, se suele invertir bastante tiempo en hacer informes, preparar nuevos proyectos, presentar trabajos, cumplir con los tramites burocráticos que exige el poder seguir en un postgrado, etc. Hasta que llega el momento en que uno mira hacia atrás, revisa los papeles y se cuestiona por la conexión que puede haber entre unas cosas y otras. Y entonces puede suceder que uno se sorprenda gratamente al comprobar cierta coherencia y continuidad con lo que se ha producido al momento actual. Junta uno las piezas -que a veces esperan esparcidas o inlocalizables en los archivos y trata uno de presentarlas en forma de marco teórico. Esto es, pues el contexto en el que surge la idea de presentar esta tesis... Para mi fortuna tuve la suerte de contar con el apoyo (tanto moral, afectivo, como académico) de personas amables que siempre me brindaron su ayuda incondicionalmente, participando generosamente con sus conocimientos, críticas y comentarios, a todos ellos les agradezco su imprescindible presencia y contribución, sería ocioso poner el nombre de todos y cada uno de ellos, pero bien saben a quien me refiero (Coloquios de Coyoacan finales del 96, para más referencias). También agradezco a las autoridades de la Facultad de Psicología por las facilidades y comisión otorgada para ausentarme de mis labores y permitir así la culminación de esta tesis. Mi deuda con el Dr. Arturo Bouzas salta a la vista de cualquiera que se desenvuelva en estos círculos, aun así me complace reconocerla expresamente aquí. A todos y cada uno de mis sinodales, Dr. Florente López, Dr. Javier Nieto y Mtro. Gustavo Bacha quienes contribuyeron notablemente con sus comentarios e indicaciones a la mejora de este trabajo y muy especialmente al Mtro. Fernando Vázquez quien corrigió y pulió este trabajo con un grado de afecto y competencia tal, que quien solo vio el trabajo antes y después puede apreciar. Finalmente hay

muchas otras personas que han contribuido directa o indirectamente a la realización de esta tesis (muy a su pesar) y el no mencionarlos aparte podría considerarse una ofensa a la gratitud. Al grupo de trabajo del Laboratorio de Comportamiento y Adaptación, Elena Durón, Adela Mendoza, Miriam Calderón, Diana Posadas, Julio Espinosa, etc. (los de antes) Vladimir Orduña, Ruben González y *et al* (los de después), al Mike Herrera por todos estos años (casi 10, aunque más bien debería de agradecer a todo el personal del laboratorio y a mi mismo por aguantarlo tanto tiempo...es bromaii, es bromaii), al grupo de Trabajo del Laboratorio de Farmacología Conductual y a su jefe muy en especial, el Dr. David Velázquez por el apoyo y la ayuda que siempre incondicionalmente me ha brindado, a los cuates de escuela y party crazy -Joven Alet´s (y dama que lo acompaña) por encontrar el poco humor y desparpajo que aún había en mi, al profesor Desorden, a Pelonius Monk, y a Hellen of Hell (por las versiones corregidas y aumentadas de los maratones Guadalupe-Reyes-Candelaria y por estar ahí siempre (a pesar de las restrcciones espaciales) preocupandote y procurandome, a los ezquisos y acapulcos por alegrar y ser parte también de las party crazys, a todos ellos **Muchas Gracias**.

Aprendizaje y Desarrollo de Preferencias

Resumen

Un resultado fundamental que se ha encontrado en la literatura contemporánea en condicionamiento operante bajo condiciones de estado estable, es que las tasas relativas de respuesta igualan las tasas relativas de reforzamiento, un aspecto al cual se le ha intentado dar respuesta es, cual es el curso que sigue este ajuste. En contraste con el interés en conducta en estado estable, ha habido pocos reportes publicados describiendo la conducta durante periodos de transición -aquellas ocasiones cuando un sujeto es expuesto por primera vez a un nuevo par de programas y esta conducta no se ha estabilizado. Este tipo de estudios son escasos (Bailey y Mazur, 1990; Mark y Gallistel, 1994; Kacelnick, Krebs y Ens, 1987; Killen 1972; Mazur, 1992, 1995; Myerson y Hale, 1988), sin embargo una extensa base de datos sobre conducta de elección durante periodos de transición ayudaría seguramente a acrecentar nuestro entendimiento de la elección y la preferencia. Por otro lado una característica que poseen los programas concurrentes es la competencia entre las opciones de respuesta por el tiempo disponible para responder, esta competencia permite que incrementos en la fuerza de una respuesta, debidos al reforzamiento, den lugar automáticamente a decrementos en la fuerza de otras respuestas. De tal manera que para estudiar la interacción entre reforzadores de manera estricta, es necesario separar las respuestas de manera que los efectos de la interacción no ocurran por la competencia del tiempo común, la manera más simple de llevar a cabo lo anterior es por medio de los programas múltiples, donde las respuestas a diferentes opciones se emiten en componentes temporalmente alejados, de forma tal que no pueden competir por un tiempo común. Sabemos que la tasa de reforzamiento en un período afecta la tasa de respuestas en otro. De manera análoga al fracaso de los modelos que no permiten interacciones del valor de reforzadores obtenidos en diferentes periodos para dar cuenta de la ejecución en estabilidad, los modelos de aprendizaje que asumen que cada respuesta es controlada por el reforzador que produce, no pueden dar cuenta de la adquisición del comportamiento en situaciones de sucesión temporal (programas múltiples). En la actualidad, aún no se cuenta ni con la evidencia empírica pertinente al problema de la adquisición en programas múltiples ni con un conjunto de modelos explícitos acerca de los algoritmos relevantes. De tal manera, el propósito del presente estudio es dilucidar el problema de la adquisición de conductas en programas múltiples (elección sucesiva), así como revisar una serie de modelos explícitos acerca de los algoritmos relevantes que podrían dar cuenta de este tipo de hallazgos. Así como evaluar el curso de los ajustes en el comportamiento ante cambios abruptos, impredecibles y no señalados, cuando las tasas de reforzamiento ocurren en distintos periodos de tiempo. En el presente experimento estudiamos con 8 pichones la respuesta de picoteo en un programa múltiple. Cada condición inició con tres días en *fases de una probabilidad* o sesiones de *entrenamiento* en ellas un programa simple RV asignaba reforzadores a los dos componentes del programa múltiple. Seguida de dos sesiones de *transición* o *fase de dos probabilidades*, donde un par de probabilidades (RVx, RVy), operaban. En la primera parte del experimento, usamos en las fases de dos probabilidades pares de probabilidades de reforzamiento que tenían razones iguales de 5 a 1, pero diferentes valores absolutos (.16, .08 y .04). Se evaluó el papel que los distintos valores absolutos tuvieron en la tasa de adquisición, encontrándose que esta última era mas rápida con los valores absolutos menores (.04 y .08). En una segunda parte, se usaron pares de probabilidad de reforzamiento de una razón de 2 a 1 y 3 a 1 con valores absolutos iguales (.04) y se evaluó el efecto de diferencias constantes en probabilidad con razones distintas. Encontrándose que la adquisición de una preferencia ocurre más rápidamente con razones mayores aún cuando la diferencia entre las dos probabilidades de reforzamiento sean constantes.

INDICE	
INTRODUCCION	1
Programas Múltiples (Elección Intertemporal)	11
METODO	13
RESULTADOS	16
DISCUSION Y CONCLUSIONES	31
ANEXO 1	39
Introducción a Manera de Justificación	39
Aprendizaje y Acción	40
Estudios de Elección	41
ANEXO 2	46
Niveles del Sistema Nervioso	46
Niveles de Análisis	46
ANEXO 3	49
¿Cómo se asocian los estímulos condicionados e incondicionados? Un acercamiento al Modelo de Rescorla y Wagner	49
El Modelo de Condicionamiento de Rescorla y Wagner	54
Explicación del Modelo de Rescorla y Wagner	56
Problemas con el Modelo de Rescorla y Wagner	56
Comprensión Neuronal: La Regla Delta	60
REFERENCIAS	62

I learned this, at least, by my experiment, that if one advances confidently in the direction of his dreams, and endeavors to live the life which he has imagined, he will meet with a success unexpected in common hours. He will put some things behind, will pass an invisible boundary; new, universal, and more liberal laws will begin to establish themselves around and within him; or the old laws be expanded, and interpreted in his favor in a more liberal sense, and he will live with the license of a higher order of beings. In proportion as he simplifies his life, the laws of the universe will appear less complex, and solitude will not be solitude, nor poverty poverty, nor weakness weakness. If you have built castles in the air, your work need not be lost; that where they should be. Now put the foundations under them.

The Portable Thoreau
Walden, 1854
Henry David Thoreau

Aprendizaje y Desarrollo de Preferencias

INTRODUCCION

De los numerosos estudios que sobre elección y preferencia se han realizado en las últimas tres décadas (Kagel, 1987; Stephens y Krebs, 1986; Williams, 1988; Davison y McCarthy, 1988), la vasta mayoría: **i**) han utilizado un modelo experimental que consiste en el estudio de la distribución de dos respuestas, cada una reforzada de acuerdo a programas de reforzamiento independientes (programas concurrentes, programas concurrentes encadenados y procedimientos de ensayo discreto) y **ii**) se enfocan sobre conducta de estado estable. Sin embargo, en los últimos años ha habido un creciente interés acerca de como se desarrollan tales preferencias en una situación de elección. Un sinnúmero de diferentes aproximaciones teóricas para la adquisición de conducta de elección han sido propuestas. Cuando las dos alternativas son programas de reforzamiento intervalo variable (**IV**), la proporción de respuestas de los sujetos con frecuencia iguala a la proporción de reforzamientos (Herrnstein, 1970, ver Anexo 1, para una revisión mas extensa), sin embargo varias desviaciones de igualación perfecta son comunes (Baum, 1974, para una revisión Williams, 1988). Cuando las dos alternativas son programas de razón variable (**RV**), los animales muestran preferencia casi exclusiva por la alternativa con las razones más pequeñas, sin embargo preferencia no-exclusiva para la mejor alternativa también es observada en algunas ocasiones (Herrnstein, y Loveland, 1975; Herrnstein y Vaughan, 1980). Estas generalizaciones se aplican en conducta de elección en estado estable, ya que describen casos donde los sujetos han sido expuestos a las dos alternativas por un tiempo suficiente, hasta que sus elecciones no muestran cambios sistemáticos de sesión a sesión. Comparada con la conducta de elección en estado estable, la adquisición de conducta de elección no ha sido bien entendida. Sin embargo, en años recientes ha habido un incremento en el interés en conducta de elección en estados de transición — períodos donde las contingencias de reforzamiento han sido cambiadas, y la conducta de los sujetos esta en proceso de adaptarse a las nuevas contingencias. Dichos experimentos han usado una variedad de diferentes especies y procedimientos experimentales (por ejemplo, Couvillon y Bitterman, 1985, 1988, 1991; Dreyfus 1991; Kacelnik, Krebs y Ens, 1987; Mark y Gallistel, 1994) y han descubierto un número de valiosos hallazgos acerca de conducta de elección en transición. Teóricamente el reto ha sido el encontrar un procedimiento computacional (algoritmo) que sea simple y que describa el curso de los ajustes en el comportamiento durante periodos de adquisición y transición.

Aprendizaje y Desarrollo de Preferencias

Alrededor del tiempo en el que en investigación operante Herrnstein publicaba sus primeros resultados acerca de la ley del efecto relativo, un grupo de investigadores en condicionamiento clásico, entre ellos principalmente Kamin (1969), Rescorla y Wagner (1972) (para una revisión más completa ver Anexo 3), reportaron los resultados de un conjunto de estudios de aprendizaje asociativo que transformaron completamente las concepciones de esta área. Dicho brevemente, sus resultados indican que la contigüidad entre un estímulo y un reforzador no es una condición ni suficiente ni necesaria para el aprendizaje (Revisar, Pappini y Bitterman (1990) y Plaud y Volgelartz (1991), para otra visión del tema.). En uno de los estudios más interesantes manteniendo la probabilidad de reforzamiento constante en la presencia de un estímulo, Rescorla (1972), encontró que la fuerza asociativa de este dependía de la probabilidad de reforzamiento en su ausencia. Para dar cuenta de estos y otros resultados similares, Rescorla y Wagner (1972) propusieron un modelo alrededor del cual ha girado la literatura en esta área en los últimos veinticinco años.

Ambas literaturas muestran la multidimensionalidad¹ del aprendizaje. El contexto de reforzamiento ganó prominencia en la adquisición de la conducta o del control de estímulos. En ambos casos, la opción natural era asumir que los estímulos que controlan la conducta son agregados. Esta fue la opción considerada por Herrnstein en el caso del estudio de la elección, por Heline y Herrnstein (1966) en el caso del estudio de la evitación y más generalmente por Baum (1979). Resulta sorprendente que esta es la propuesta más reciente desde una perspectiva computacional presentada por Gallistel (1990). Sin embargo, la gran mayoría de las explicaciones en ambas literaturas han evitado escrupulosamente el trabajar con variables de agregados. En palabras de Rescorla y Wagner:

“Le presentamos al animal acontecimientos individuales, no correlaciones o información; y una teoría adecuada debe detallar como es que estos acontecimientos individuales afectan al animal. Es decir, necesitamos una teoría basada en acontecimientos individuales.”

1. Multidimensionalidad postula que el cambio en la conexión asociativa entre un EC y un EI en un cierto ensayo depende de la fuerza del pre-ensayo de las asociaciones entre el EI y todos los ECs activos, es decir, los otros ECs presentes en el ensayo. Este postulado ha sido utilizado para explicar, ensombrecimiento, bloqueo, pre-condicionamiento, y los efectos de suficiencia predictiva, todos los cuales demuestran que lo que un animal aprende acerca de una relación entre ese EI y los otros ECs. La esencia de este hallazgo es que dicha asociación, no se desarrolla entre un cierto EC y el EI si no que otros ECs bastan para predecir la ocurrencia del EI.

Aprendizaje y Desarrollo de Preferencias

En resumen, existe una amplia evidencia experimental que muestra que la conducta de los organismos se ajusta finamente a las tasas relativas de ocurrencia de reforzadores. La casi totalidad de los investigadores trabajando en este problema ha evitado suponer que los organismos computan tasas de ocurrencia de sucesos. En su lugar se han propuesto modelos de condicionamiento, en los que se asume que la fuerza de respuesta es determinada por efectos episódicos de incremento-decremento de reforzamiento-extinción (Davis, Staddon, Machado y Palmer, 1993; Killen, 1994; Williams, 1994b). Estos modelos tienen su origen en los operadores lineales introducidos en la literatura psicológica del aprendizaje por Bush y Mosteller (1955) y descritos más adelante en el contexto de la literatura operante por Davis *et al.* (1993)

Para entender el modelo de Operador Lineal, considere una situación experimental que ilustre la conducta adaptativa de un pichón en una caja de condicionamiento operante con dos teclas de respuesta en un programa concurrente, en donde se tienen dos conductas alternativas posibles para el individuo y dos posibles consecuencias en el ambiente. Así tenemos cuatro posibles eventos: (1) picar la tecla izquierda y recompensa; (2) picar la tecla izquierda y no recompensa; (3) picar la tecla derecha y recompensa; (4) picar la tecla derecha y no recompensa. Este modelo trata de explicar la forma en la cual la probabilidad de picar la tecla derecha $P_D(t)$ y la probabilidad de picar la tecla izquierda $P_I(t)$ cambian con el tiempo a $P_D(t+1)$ y $P_I(t+1)$, dependiendo de la consecuencia que ocurra.

De acuerdo al modelo, al inicio de la primera sesión, podemos asumir que el animal tiene una probabilidad de .50 de picar la tecla derecha, $p_1 = 0.50$. Ahora supóngase que el pichón pica la tecla derecha y es reforzado. Lo que se esperaría es que la probabilidad de esa respuesta incrementara. Suponga que en el siguiente ensayo el pichón pica la tecla izquierda y el resultado es negativo, esto es, no recibe recompensa o incluso recibe algún castigo. Bajo estas circunstancias se esperaría que la probabilidad de que el animal picara la tecla derecha en el tercer ensayo, p_3 , sería mayor que p_2 .

Podemos asumir que el efecto picar la tecla derecha y ser reforzado en el ensayo n da como resultado un incremento en p_n , que es proporcional a la cantidad que deba ser aprendida, $1-p_n$, digamos $a(1-p_n)$. Y en este caso, la probabilidad en el ensayo $n+1$ er ensayo es:

Aprendizaje y Desarrollo de Preferencias

$$\begin{aligned} P_{n+1} &= P_n + a(1-p_n) \\ P_{n+1} &= (1-a)p_n + a; \end{aligned} \quad (1)$$

es decir, la probabilidad en el n -ésimo ensayo se desarrolla por esta transformación lineal y se convierte en la probabilidad en el siguiente ensayo. Entonces esta expresión es llamada *operador lineal* y puede ser aplicada en cualquier ensayo en el que el animal pique la tecla derecha y sea reforzado.

El modelo de operador lineal se puede aplicar directamente a proporciones de respuesta. Sin embargo, no hay evidencia de que los organismos aprendan proporciones, pero si hay mucha evidencia que apoya que aprenden respuestas particulares. El modelo incluye dos ecuaciones, una para las respuestas a la tecla derecha (D) y otra para las respuestas a la tecla izquierda (I), con una regla para combinarlas:

$$\begin{aligned} V_{I(t+1)} &= aV_{I(t)} + (1-a)X_{I(t+1)} \quad \text{y} \\ V_{D(t+1)} &= aV_{D(t)} + (1-a)X_{D(t+1)} \end{aligned} \quad (2)$$

Con la regla de respuesta: $s = V_D / (V_D + V_I)$

Donde $V_{I(t+1)}$ y $V_{D(t+1)}$ son la fuerza de la respuesta derecha e izquierda en el siguiente instante en tiempo; $V_{I(t)}$ y $V_{D(t)}$ representan la fuerza de la respuesta derecha e izquierda actual; a es el tiempo constante; X_I y X_D son las asíntotas de las respuestas con un valor de 1 cuando la respuesta es reforzada y de 0 cuando no lo es, y s es la proporción predicha de elecciones a la tecla derecha.

Nótese que en este modelo sólo se actualiza la ecuación que corresponde a la respuesta ejecutada, conservándose la otra igual. Esto es, no hay interacción entre reforzadores.

Intentar entender el fenómeno conductual a través de la teoría implica la exploración simultánea de dos campos o dominios: El dominio de los resultados experimentales y el dominio de los posibles sistemas teóricos. Una teoría afortunada representa la identificación correcta de

Aprendizaje y Desarrollo de Preferencias

regiones isomórficas en estos dos dominios : un sistema teórico que estrictamente replique un grupo de resultados experimentales. Esta búsqueda de dos dominios infinitos es imposible sin una clasificación. Los resultados experimentales son comunmente clasificados por procedimiento -condicionamiento clásico versus condicionamiento instrumental, procedimientos de elección versus procedimientos de no-elección-, y así sucesivamente. Hay menos consenso sobre la propia clasificación de los sistemas teóricos. Una manera de clasificar los modelos dinámicos de elección es como sigue (ver, Davis *et al*; 1993.)

Considerando solo modelos computables determinísticos, que son modelos en los cuales un estado en el instante siguiente es una función bien definida del estado actual y del insumo (input, De acuerdo a Davis *et al* (1993) y Stenberg (1963)). Cualquier modelo computable para elección recurrente (periódica) puede ser reducido a la siguiente forma discreta:

$$\mathbf{Y}(t+1) = f[\mathbf{Y}(t), \mathbf{R}(t)], \quad (3)$$

Donde $\mathbf{Y}(t)$ es un vector que representa el estado del modelo y $\mathbf{R}(t)$ es un vector que representa las condiciones de reforzamiento en tiempo (iteración) t . En otras palabras, un modelo computable de elección indica el estado del modelo en la siguiente iteración como una función del estado y las condiciones de reforzamiento en la iteración anterior.

La anterior ecuación es “el modelo por definición”. Una segunda función, la regla de respuesta, mapea el estado del modelo sobre la conducta:

$$\mathbf{B}(t) = g [\mathbf{Y} (t)] \quad (4)$$

Donde \mathbf{B} es alguna propiedad conductual medible tal como la proporción de elección u ocurrencia de respuesta. Una distinción importante entre los modelos de ejecución o de estado es si el vector \mathbf{Y} puede ser eliminado de estas dos ecuaciones y ser remplazado por una función de \mathbf{B} sola. Si \mathbf{Y} puede ser eliminada, el modelo es un modelo de ejecución. La posibilidad más simple es que la función g en la ecuación anterior es simplemente un mapeo de uno a uno, ya que el estado del modelo es unicamente definido por algunas propiedades medibles de la conducta. Por ejemplo, \mathbf{Y} podría ser definido como la tasa de respuesta o la proporción de elección medida

Clasificación y Desarrollo de Preferencias.

ada. La ecuación también es obviamente redundante para los casos en los que $Y = 0$. La ecuación 3 puede ser rescrita en términos de las cantidades

que puede ser un mapeo de -varios a uno-; estos es, un cierto modelo de conducta, sin embargo, dicha conducta puede ser compatible con más modelos si y sólo si al menos que las variables puedan ser redefinidas de tal modo que los nuevos modelos son modelos de estado (ver Davis *et al.* 1993) para los cuales bajo las cuales esta redefinición es posible). Por ejemplo, si una respuesta ocurre siempre que alguna variable exceda un cierto nivel, entonces a causa de Y , el estado del modelo, no puede ser únicamente el resultado de una respuesta sobre una iteración particular.

Los modelos corresponden a lo que ha sido denominado como la suposición de "camino" (path), esto es, la idea de que la conducta futura depende sólo del estado actual (input), es decir que la probabilidad de cualquier respuesta en el intento n depende de su probabilidad en el intento anterior y el hecho de que se haya ocurrido o no. La historia anterior al último ensayo está incluida por completo en las condiciones de las puestas a ese ensayo. Si por ejemplo dos animales tienen la misma respuesta, no importa que sus historias detalladas sean diferentes. Las longitudes de sus respectivas secuencias son irrelevantes.

Interpretar el significado de esta hipótesis es decir que un suceso no tiene efecto de un suceso en un ensayo se manifiesta inmediata y completamente en el ensayo siguiente. Lo que esta clasificación agrega es una distinción entre los modelos de estado (las cuales siempre muestran independencia del camino en el espacio de los estados) y los modelos de camino. Entre el modelo de estado y las cantidades observables, lo cual depende de la interpretación (ver el apéndice en Davis *et al.* 1993).

Como se ha mencionado anteriormente casi en forma exclusiva estos modelos sólo han sido utilizados para explicar, los resultados obtenidos en programas concurrentes, donde el sujeto responde ante dos o más opciones de respuesta en forma simultánea y puede

Aprendizaje y Desarrollo de Preferencias

moverse libremente de una a la otra. Las pruebas empíricas han consistido principalmente de estudios de transición, en los cuales se analiza el curso que sigue la conducta al momento de cambiar los parámetros de reforzamiento de un programa concurrente.

Dos han sido las principales transiciones que se han estudiado **i)** alternaciones en los programas de reforzamiento y **ii)** cambios a extinción. El procedimiento experimental típico consiste en presentarle al animal dos opciones de respuesta, solo una de las cuales es reforzada en un cierto día, de acuerdo a un programa de reforzamiento probabilístico (programa de razón). En cada sesión diaria se asigna probabilísticamente cual de las dos respuestas será reforzada. Los modelos de operadores lineales (ecuación 3) describen satisfactoriamente estos experimentos. Sin embargo, recientemente se ha demostrado que fracasan cuando el procedimiento se modifica para evaluar efectos históricos y de acarreo (Davis *et al*, 1993). Si en lugar de alternar diariamente cuál de las dos respuestas es reforzada, se les refuerza en dos bloques de varias sesiones y luego se extinguen, es frecuente encontrar una regresión a preferir la respuesta reforzada en el primer bloque. De igual forma, estos modelos no pueden dar cuenta de la mejora en ejecución en procedimientos de inversión diaria sucesiva. Para dar cuenta de estos y otros resultados relacionados (Davis *et al*, 1993) y (Mark y Gallistel, 1994), han propuesto un modelos alternativos que toma en cuenta la historia lejana de reforzamiento uno de ellos llamado el modelo de efectos acumulativos, cuyas suposiciones básicas establecen que la elección ocurre de una forma de todo o nada (winner-take-all) para la respuesta alterna con la probabilidad alta de reforzamiento y que la probabilidad de reforzamiento de cada respuesta alterna es calculada de la historia completa de entrenamiento (número total de respuestas reforzadas / número total de respuestas reforzadas y no reforzadas). Siendo más específicos. El modelo es computado respuesta por respuesta, y la actividad con el valor más alto de V es la única que ocurre, como en la ecuación 2. Los valores de V son simplemente probabilidades de reforzamiento (frecuencias relativas) en las cuales el numerador es el reforzamiento y el denominador la respuesta, los cuales son acumulados desde el inicio de l experimento; en términos formales:

$$V_i(t+1) = [R_i(t) + R_i(0)] / [N_i(t) + N_i(0)] \quad (5)$$
$$N_i(0) \geq R_i(0) > 0,$$

Aprendizaje y Desarrollo de Preferencias

Donde $R_i(t)$ es el número total de reforzamiento por respuesta i desde el inicio del experimento a través de la iteración t , y $N_i(t)$ es el número total de veces que la respuesta i , ha ocurrido. $R_i(0)$ y $N_i(0)$ son constantes que representan las condiciones iniciales -número de respuestas y reforzamientos que representan el efecto de la experiencia pasada del animal. Dado que siempre estamos tratando con situaciones simétricas de dos elecciones. El modelo asume que $R_L(0) = R_D(0) = R_0$ y $N_L(0) = N_D(0) = N_0$, de tal manera que todas las predicciones están basadas en solo dos parámetros o valores libres para las condiciones iniciales. Los estados -parámetros para R y N - en este modelo refleja la historia entera del modelo.

Mazur y sus colegas han usado un procedimiento en el cual es posible obtener varias curvas de adquisición de cada sujeto (Bailey y Mazur, 1990; Mazur, 1992; Mazur y Ratti, 1991). Cada condición comenzaba con un periodo con iguales contingencias de reforzamiento para dos teclas de respuesta y entonces el desarrollo de la preferencia es observado, cuando sorpresivamente una tecla empieza a dar más reforzadores que la otra. Estudios previos con este procedimiento examinaban como las tasas de transición eran afectadas por un número de diferentes factores, tales como la probabilidad promedio de reforzamiento, el tamaño de un cambio en las probabilidades y la tasa promedio de reforzamiento.

En el estudio reportado por Bailey y Mazur (1990) en un primer experimento la tasa de adquisición de una preferencia en 8 palomas se observó durante 10 periodos de transición en un procedimiento de ensayos discretos de dos elecciones (en este tipo de procedimiento, al sujeto solo se le permite hacer una sola respuesta por ensayo, la cual puede o no ser reforzada). Cada condición comenzaba con la misma probabilidad de reforzamiento en las 2 teclas y luego se le cambiaba a probabilidades diferentes, de tal manera que una tecla tenía una alta probabilidad de reforzamiento (la tecla derecha en la mitad de las condiciones y la tecla izquierda en la otra mitad), y se observó el desarrollo de la preferencia por esta tecla en particular. La tasa de adquisición de la preferencia para la tecla que señalaba mayor densidad de reforzamiento fue más rápida cuando la razón de ambas probabilidades de reforzamiento fue más alta. Además observaron que los animales fueron más sensibles a las razones (δ/δ) y no a las diferencias ($\delta-\delta$) en las probabilidades de reforzamiento. En un análisis de las respuestas individuales y los reforzadores ensayo por ensayo encontraron dos efectos sobre la elección. El efecto a corto plazo era un incremento en la probabilidad de regresar a la misma tecla en el primer o segundo ensayos

Aprendizaje y Desarrollo de Preferencias

seguidos de un reforzador, el efecto de largo plazo era un incremento gradual en la producción de respuestas en la tecla con alta probabilidad de reforzamiento, un incremento que continuaba generalmente después de varios cientos de ensayos.

En un segundo experimento con operante libre (Mazur, 1992) cada condición comenzaba con 2 programas IV iguales, pero luego una de las teclas entregó 60%, 75%, o 90% de los reforzadores. La tasa de aproximación a la asíntota fué aproximadamente la misma en los 3 porcentajes de reforzamiento. Los datos son bien descritos por un modelo que establece que la fuerza de cada respuesta es independientemente incrementada por reforzamiento y decrementada por no reforzamiento (Una adaptación del Modelo de Operador Lineal).

En síntesis los resultados de los experimentos de Mazur y sus colaboradores contradicen las predicciones de varios modelos de conducta de elección transicional, incluyendo el modelo cinético de Myerson (Myerson y Miezin, 1980; ver también Myerson y Hale, 1988), el modelo de razón de invarianza de Staddon (Staddon, 1988; Staddon y Horner, 1989), y el modelo de operador lineal (Bush y Mosteller, 1955). Dichos modelos no pueden predecir cuales condiciones en transición tendrían tasas de adquisición rápidas y cuales tendrían tasas de adquisición lentas. Como ejemplo consideremos dos condiciones del experimento de Mazur y Ratti (1991) con programas RR (razón probabilística). En una condición, las probabilidades de reforzamiento para dos opciones de respuesta fueron .16 y .10 y en otra condición la probabilidades de reforzamiento fueron .07 y .01. ¿Qué tan rápidamente los sujetos desarrollaban una preferencia por la probabilidad alta de reforzamiento, en estas dos condiciones? El modelo Cinético de Myerson predecía tasas iguales de adquisición en ambas condiciones, ya que las probabilidades de reforzamiento diferían por la misma cantidad (.06) en ambas condiciones. El modelo de Razón de Invarianza de Staddon y el modelo de Operador Lineal Simple predecían una adquisición de preferencia levemente más rápida en la condición .16-.10. Sin embargo los resultados actuales contradicen todas estas predicciones, con adquisición sustancialmente más rápida en la condición .07-.01. En general Mazur y Ratti (1991) encontraron adquisición de preferencia más rápida cuando la razón de las dos alternativas de las probabilidades de reforzamiento eran más grandes. No obstante dichos resultados pueden ser descritos bastante bien por un simple modelo matemático similar a uno propuesto por Couvillon y Bitterman (1985). El modelo establece que cada opción de respuesta, i , tiene un fuerza separada V_i . Incrementos en la fuerza de respuesta

Aprendizaje y Desarrollo de Preferencias

cada vez que la respuesta es reforzada y decremos cada vez que la respuesta no es reforzada. Cada vez que una respuesta es reforzada, V_i incrementa como sigue:

$$\Delta V_i = r (1 - V_i) \quad (6)$$

Donde ΔV_i es el cambio en la fuerza V_i , y r es un parámetro de reforzamiento que puede ir de 0 a 1. Después de cada respuesta que no es *reforzada*, la fuerza de respuesta decremos como sigue:

$$\Delta V_i = n (-V_i) \quad (7)$$

Donde n es un parámetro de no reforzamiento que puede ir de 0 a 1. Una simple regla de igualación es usada para trasladar desde las fuerzas independientes de dos respuestas (V_1 y V_2) para la probabilidad de que una pueda ocurrir:

$$P_i = V_i / (V_1 + V_2) \quad (8)$$

Aunque los resultados de varios experimentos son bien descritos por este modelo, es demasiado sencillo para acomodar algunas de las complejidades de conducta observada en situaciones de elección simple. Por ejemplo Mazur (1995) encontró un tipo de recuperación espontánea que no era predicho por este modelo (o por muchos otros modelos de conducta de elección transicional, incluyendo todos aquellos mencionados anteriormente) dichos resultados fueron encontrados al utilizar tres diferentes proporciones de reforzamiento (.10-.90, .25-.75 y .40-.60) para las sesiones de transición y para cuatro sesiones post-transición, al inicio de cada sesión post-transición la proporción de respuesta para la opción rica fue menor que al final de la sesión precedente (última en transición). Esta recuperación espontánea fue evidente para el total de las tres proporciones de reforzamiento utilizadas y era mayor en la primera sesión de post-transición y progresivamente fue tornándose menor en las sesiones subsecuentes.

En resumen los experimentos de Mazur y colaboradores (Bailey y Mazur, 1990; Mazur, 1992; Mazur, 1995; Mazur, 1996 y Mazur y Ratti, 1991), ilustran varias limitaciones en las predicciones de varios modelos tales como el modelo cinético (Myerson y Hale, 1988; Myerson y

Aprendizaje y Desarrollo de Preferencias

Miezin, 1980), el modelo de operador lineal simple (Bush y Mosteller, 1955), el modelo de razón-invarianza (Staddon, 1988; Staddon y Horner, 1989) y la versión original de la teoría de mejoramiento (Herrnstein y Vaughan, 1980; Vaughan y Herrnstein, 1987). Además de que estos modelos sólo han sido utilizados para explicar casi en forma exclusiva los resultados obtenidos en programas concurrentes, (ver Anexo 1). Dada la simultaneidad de las opciones de respuesta en los programas concurrentes, el tiempo dedicado a responder en una opción hace que se reste tiempo dedicado a responder a la otra opción. La competencia entre respuestas por el tiempo disponible, permite que el incremento en la fuerza de una respuesta, de lugar a decrementos en las tasas de otras respuestas. Para estudiar la posible interacción entre reforzadores, es necesario separar las respuestas de manera que los efectos de la interacción no ocurran por la competencia del tiempo común. La forma más simple de lograr lo anterior, tal y como se emplea en esta investigación, es utilizando

Programas Múltiples (Elección Intertemporal).

En los entornos naturales, el acceso de bienes necesarios para sobrevivir (la tasa de reforzamiento o ganancias) varían, frecuentemente en forma señalada, a lo largo del tiempo. Los programas de reforzamiento múltiples son la mejor manera de modelar experimentalmente estos entornos. Algunos experimentos han mostrado que la tasa de respuesta en un componente con un programa de intervalo variable (IV) constante, depende inversamente de la tasa de reforzamiento en el componente alterno, fenómeno conocido como contraste conductual (ver para una revisión general de contraste conductual, Williams, 1988; Williams, 1983) y se han reportado dos tipos: **contraste global positivo y contraste global negativo**. En el primer caso la tasa de respuesta en un componente incrementa como función de un decremento en la tasa de reforzamiento en el componente alterno, relativo a una condición previa de igualdad entre tasas de reforzamiento entre los componentes. El componente donde se observa el incremento de la respuesta mantiene, sin embargo, su tasa de reforzamiento fija. En el segundo caso la tasa de respuesta en un componente decreta como función de un incremento en la tasa de reforzamiento en el componente alterno, relativo a una condición previa de igualdad de tasa de reforzamiento entre los componentes. Nuevamente, el componente donde se observa el decremento de la respuesta mantiene su tasa de reforzamiento fija.

Aprendizaje y Desarrollo de Preferencias

Además que en términos del problema de aprendizaje que enfrenta un organismo, las situaciones temporalmente sucesivas son más similares a las estudiadas bajo el nombre de condicionamiento clásico que a los programas simultáneos.

En la actualidad aún no se cuenta ni con la investigación empírica pertinente al problema de la adquisición en programas múltiples ni con un conjunto de modelos explícitos acerca de los algoritmos relevantes. Por tanto, el siguiente estudio tiene como objetivo el generar la base de datos necesaria para el desarrollo y evaluación de las distintas propuestas teóricas acerca de la representación de sucesos intertemporales.

Al igual que en el caso de los estudios de elección simultánea, el núcleo del experimento consistió de estudios de transición en las condiciones de reforzamiento de las alternativas. Dicho experimento tuvo como objetivo fundamental el evaluar el curso de los ajustes en el comportamiento ante cambios abruptos, impredecibles y no señalados cuando las tasas de reforzamiento ocurren en diferentes períodos de tiempo, esto es, en programas múltiples.

El procedimiento general de este experimento fue similar a los utilizados por Mazur y sus colegas (Bailey y Mazur, 1990; Mazur, 1992; y Mazur y Ratti, 1991.), aplicado a programas múltiples. Cada condición consistió de:

Tres días de *entrenamiento* en los cuales las respuestas en los dos componentes fueron reforzadas con la misma probabilidad de reforzamiento (RV), *-fase de una sola probabilidad-* la cual estaba constituida por un programa simple RV que asignaba reforzadores a los dos componentes de los programas múltiples, seguida de dos sesiones de *transición*, o *fase de dos probabilidades*, donde un par de probabilidades (RVx, RVy), operaban. Cada condición se repitió cuatro veces y los sujetos pasaron por las condiciones en distinto orden. Los valores para las razones utilizadas fueron creados a partir de las progresiones sugeridas por Freshler y Hoffman (1962). Los componentes duraban 60 seg. cada uno y las sesiones finalizaban a los 30 minutos. En esta primera parte del presente experimento, usamos en *las fases de dos probabilidades* pares de probabilidades de reforzamiento que tenían igual razón (5 a 1), pero diferentes valores absolutos (.16, .08 y .04). En la segunda, se usaron pares de probabilidad de reforzamiento de una razón de 2 a 1 y 3 a 1 con iguales valores absolutos (.04), lo anterior para comparar el efecto de una razón de probabilidades diferentes, y el efecto de diferencias constantes en probabilidad. (Las

Aprendizaje y Desarrollo de Preferencias

cinco condiciones con valores de los RVs y con las probabilidades de reforzamiento en *entrenamiento y transición* se indican en la Tabla 1).

M E T O D O

Sujetos

Se utilizaron ocho palomas domesticas adultas Columba Livia privadas al 80% de su peso corporal, experimentalmente ingenuas. Para mantenerlos en el peso. Después de cada sesión, se les dio alimento adicional para que al día siguiente tuvieran el peso del criterio. A lo largo del estudio las palomas tendían acceso libre al agua.

Aparatos

Ocho cámaras idénticas de condicionamiento operante para pichones diseñadas por el personal del Laboratorio de Comportamiento y Adaptación, con las siguientes dimensiones interiores, 37 cm. de alto, 30 cm. de ancho y 35cm. de fondo, cada una contaba con los siguientes componentes: **i)** una luz general que se localizaba en el centro del techo de la caja, **ii)** tres teclas de respuestas translúcidas con su correspondiente proyector de estímulos, las cuales se encontraban en la pared derecha de la cámara y estaban separadas por 7 cm; cada tecla media 2.7 cm. de diámetro, las teclas izquierda y central no fueron utilizadas en este experimento, la tecla derecha podía ser iluminada por una luz roja o verde producidas por dos focos de 2 Watts fijados detrás de la tecla, Una fuerza de aproximadamente 0.15 N era requerida para operar cada tecla. **iii)** un dispensador electromagnético de alimento (grano balanceado) que se localizaban en la pared derecha de la cajas a 5.5 cm. del piso y a 10 cm. de la tecla central. El reforzador fue dos segundos de acceso libre a una mezcla de granos en el dispensador. Durante ese tiempo se encendió una luz que iluminaba el dispensador (comedero) con una luz blanca proveniente de un foco de 2 watts fijado sobre este; y se apagaban tanto la luz general como la de las teclas, **iv)** las cámaras experimentales estaban ventiladas todo el tiempo que duraba la sesión cada una con un extractor que sacaba el aire del interior de la cámara, y que a su vez servía como generador de ruido blanco que atenuaba el sonido proveniente del exterior. El control de la secuencia de

Aprendizaje y Desarrollo de Preferencias

presentación de los estímulos dentro de las cámaras experimentales, así como el registro de las respuestas de cada animal, se realizó empleando una computadora PC por medio de una interface y lenguaje de programación (MEDSTATE). Todos los eventos de una sesión se registraron con una exactitud de mili segundos. Los datos fueron almacenados en discos flexibles para su posterior tratamiento.

Procedimiento

Los sujetos fueron alojados en cajas habitación individuales teniendo los primeros veinte días acceso libre tanto al agua como al alimento. Durante este período, se registró diariamente su peso, posteriormente se tomó la media de este registró como punto de comparación posterior (peso *ad lib.*).

Tabla 1. Probabilidades de Refto. y los Valores de los RVs., para cada uno de los componentes en cada una de las fases y condiciones experimentales.

Condiciones del Experimento (Valores de los Rvs, Probabilidades de Reforzamiento. Razones y Diferencias utilizados)								
Condición	Fase de 1 Probabilidad		Fase de 2 Probabilidades				Razón	Dif.
	Probabilidad en Entrenamiento		Probabilidad Alta		Probabilidad Baja			
A	0.12	RV 8 - RV 8	0.2	RV 5	0.04	RV 25	5a1	0.16
B	0.06	RV 17 - RV 17	0.1	RV 10	0.02	RV 50	5a1	0.08
C	0.03	RV 33 - RV 33	0.05	RV 20	0.01	RV 100	5a1	0.04
D	0.06	RV 17 - RV 17	0.08	RV 13	0.04	RV 25	2a1	0.04
E	0.04	RV 25 - RV 25	0.06	RV 17	0.02	RV 50	3a1	0.04

Después de este período, se restringió paulatinamente la cantidad de alimento disponible hasta que se alcanzó el criterio de 80% (+-5gm) de peso con respecto al valor *ad lib.* Una vez mantenidos los sujetos diez días al 80% de su peso normal, se pasó a la fase de entrenamiento del picoteo mediante un programa de automoldeamiento del picoteo a la tecla, hasta conseguir de manera consistente la respuesta de picoteo en las tres teclas, ante los tres colores (amarillo, verde y rojo). Posteriormente se les presentaron programas de razón fija para una tecla lateral (derecha) con tasas de reforzamiento altas (RF 3) que se iban disminuyendo paulatinamente hasta llegar al valor más bajo de reforzamiento que se usó en una de las condiciones experimentales (RV 100).

Aprendizaje y Desarrollo de Preferencias

Cada condición comenzaba con sesiones de una probabilidad (tres sesiones), en las cuales un programa probabilístico simple asignaba reforzadores a ambos componentes del programa múltiple con igual probabilidad, cada componente. Estas sesiones eran diseñadas para elicitarse ampliamente igual número de respuestas en ambos componentes. Cada condición finalizaba con dos sesiones de dos probabilidades, en la cual dos programas probabilísticos independientes asignaban reforzadores a la respuesta en cada uno de los dos componentes del programa múltiple, y donde la probabilidad de reforzamiento era más alta para un componente que para el otro. Cada sesión finalizaba a los treinta minutos. En el comienzo de cada sesión, la luz general comenzaba y permanecía encendida a lo largo de la dicha sesión, excepto durante los periodos de reforzamiento.

Los datos fueron analizados tanto globalmente, en términos de tasas de respuesta y tasas relativas de respuesta a lo largo de las sesiones, como en términos de las tasas de respuesta y tasas relativas de respuesta componente a componente.

Aprendizaje y Desarrollo de Preferencias

RESULTADOS

Obtuvimos varias curvas de adquisición para cada una de las condiciones a las que fueron sometidos los sujetos y comparamos dichas curvas para ver si eran distintas, ya sea por las razones utilizadas, o por las diferencias o ambas, además de ver el patrón de ajuste de la distribución de respuestas de los sujetos, a lo largo de las sesiones. Para facilitar la descripción de los siguientes resultados al color de la tecla asociado con la probabilidad de reforzamiento más alta en las fases de transición le llamamos “componente rico”, en cada condición (aún, en las fases de entrenamiento donde este color de la tecla estaba asociado a una probabilidad de reforzamiento mas baja), en consecuencia el color de la tecla asociado a la probabilidad de reforzamiento mas baja le llamamos “componente pobre”. De tal forma la presentación de los siguientes resultados se divide en dos partes: 1) Un análisis del ajuste de las preferencias a un nivel molar, es decir analizamos las tasas absolutas de respuestas, y las tasas relativas de respuestas como una función del día en cuestión para cada una de las condiciones señaladas en el experimento (Ver Tabla 1), y 2) a través de un análisis más molecular, es decir, de componente a componente y de sesión a sesión e igualmente para cada una de las condiciones (curvas de adquisición).

La figura 1 muestra, el promedio de las tasas absolutas de respuesta ($\#$ de Respuesta / Tiempo de exposición al componente a lo largo de la sesión {ordenada}), como una función del día y fase correspondiente para cada una de las condiciones {abcisa}, los tres primeros pares de cajas (llenas y vacías) en cada una de las cinco gráficas, corresponden al promedio de los tres primeros días de cada condición para el total de los sujetos y para el total de sesiones que los sujetos pasaron por dicha fase (4 ciclos), los cuales están en las fases de una probabilidad o entrenamiento (donde un simple programa de reforzamiento variable RVx asignaba reforzamientos con igual probabilidad a ambos componentes del programa múltiple), los siguientes dos pares de cajas (día t1 y día t2) son el promedio para el total de los sujetos de los dos días que pasaban en las fases de dos probabilidades o transición (donde dos programas de reforzamiento RVx y RVy asignaban signaban reforzamientos con diferente probabilidad a cada uno de los componentes del programa múltiple). Las cajas llenas representan el color del componente que estaría asociado en las fases de transición al componente pobre y las cajas vacías representan el color del componente que en las fases de transición estaría asociado al componente rico. Como se puede observar de manera muy general en la mayoría de las cinco condiciones en

Aprendizaje y Desarrollo de Preferencias

los días de entrenamiento, la asignación de las respuestas para cada uno de los componentes del programa múltiple fue muy similar a excepción de la condición C (gráfica de en medio a la izquierda) donde el valor de los RVs fue de 33 y fue el valor más extremo utilizado en las fases de una probabilidad, se puede ver que hubo preferencia por un color en particular. En cuanto a las fases de transición se puede apreciar casi en todas las condiciones una preferencia en la tasa de respuestas por el componente asociado al RV pequeño (rico). Los tres primeros paneles los dos de arriba y el de en medio a la izquierda (Condiciones A, B y C respectivamente) en los cuales las razones en estas fases de transición eran de 5 a 1 para las tres condiciones, donde los valores absolutos o diferencias variaban en cada una de ellas. Se puede apreciar que en las condiciones B y C la preferencia por el color de la tecla asociada al componente rico es más pronunciada además de que las tasas absolutas de respuesta en dicho componente son más elevadas en las dos condiciones que las de la condición A, ya que en las tasas de respuestas en el componente rico en B y C están sobre 100 y en A esta por debajo de 100, además de que en términos generales las tasas absolutas de respuesta tanto en el componente rico como en el componente pobre, en la condición A, en comparación con las otras 4 condiciones, siempre fueron menores.

Ahora bien, comparando las dos últimas condiciones o gráficas, la de en medio a la derecha y la de abajo a la izquierda (Condiciones D y E) con la gráfica de en medio a la izquierda (Condición C) las razones en las fases de transición variaban, pero donde las diferencias permanecían constantes, se puede apreciar que cuando la razón es mayor 5 a 1, (Condición C) -D y E tenían razones de 2 a 1 y de 3 a 1 respectivamente- la preferencia por la tecla rica es claramente pronunciada y donde la razón es menor o sea la condición D dicha preferencia es menos acentuada.

Aprendizaje y Desarrollo de Preferencias

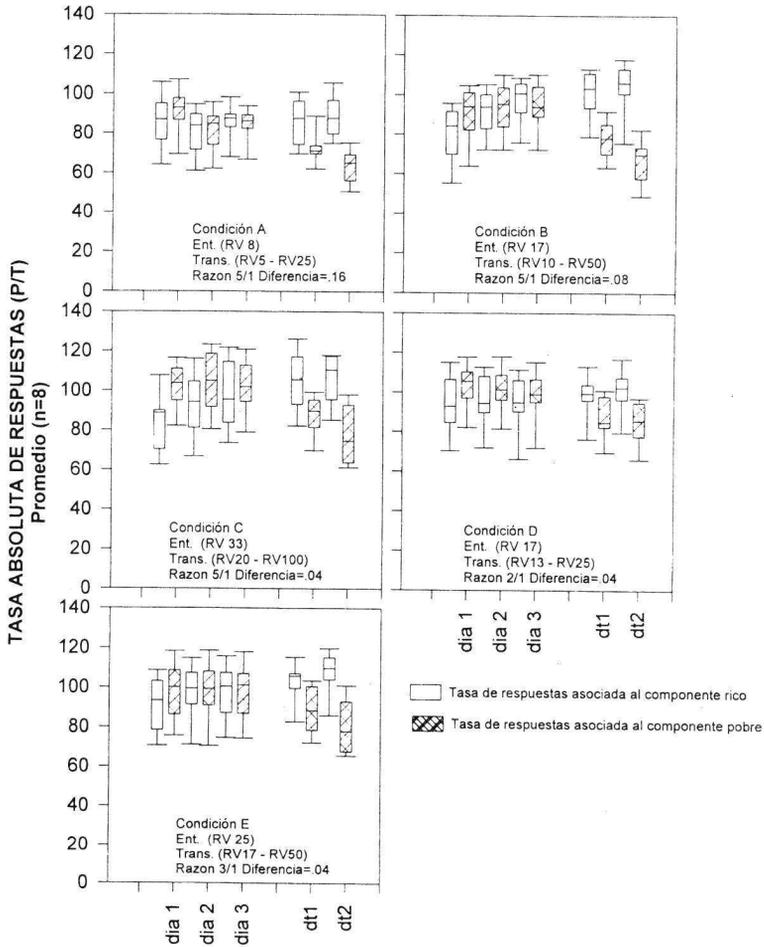


Fig. 1. Muestra los promedios (línea dentro de las cajas), las desviaciones estándar (extremos de las cajas) y los errores estándar (líneas fuera de las cajas) de las tasas absolutas de respuesta para el total de los sujetos en todas las condiciones y en los cuatro ciclos, a través de las sesiones o días de que consistió cada condición experimental.

Aprendizaje y Desarrollo de Preferencias

La figura 2 muestra los promedios del total de los sujetos de las tasas relativas de respuesta para cada una de las condiciones, se presentan las tasa relativas para los promedios de los dos días por separado, tanto el efecto de condición como en la interacción entre condición y día, los resultados fueron estadísticamente significativos al aplicarse un MANOVA a través de todas las condiciones $F(4,12) = 16.86; p < .0001$ y $F(4,12) = 14.78; p < .0001$ para la interacción condición día (ver Tabla 2, para los resultados generales del MANOVA), se puede apreciar que cuando la razón fue de 5 a 1 las diferencias en el factor día fueron más significativas.

En la figura 3 se muestran el promedio general de las tasas relativas de respuestas para el total de los sujetos, en los 4 ciclos y para cada una de las condiciones del experimento.

Para cada gráfica de la figura 3, en el eje de las abscisas se muestran intervalos de 2 componentes o de 2 minutos, en los cuales los sujetos respondían a uno y a otro componente (rico y pobre) del programa múltiple. Y en tanto en el eje de las ordenadas se muestra la tasa relativa de respuestas, es decir la tasa de respuestas en el componente rico (P_r) dividida por la tasa de repuestas en el componente rico más la tasa de respuestas en el componente pobre ($P_r + P_p$) dicha medida fue obtenida para cada uno de los 15 intervalos de dos componentes que constituían cada sesión. Los círculos cerrados en cada una de las gráficas representan el promedio total de los sujetos y en los cuatro ciclos para los tres días de entrenamiento o fase de una probabilidad, de la tasa relativa de respuestas, los triángulos abiertos representan el promedio total de todos los sujetos y en los 4 ciclos para el 1^{er} día de transición o fase de dos probabilidades, de la tasa relativa de respuestas, y los rombos abiertos representan el promedio total de todos los sujetos y en los 4 ciclos para el 2^o día de transición o fase de dos probabilidades, de la tasa relativa de respuestas. El orden de presentación para cada una de las gráficas es el mismo que se adoptó para la figura 1.

Aprendizaje y Desarrollo de Preferencias

Tabla 2. Se presentan los valores obtenidos al aplicar el MANOVA general con tres factores, Condición, Día e Intervalo con 5, 2 y 15 niveles respectivamente.

MANOVA GENERAL		Tabla de todos los efectos e interacciones 1 = Condición, 2 = Días, 3 = Intervalos				
Efectos e Interacciones	gl Efecto	SC Efecto	gl Error	SC Error	Valor de F	Nivel de p
1	4	0.0826	12	0.0049	16.855	.0001*
2	1	0.2691	3	0.0138	19.447	.0216*
3	14	0.0714	42	0.0027	26.432	.0000*
12	4	0.0404	12	0.0027	14.785	.0001*
13	56	0.0058	168	0.0016	3.521	.0000*
23	14	0.0047	42	0.0013	3.66	.0006*
123	56	0.0057	168	0.0016	3.686	.0000*

De una manera muy general se puede apreciar en cada una de las condiciones, que la tasa relativa de respuestas en el 2º día de transición es más extrema y más alejada de la tasa relativa de respuestas en los días de entrenamiento (círculos cerrados), que la tasa relativa de respuestas del 1º día. El MANOVA revela el efecto significativo de intervalo [$F(14, 42)=26.43; p<.00$] así como un efecto levemente significativo de día a través de los intervalos [$F(14,42)=3.66, p<.00$]. En las fases de entrenamiento o de una probabilidad es muy claro que los sujetos responden de igual manera en ambos componentes (valores cercanos a .5) casi en todas las condiciones, si acaso a excepción de la condición C, gráfica de enmedio a la izquierda en donde los valores están al principio de la sesión por debajo de .5, pero conforme avanza la sesión estos valores son muy cercanos a .5. En relación con la diferencia de las tasas relativas de las fases de transición y las tasas relativas en las fases de entrenamiento, ésta como se puede apreciar es mayor en la Condición C donde los valores utilizados son mas extremos (RV 33 en entrenamiento y RV20-RV100 en transición, razón de 5 a 1 y diferencia de .04) y es menor en la condición D gráfica de enmedio a la derecha, donde los valores utilizados tanto en razón como en diferencia fueron los menores (razón de 2 al 1 y diferencia de .04). También se puede apreciar en el último intervalo por lo menos en cuatro condiciones (A,B,D y E), un incremento en la tasa relativa tanto en transición como en entrenamiento. Tal incremento podría ser atribuido a un efecto de saciedad al final de las sesiones, efecto que no es tan pronunciado en la condición C donde los valores a los que fueron sometidos los sujetos, como anteriormente mencionamos, fueron los más extremos o

Aprendizaje y Desarrollo de Preferencias

pobres de todo el experimento, sin embargo este efecto en el último intervalo no influyó significativamente en los análisis estadísticos realizados, ya que éstos se realizaron con o sin éste intervalo y los resultados fueron similares.

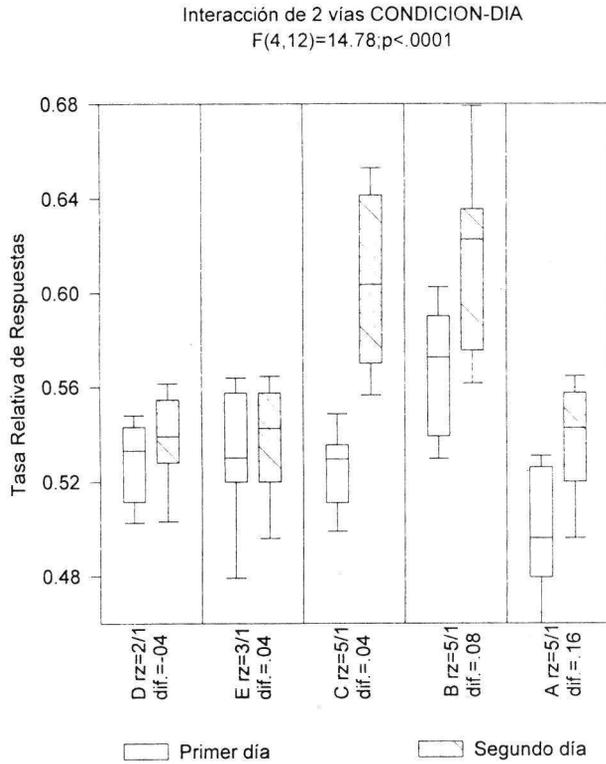


Fig. 2. Se presenta la interacción entre condición y día, para todas las condiciones, las cajas vacías representan el promedio, desviación estándar y el error estándar de las tasas relativas de repuesta para el primer día de transición y las cajas llenas representan los mismos parámetros pero para el segundo día de transición.

Aprendizaje y Desarrollo de Preferencias

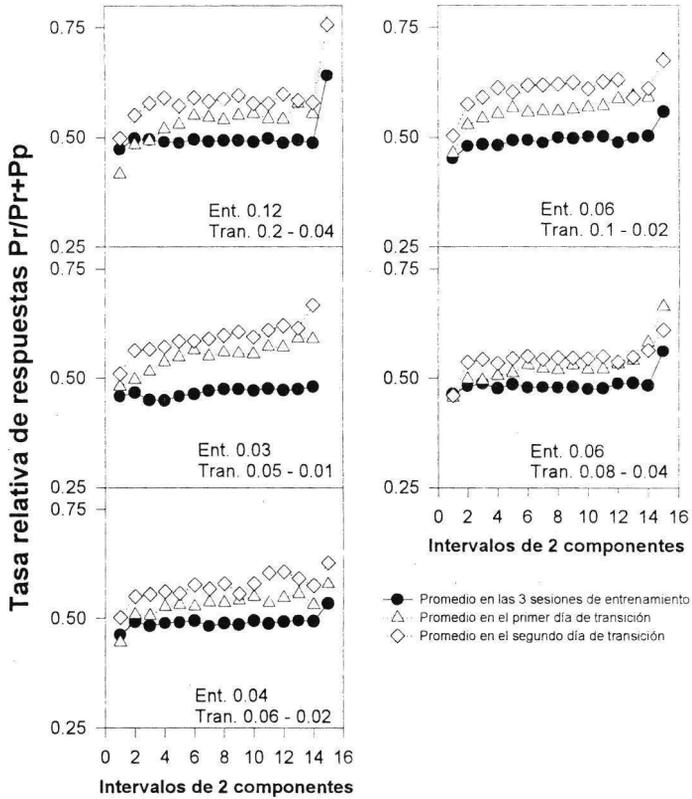


Fig. 3. Se presentan los promedios de las tasas relativas de respuesta como a función de los intervalos de 2 componentes o dos minutos, para el total de los sujetos en todas las condiciones y para los cuatro ciclos del experimento. Los círculos representan las tasas relativas en los tres días de entrenamiento. Los triángulos representan las tasas relativas en el 1^{er} día de transición. Los rombos representan las tasas relativas en el 2^o día de transición.

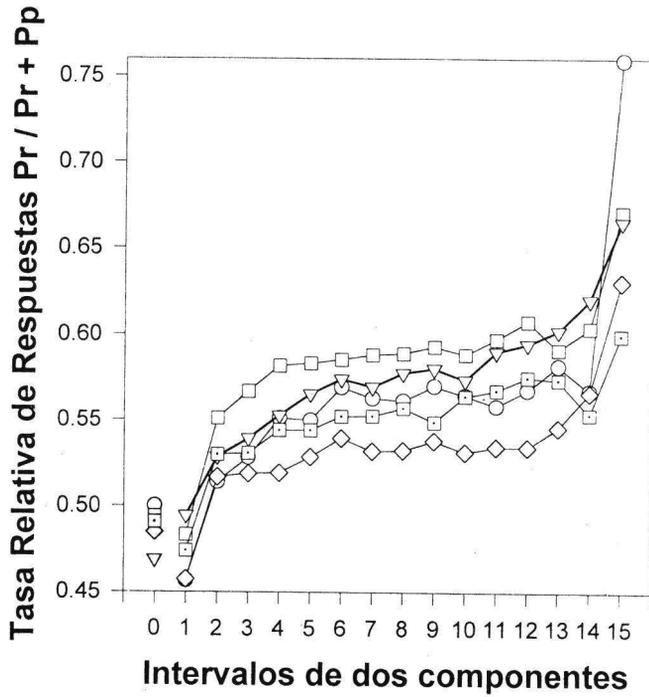
Aprendizaje y Desarrollo de Preferencias

En la figura 4 se muestran los promedios para el total de los sujetos en los dos días de transición en todas las condiciones para los cuatro ciclos del experimento de la tasa relativa de respuestas ($P_r / P_r + P_p$). En el eje de las abscisas se muestra cada uno de los 15 intervalos de dos componentes o de dos minutos de que consta cada sesión y en el eje de las ordenadas se grafica el promedio de la tasa relativa de respuestas para el total de los sujetos, los símbolos sombreados ubicados en el intervalo 0, representan el promedio total de las sesiones de entrenamiento para cada una de las condiciones y los símbolos abiertos representan el promedio total de la tasa relativa de respuestas para cada uno de los intervalos de la sesión en las fases de transición para cada una de las condiciones. Se puede apreciar que los promedios que vienen de entrenamiento -símbolos sombreados- caen entre valores de (.45 y .50), es decir los para cada uno de los intervalos de la sesión en las fases de transición para cada una de las condiciones (círculos-Condición A, triángulos-Condición B, cuadrados-Condición C, rombos-Condición D y hexágonos-Condición E.). Se puede apreciar que los promedios que vienen de entrenamiento (símbolos sombreados) caen entre valores de (.45 y .50) es decir los sujetos respondían igual en ambos componentes del programa múltiple de una probabilidad, mientras que los valores de los intervalos en transición (intervalos del 1 al 15, figuras abiertas), en el primer intervalo de la sesión empiezan respondiendo de manera muy similar a las sesiones de entrenamiento, sin embargo ya para el segundo componente se ve un claro incremento en la tasa relativa de respuestas, que se mantiene hasta el final de la sesión, es decir el animal muestra una preferencia por la tecla rica o por el valor del RV menor en las fases de transición para el total de las condiciones. La tasa relativa mayor o extrema se da en la condición C (cuadrados) y la menor se da en la Condición D (rombos), igual que en la figura 2, se puede apreciar también de manera muy general que las tasas relativas aumentan conforme aumenta la razón y disminuye la diferencia, es decir la tasa relativa mayor como mencionamos es la de la Condición C la cual tiene razón de 5 a 1 y diferencia de .04, la que le sigue es la condición B (triángulos) la cual tiene una razón igual de 5 a 1 pero una diferencia de .08, la Condición A (círculos) tiene la misma razón y una diferencia de .16, las dos tasas relativas menores son las dos últimas condiciones D y E (rombos y hexágonos respectivamente) las cuales tenían razones de 2 a 1 y de 3 a 1 respectivamente y diferencias de .04, se puede apreciar además de que las condiciones más parecidas en cuanto a valores utilizados A y D (RV8-RV8 y RV17-RV17 en entrenamiento, y RV5-RV25 y RV13-RV25 en transición, respectivamente) de hecho fueron muy parecidas en las tasas relativas de respuesta obtenidas.

Aprendizaje y Desarrollo de Preferencias

En la figura 5 se presentan los promedios totales de las tasas absolutas de respuestas de los tres días de entrenamiento en el total de los sujetos para los cuatro ciclos del experimento (panel superior) como una función del número de intervalo de 2 componentes o dos minutos (abscisas). Los promedios totales de las tasas absolutas para el promedio de los dos días en transición (panel inferior) igualmente como una función del número de intervalos de 2 componentes o dos minutos. En la figura 6 se presentan los promedios totales de las tasas absolutas de respuesta en el 1^{er} día (panel superior) y 2^o día (panel inferior) de transición en el total de los sujetos para los cuatro ciclos del experimento como una función del número de intervalos de 2 componentes o dos minutos. Los símbolos cerrados representan el promedio total de la tasa absoluta de respuestas en el color de la tecla que en las fases de transición estuvieron asociadas al componente pobre y los símbolos abiertos representan el promedio total de la tasa absoluta de respuestas en el color de la tecla que en las fases de transición estuvieron asociadas al componente rico. En el panel superior se puede observar que no hay diferencias significativas en cuanto a la asignación de respuestas para cada uno de los componentes, en cada una de las condiciones del experimento. Todas las tasas de respuesta caen en un rango de entre 80 y 110 respuestas por minuto, se puede observar al principio una pequeña diferencia en el primer intervalo pero después de dicho intervalo el rango se hace aun más pequeño (85 a 105 aproximadamente). Se puede observar que en el intervalo final en la mayoría de las condiciones, hay una disminución en la tasa de respuesta de ambos componentes del programa múltiple, a excepción de la condición C (cuadrados) donde no se ve una disminución muy acentuada, el caso extremo son las condiciones A y D (círculos y rombos), que al parecer, en el intervalo final, en el color del componente que estaría asociado al componente pobre en las fases de transición las tasas disminuyeron notablemente. El panel inferior de dicha figura es un resumen de los dos paneles de la figura 6 ya que presenta los promedios de los dos días en transición (1^{eros} y 2^{os} días) para el total de los sujetos en todas las condiciones para los 4 ciclos del experimento. Los resultados descritos en esta figura son congruentes con los encontrados y descritos en las dos figuras anteriores (figuras 3 y 4) tanto en las diferencias en las tasas de respuesta ya sea para los componentes ricos y pobres en todas las condiciones, como para las tasas de respuesta obtenidas en el 1^{er} y último intervalo en todas las condiciones.

Aprendizaje y Desarrollo de Preferencias



ENTRENAMIENTO		TRANSICION
○	CONDICION A RV8 - RV8	RV5 - RV25
□	CONDICION B RV17 - RV17	RV10 - RV50
▽	CONDICION C RV33 - RV33	RV20 - RV100
◇	CONDICION D RV17 - RV17	RV25 - RV50
◻	CONDICION E RV25 - RV25	RV17 - RV50

Fig. 4. Se presenta el promedio de la tasa relativa de respuestas en los dos días de transición como una función de los intervalos de 2 componentes. Los símbolos cerrados son los promedios de los días de entrenamiento. Los símbolos abiertos son los promedios de los días de transición.

Aprendizaje y Desarrollo de Preferencias

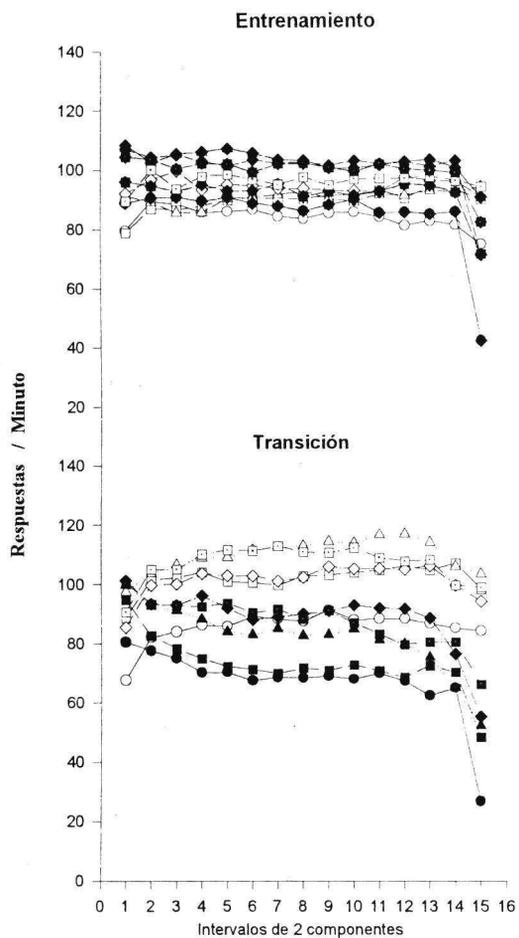


Fig. 5. Muestra el promedio total de las tasas absolutas de respuesta como una función de los intervalos de 2 componentes para los tres días de entrenamiento (panel superior) y para los dos días de transición (panel inferior) para el total de los sujetos en todas las condiciones y para los cuatro ciclos del experimento. Los símbolos abiertos representan las tasas de respuesta en los componentes ricos y los símbolos cerrados representan las tasas de respuesta en los componentes pobres.

Aprendizaje y Desarrollo de Preferencias

La figura 6 en el panel superior se muestran los promedios en el 1^{er} día de transición para los cuatro ciclos de que consistía el experimento. En el primer intervalo los sujetos responden muy parecido en ambos componentes, como si todavía continuaran en las fases de entrenamiento o de una probabilidad y los valores para ambos componentes en transición o fase de dos probabilidades fueran iguales, sin embargo conforme pasan los intervalos, las tasas para cada uno de los componentes van acrecentando su diferencia, se puede ver que las diferencias son más notables en las condiciones B y C (cuadrados y triángulos respectivamente) y las menos se dan en la condición D (rombos), también se puede ver que en todas las condiciones en el último intervalo una significativa disminución en la tasa de respuestas del componente pobre o de baja probabilidad de reforzamiento, al igual que en entrenamiento. El caso extremo se dio en la Condición A donde la tasa de respuestas es de 25 respuestas por minuto aproximadamente en el componente anteriormente mencionado, la tasa de respuestas mayor en el componente rico o de alta probabilidad de reforzamiento fue la obtenida en la Condición C (triángulos abiertos), la cual fue de 115 respuestas por minuto aproximadamente para la mayoría de los intervalos (los valores de los componentes RVrico y RVpobre en dicha condición fueron de 20 y 100 respectivamente), por el contrario la tasa de respuestas menor en dicho componente del total de condiciones fue la obtenida en la condición A (círculos abiertos), 80 respuestas por minuto aproximadamente (los valores de los componentes RVrico - RVpobre en esta condición fueron de 5 y 25 respectivamente). El panel inferior muestra los promedios en el 2^o día de transición en el total de los sujetos en cada una de las condiciones, para los cuatro ciclos de que consistía el experimento.

Podemos apreciar que las diferencias entre las tasas de respuesta para los componentes ricos o de alta probabilidad de reforzamiento y pobres o de baja probabilidad de reforzamiento, en comparación con la figura anterior son significativamente más acentuados, de hecho, la mayoría de las tasas de respuesta en los componentes ricos o de alta probabilidad de reforzamiento (símbolos abiertos) caen en un extremo de la gráfica (parte superior) a excepción de la condición A (círculos), y las tasas de respuesta en los componentes pobres o de baja probabilidad de reforzamiento caen en el otro extremo de la gráfica (parte inferior). Al igual que en la figura anterior en el primer intervalo, las tasas de respuesta son muy parecidas para ambos componentes (pobre y rico) en todas las condiciones, pero después de este intervalo las diferencias en las tasas de respuesta comienza a ser visiblemente distintas para ambos componentes, en todas las condiciones, e insistimos, en comparación con las tasas de respuesta del 1^{er} día las diferencias son

Aprendizaje y Desarrollo de Preferencias

notablemente más acentuadas, podemos observar además que las tasas de respuesta más altas fueron las obtenidas en las condiciones C y E (donde los requisitos de respuesta fueron los más extremos o altos del experimento, RV20-RV100 y RV17-RV50 respectivamente) y por el otro lado las tasas de respuesta menores fueron las obtenidas en las Condiciones A y B (donde los requisitos de respuesta utilizados fueron de RV5-RV25 y RV10-RV50 respectivamente). La diferencia más acentuada en la presente figura, en cuanto a las tasas de respuesta para el componente rico o de alta probabilidad de reforzamiento y para el componente pobre o de baja probabilidad de reforzamiento, fueron las obtenidas en la condición B (componente rico RV10, componente pobre RV50), así la tasa para el componente rico fue en promedio de aproximadamente 110 respuestas por minuto y para el componente pobre fue de 60 respuestas por minuto aproximadamente. Y la diferencia menos acentuada fue la obtenida en la condición D (componente rico RV13, componente pobre RV25) en la cual la tasa para el componente rico fue en promedio de aproximadamente 100 respuestas por minuto y para el componente pobre fue de 90 respuestas por minuto aproximadamente. Se puede apreciar también la disminución de las tasas de respuesta en el componente pobre en el intervalo final, en la mayoría de las condiciones, el caso más extremo al igual que en la figura anterior se dio en la condición A (círculos cerrados) y el menos fue para la condición C (triángulos cerrados).

Aprendizaje y Desarrollo de Preferencias

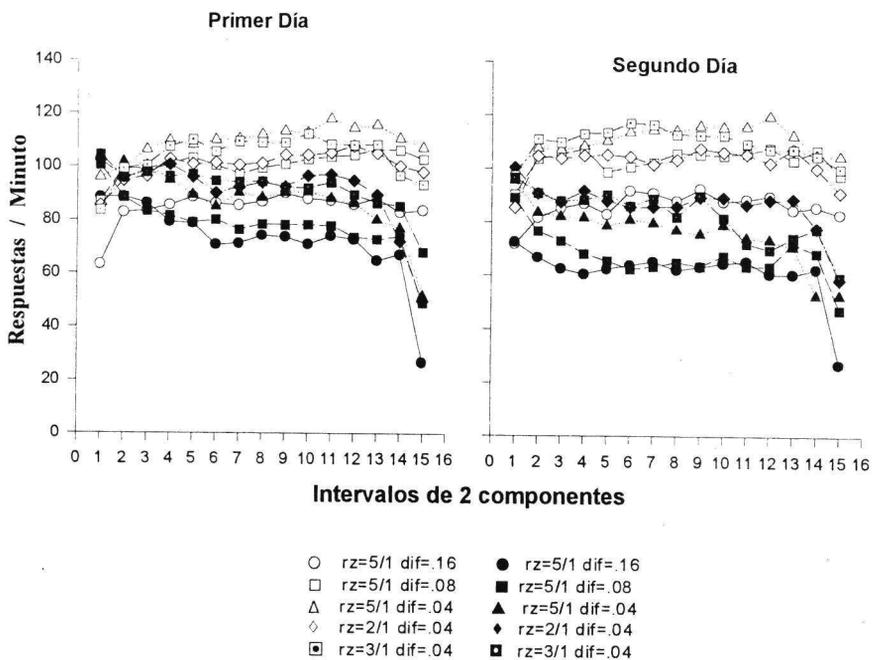


Fig. 6. Muestra el promedio total de las tasas absolutas de respuesta como una función de los intervalos de componentes para el 1^{er} día en transición (panel superior) y para el 2^o día en transición (panel inferior) para total de los sujetos en todas las condiciones y para los cuatro ciclos del experimento. Los símbolos abiertos representan las tasas de respuesta en los componentes ricos y los símbolos cerrados representan las tasas de respuesta en los componentes pobres.

Aprendizaje y Desarrollo de Preferencias

Por último la figura 7 muestra una interacción entre condición, día e intervalo para el total de los sujetos en cada una de las condiciones. El MANOVA refleja un efecto levemente significativo de intervalo [$F(56,168) = 3.52$, $p < .000$], se realizó un análisis por contrastes de cada una de las condiciones a través de los intervalos y se pudo ver un efecto lineal estadísticamente significativo para cada una de las condiciones por separado. En la interacción entre condición, día e intervalo, el MANOVA también refleja un efecto leve estadísticamente significativo [$F(56,168) = 3.69$, $p > .000$] se pudo apreciar que las tasas relativas de respuesta fueron más extremas en el segundo día que en el primero y se pudo apreciar un efecto de recuperación espontánea parecido al reportado en una de las investigaciones de Mazur (1995) ya que en el presente estudio las tasas relativas de respuestas en el segundo día, en el primer intervalo tienden a .50 o menos en todas las condiciones, como en los primeros intervalos del primer día, lo cual sucede en el estudio reportado por Mazur donde las tasas de adquisición de las primeras sesiones post-transición se mantenían constantes o parecidas a las primeras sesiones de transición, al inicio de cada sesión post-transición la proporción de respuesta para la opción rica fue menor que en al final de la sesión precedente (última en transición). Esta recuperación espontánea era mayor en la primera sesión de post-transición y progresivamente fue tornándose menor en las sesiones subsiguientes.

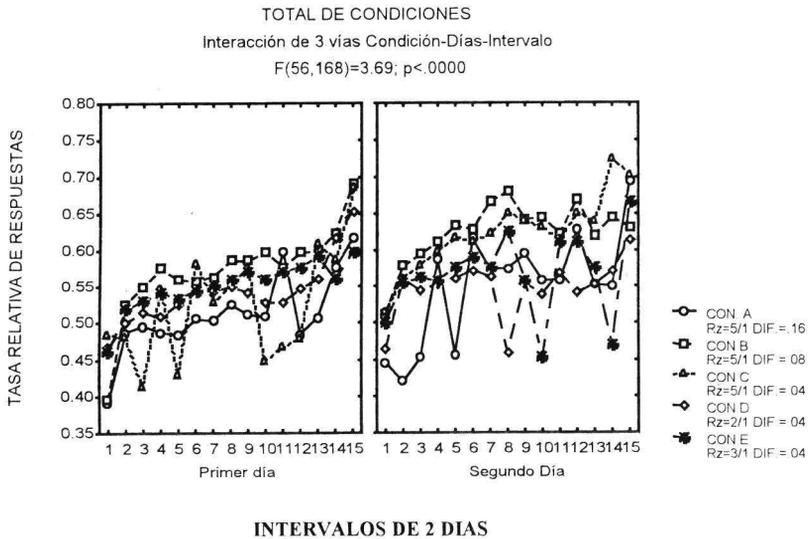


Fig. 7 Se presenta la interacción Condición-Día-Intervalo para cada una de las condiciones, la gráfica de la izquierda es la Condición-Intervalo en los primeros días de transición y la gráfica de la derecha es la misma interacción pero para los segundos días de transición.

Aprendizaje y Desarrollo de Preferencias

DISCUSION Y CONCLUSIONES

El objetivo del presente trabajo fue probar si bajo las condiciones propuestas en el presente estudio, nuestros resultados se ajustaban a los hallazgos tanto empíricos como teóricos, particularmente en relación con los resultados reportados en programas concurrentes por Mazur y sus colegas. Nuestra estrategia fue examinar los factores que afectan la tasa de adquisición de una preferencia cuando dos alternativas temporalmente distantes proporcionan reforzadores con diferentes probabilidades, y observar si nuestros sujetos eran sensibles a las razones o a las diferencias o a ambas, en las probabilidades de reforzamiento.

Los principales hallazgos del presente experimento fueron que cuando las diferencias ($\delta_1 - \delta_2$) entre dos probabilidades de reforzamiento (alta y baja) son constantes, el desarrollo de la preferencia fue mayor o más rápido cuando la razón de probabilidades (δ_1 / δ_2) fue mayor (5 a 1) lo cual es congruente con los resultados reportados por Mazur. Sin embargo, cuando la razón se mantiene constante dicha tasa de adquisición fue mayor o más rápida cuando la diferencia entre las probabilidades de reforzamiento era menor (.08), al contrario de los que sucede con los datos reportados por Mazur donde la condición con la diferencia más alta era la que más rápido se adquiría. La diferencia mayor en el desarrollo de la preferencia se dio en la condición C (RV20-RV100) en el programa con la razón de reforzamiento más alta y con la diferencia más pequeña, pero el desarrollo de preferencia más rápido fue para la condición B (RV10-RV50) y el más bajo y más lento fue para la condición D (RV13-RV25). Varias conclusiones pueden extraerse acerca del desarrollo de preferencias en situaciones de elección sucesiva. Primero, con programas múltiples RV-RV, la adquisición de una preferencia ocurre más rápidamente con razones mayores aun cuando la diferencia entre las dos probabilidades de reforzamiento, hayan sido constantes. Este resultado ha sido obtenido en programas concurrentes en procedimientos de ensayo discreto (Bailey y Mazur, 1990) y procedimientos de operante libre (Mazur y Ratti, 1991; Mazur, 1992).

En términos generales los resultados reportados con programas concurrentes de razón variable en estado estable (Herrnstein y Loveland; 1975) han mostrado que los sujetos responden exclusivamente en el programa con el requisito de respuesta menor, esto es, minimizan el

Aprendizaje y Desarrollo de Preferencias

número de respuestas emitido por reforzador obtenido y también maximizan la tasa promedio de reforzamiento. Resultados reportados en programas RV o RR (estos últimos de razón probabilística, programas en los cuales cada respuestas sobre cierta tecla tiene una probabilidad fija de ser reforzada) en transición como los reportados por Mazur y colaboradores., no hacen mención de las tasas absolutas, dado que su objetivo principal era examinar los distintos factores que pudieran afectar la tasa de adquisición de una preferencia cuando dos opciones proporcionan reforzadores con diferentes probabilidades, por lo tanto sus análisis eran más moleculares tales como los porcentajes de respuesta en cierto componente en un bloque de respuestas es decir dividían las sesiones en bloques de x número de respuestas.

Con respecto a programas múltiples RV-RV utilizados en este estudio no se encontró preferencia exclusiva. Tal y como lo demuestra un estudio realizado por Bouzas (no publicado) con programas múltiples RV-RV en estado estable donde se estudiaron 4 distintos valores de RV (10, 30, 60 y 90) y sus diferentes interacciones. Los resultados muestran una disminución en la tasa de respuesta cuando en el componente con mayor valor de RV ésta disminución era más acentuada cuando las diferencias entre los valores de los componentes era mayor (por ejemplo RV10-RV90), aun así los animales de dicho experimento nunca dejaron de responder en el componente con densidad de reforzamiento baja, en ninguna de las condiciones. Como se puede ver en los resultados encontrados a nivel molar nuestros sujetos nunca dejaron de responder al color asociado al componente pobre o de baja probabilidad de reforzamiento (preferencia exclusiva) en ninguna de las condiciones como sucede en los programas concurrentes RV-RV, tal como menciona Reynolds (1963) “la relación entre la tasa de respuesta y la tasa de reforzamiento en un componente en un programa múltiple es muy diferente de la relación encontrada en la ejecución en un programa concurrente”, inclusive las tasas de respuesta en el componente pobre fueron muy altas en la mayoría de las condiciones, de hecho la tasa más baja en dicho componente fue la obtenida en la condición A donde el requisito de respuesta en dicho componente en esta condición no fue muy alto, de hecho fue el requisito de respuesta mas bajo en comparación con todas las otras condiciones. Por otro lado las diferencias en las tasas de respuesta en los componentes ricos como en los componentes pobres en casi todas las condiciones es notoria, se aprecia que conforme la razón de un programa a otro (δ/δ) es mayor (la mayor razón utilizada fue de 5 a 1, Condiciones A,B y C) la diferencia en la preferencia de los

Aprendizaje y Desarrollo de Preferencias

componentes también es mayor conforme esta razón disminuye (la menor razón utilizada fue de 2 a 1, Condición D) esta diferencia disminuye. Por el lado de los valores absolutos o diferencia utilizadas ($\delta - \delta$) se manipularon tres diferencias y se pudo observar que si hay una diferencia significativa en todas las condiciones, en las condiciones A, B y C las diferencias entre las tasas de respuesta son mayores, notese que en estas condiciones la razón fue de 5 a 1, la condición C fue la de menor diferencia o valores absolutos (.04) y es donde la diferencia en la asignación de respuestas para un componente u para otro es más notoria, aun cuando la tasa relativa mas alta a través de los intervalos fue en la condición B, ahora si comparamos las condiciones C, D y E donde la diferencia es constante (.04) pero la razón varía (5 a 1, 2 a 1 y 3 a 1 respectivamente) la Condición C sigue siendo la de mayor diferencia en la asignación de respuestas en un componente u otro.

¿Cómo podemos explicar los anteriores resultados y que podemos concluir de ellos?
¿Porqué los sujetos siguen respondiendo en el componente pobre de baja probabilidad de reforzamiento a diferencia de lo que sucede con programas concurrentes ?. Bien, la anterior pregunta puede ser respondida en una primera instancia de una manera muy intuitiva, a los sujetos no les cuesta nada responder en el componente pobre y de cierta forma maximizan la tasa total de reforzamiento, pero, si comparamos estos resultados con lo reportado en programas concurrentes donde los sujetos muestran preferencia exclusiva por la opción de menor requisito de respuesta o de mayor probabilidad de reforzamiento, podemos responder que en los programas concurrentes las respuestas de los sujetos están compitiendo por un tiempo común, es decir si el sujeto responde el componente o en la opción de menor probabilidad de reforzamiento le esta restando tiempo a la otra opción dado que las dos opciones son simultáneas, en términos más coloquiales digamos que hay una restricción temporal en los programas concurrentes, el tiempo que el sujeto emplea en responder en el componente con baja probabilidad de reforzamiento lo podría utilizar en la opción de alta probabilidad de reforzamiento de hecho se podría interpretar como una penalización “si respondes en el pobre dejas de ganar en el rico), en los programas múltiples como mencionamos en la introducción tal restricción no existe, al sujeto no le está costando nada responder en el componente pobre de hecho en términos de reforzamiento si no respondiera estaría perdiendo reforzadores aunque ganaría en términos de “ocio” una cuestión interesante como propuesta de experimento sería ver qué sucede si al sujeto

Aprendizaje y Desarrollo de Preferencias

se le impusiera una restricción, es decir que pasaría si alterno a los componentes del programa múltiple estuviera corriendo un programa IV constante para ambos componentes con el valor promedio de ambos componentes (digamos como un programa concurrente-múltiple) al sujeto le costaría responder al componente pobre ya estaría perdiendo tiempo de responder en el componente alterno y por lo tanto reforzadores.

En cuanto las posibles explicaciones en las diferencias de las tasas de respuesta para ambos componentes estas se darán más adelante en conjunto con las posibles explicaciones para los resultados a nivel molecular.

Como se menciona en la introducción en los estudios de Mazur y cols. En los experimentos reportados (Bailey y Mazur, 1990; Mazur, 1992, 1995, 1996; Mazur y Ratti, 1991) con programas concurrentes de reforzamiento de razón probabilística tanto de operante libre como de ensayos discretos examinaban como las tasas de respuestas en las fases de transición o de dos probabilidades eran afectadas por un número de diferentes factores, tales como la probabilidad promedio de reforzamiento, el tamaño de un cambio en las probabilidades y la tasa promedio de reforzamiento.

En cuanto a los resultados encontrados en las tasas relativas de adquisición en comparación con los hallazgos reportados por Mazur y cols. En una primera instancia parecerían ser congruentes ya que si comparáramos los resultados de los porcentajes de respuesta por bloques de cien respuestas (experimento de Mazur (1992)) y las tasas relativas de respuesta por intervalos de dos componentes o dos minutos (nuestro experimento), en el primer bloque o intervalo de las sesiones los sujetos responden en promedio a nivel de .5, es decir muy parecido en ambos componentes dado que vienen de las fases de entrenamiento o de una probabilidad hay que recordar que en estas fases la probabilidad de reforzamiento es la misma para ambos componentes, pero pasando este primer bloque o intervalo, las tasas de adquisición van incrementando y siguen así hasta el final de los intervalos o bloques o de manera más general hasta el final de la sesión, en cada una de las condiciones, así mismo en cuanto a las diferencias en las tasas de adquisición en cuanto a las razones utilizadas (Mazur solo utilizó dos de 5 a 1 y de 2 a 1, nosotros utilizamos tres 5 a 1, 3 a 1 y 2 a 1), se puede ver de manera muy general que las

Aprendizaje y Desarrollo de Preferencias

tasas de adquisición son más lentas cuando las razones utilizadas son menores (2 a 1) a nivel de las diferencias (Mazur utilizó 6: .16, .12, .08, .06, .04 y .02 nosotros solo 3: .16, .08 y .04) se pudo observar un hallazgo raro cuando las razones eran de 5 a 1 la tasa de adquisición más rápida no fue la de mayor diferencia (.16) en ambos casos, si no la de un valor inmediato anterior (.12 en el experimento de Mazur) y (.08 en nuestro experimento), las tasas de adquisición más lentas en ambos experimentos fueron las de menor diferencia (.04). Sin embargo un análisis más detallado muestra ciertas diferencias. Se puede observar que las curvas de adquisición en nuestro experimento no son tan rápidas, y no alcanzan los valores reportados por Mazur, nuestros valores máximos en estas tasas de adquisición son si acaso del .6 o 60% parecidas a las de Mazur solo cuando las razones son de 2 a 1, pero los valores en las razones de 5 a 1 en su experimento alcanzan un valor máximo de 75% y 85%, al parecer de igual manera que con las tasas absolutas, las posibles causas son debido, a que los sujetos no les cuesta nada responder en el componente pobre en los programas múltiples, de hecho los valores extremos obtenidos por Mazur son alcanzados cuando las diferencias en los valores es mayor (.16 y .12) en ambas razones, sin embargo, en nuestro experimento en la condición (A) donde utilizamos una de estas diferencias (.16) los valores obtenidos en la tasa relativa de adquisición estuvieron muy por debajo de los resultados reportados por Mazur (1992), de hecho dos condiciones con diferencias menores estuvieron por arriba de las tasas relativas de esta condición, nosotros atribuimos a que dicha condición (con valores RV5-RV25) los valores utilizados en los RVs eran extremadamente ricos (ambos) y dado que en cada componente y sesión descontábamos el tiempo de reforzamiento de la sesión completa, las sesiones en esta condición duraban en tiempo real más que en todas las otras condiciones, ya que se obtenían bastantes reforzadores, por lo tanto había que descontar ese tiempo de la duración del componente y de la sesión. Por otro lado los efectos que se ven en el último intervalo en casi todas las condiciones, a excepción de la condición C donde este efecto no es tan extremo, lo atribuimos a un efecto de saciedad, ya que como se puede apreciar en las otras 4 condiciones, las probabilidades de reforzamiento no eran tan bajas como en la condición C, de tal manera que los sujetos al llegar al último intervalo de las sesiones estaban saciados, de hecho en la condición A hubo sujetos que al final de una sesión volvían a su peso ad lib de 95%, caso que nunca ocurrió en la condición C donde las probabilidades de reforzamiento fueron bajas en ambos componentes, en comparación con las otras condiciones. Sin embargo, dichos intervalos o efectos no fueron decisivos para los análisis reportados ya que se hicieron

Aprendizaje y Desarrollo de Preferencias

análisis tanto con esos intervalos como sin ellos y no hubo una diferencia estadísticamente significativa entre uno y otro análisis.

En relación con las tasas absolutas de respuesta por intervalo de dos componentes o dos minutos es obvia la diferencia entre el número de respuestas que los sujetos emiten en las fases de una probabilidad o entrenamiento y las fases de dos probabilidades o transición para cada una de las condiciones y también la diferencia entre el 1^{er} día de transición y el 2^o día, en este último la diferencia entre las tasas de respuesta para el componente pobre y para el componente rico son más acentuadas que en el 1^{er} día, el efecto tal y como sucede en los programas concurrentes en donde los sujetos incrementan sus tasas de respuesta en el componente rico y la disminuyen en el componente pobre, no es tal cual, sino, más bien lo último, los sujetos solo disminuyen su tasa de respuestas en el componente pobre, ya que no hay un incremento notable en la tasa de respuestas en el componente rico a diferencia del primer día y obviamente ni el promedio total de ambos días, suponemos que si las sesiones de transición o fases de 2 probabilidades se hubieran corrido un día más es decir 3 días en transición en lugar de 2 estas diferencias serían aun más notables.

No obstante que estos hallazgos son fáciles de resumir, ellos poseen dificultades para varias teorías de adquisición. Cada uno de los modelos mencionados en la introducción pueden en primer lugar predecir correctamente algunos de los hallazgos básicos acerca de la adquisición de preferencia en programas concurrentes, sin embargo, muchos de ellos no pueden explicar la totalidad de los resultados en los mismos programas concurrentes como lo señalan Bailey y Mazur (1990) ; Mazur y Ratti (1991) y Mazur (1996) sus resultados son inconsistentes con el modelo cinético, el modelo de operador lineal simple, el modelo de razón-invarianza y la versión original de la teoría de mejoramiento, en relación con nuestro experimento estos modelos también son insuficientes para explicar los hallazgos encontrados, en primera por una restricción formal en torno a las ecuaciones de los modelos mencionados ya que solo utilizan una ecuación dada la naturaleza de los programas concurrentes, ya que los sujetos si responden a una opción dejan de responder a la otra, por lo cual sería irrazonable utilizar dos ecuaciones en los modelos, caso contrario ocurre con los programas múltiples donde los sujetos pueden responder a ambas opciones sin que sus respuestas compitan por un tiempo común. Además de que la habilidad de los sujetos según los resultados hallados, para discriminar y responder a ciertas propiedades

Aprendizaje y Desarrollo de Preferencias

locales de los programas es interesante por sí misma y es consistente con un amplio cuerpo de investigación previa. Sin embargo, las ecuaciones básicas que intentan describir los efectos de reforzamiento y no-reforzamiento en el modelo de operador lineal, en el modelo de razón-invarianza, la teoría cinética y la teoría de Couvillon y Bitterman (1985) no están diseñadas para tratar con tales efectos y sería irrazonable esperar que tales teorías predigan nuestros resultados.

Los problemas con estos modelos no significa, sin embargo, que modelos más complejos o elaborados sean necesarios para explicar los principales hallazgos aquí reportados, sino que tal vez la utilización de algún modelo simple que responda cuestiones básicas acerca de las situaciones de elección estándar tales como : ¿ Cómo es afectada la tasa de transición por las probabilidades de reforzamiento o por diferentes tasas de reforzamiento ? Si un modelo matemático puede predecir alguna cosa acerca de conducta de elección en transición este sería capaz de responder dichas cuestiones básicas correctamente. El hecho de que no todos los modelos hagan predicciones correctas sugiere que puede ser posible distinguir entre modelos de elección de adquisición adecuados e inadecuados usando situaciones de elección muy simples.

Resumiendo para los propósitos del presente estudio, la observación crítica es que los modelos actuales de aprendizaje que dependen principalmente del patrón de probabilidad de reforzamiento obtenido para opciones de respuesta individuales, incluyendo el modelo de efectos acumulativos propuesto por Davis *et al.* (1993), no pueden dar cuenta de las características críticas de la literatura experimental existente. No obstante, esta en disputa el tipo de modificación que podría ser requerido. Una opción, es añadir en detalle a los modelos basados en probabilidad de reforzamiento suposiciones adicionales acerca de la naturaleza de la unidad de respuesta. Una segunda opción pero más drástica según Williams (1994b) es abandonar las suposiciones de fuerza de respuesta incremental, que subyacen a tales modelos, en favor de informes representacionales que asumen que los sujetos en experimentos de condicionamiento adquieren conocimiento acerca de las fuentes de recompensa en su medio ambiente. En este sentido nuestro siguiente paso sería postular a partir de la evidencia obtenida en este y otros trabajos, un modelo que asuma que los organismos aprenden (representan) tres tasas de reforzamiento, las dos asociadas con cada uno de los componentes de los programas múltiples y

Aprendizaje y Desarrollo de Preferencias

otra que es el resultado de computar la tasa de reforzamiento a lo largo de la sesión y se propondría que en las situaciones de elección intertemporal, el efecto de un reforzador es proporcional a la diferencia entre la tasa de reforzamiento estimada para un componente y la estimación de la tasa global computada a lo largo de los diferentes periodos de tiempo.

Aprendizaje y Desarrollo de Preferencias

ANEXO 1

Introducción a manera de Justificación.

Dado que el estudio del aprendizaje siempre ha ocupado una posición privilegiada en biología y psicología ya que es una característica distintiva de la conducta humana. Y si bien una de las principales justificaciones para cualquier investigación científica siempre ha sido su aplicación potencial. Podemos asumir que el estudio del aprendizaje en animales siempre ha sido un esfuerzo popular que puede ser justificado no solo por su interés intrínseco, sino también por su promesa de mejoramiento humano.

Posterior al auge en los 50s y 60s y de eminentes investigadores como Hull, Spence, Mowrer, Tolman, Miller y Skinner, otras maneras de aproximarse al estudio de la inteligencia humana, como psicología fisiológica, la genética de la conducta, la psicolingüística, la ecología conductual y muy particularmente la psicología cognitiva en sus diversas formas, desplazaron al aprendizaje animal de la luz pública. No obstante, casi desde los inicios de los 90s, el estudio contemporáneo del aprendizaje animal ha sufrido un impresionante aumento, reafirmando que sigue siendo un área vital de investigación que se ocupa de ciertos aspectos fundamentales del modo en que la conducta es gobernada por los acontecimientos del medio. Este estudio moderno del aprendizaje, se ve enriquecido por nuevos y numerosos hallazgos y por modos de pensar e interpretar también novedosos. Nuestras ideas básicas en torno al condicionamiento clásico y al condicionamiento instrumental han experimentado profundos cambios en los últimos 25 años, y este vigoroso proceso continúa. En suma, el estudio del aprendizaje y del condicionamiento se integra cada vez mejor con las investigaciones afines en torno a las bases biológicas de la conducta y al estudio de los procesos cognitivos. Además dicho estudio también aporta técnicas para el análisis de la conducta que resultan de gran valor en diversos campos de reciente aparición, como la neurociencia conductual, la psicobiología evolutiva, la psicofarmacología, la medicina conductual, etc. Los asombrosos hallazgos de la neurofisiología renuevan la esperanza de un rápido descubrimiento de los substratos neurológicos del aprendizaje, y este renovado interés por la neurofisiología del aprendizaje da lugar, a su vez, a un resurgimiento del interés por los mecanismos conductuales básicos del aprendizaje. Por lo tanto, aunque el estudio del aprendizaje y el condicionamiento empezó siendo asunto primordialmente de psicólogos, ahora

Aprendizaje y Desarrollo de Preferencias

es parte integral de una amplia red de enfoques interdisciplinarios que se ocupan del estudio de la conducta.

Aprendizaje y Acción.

El aprendizaje es uno de los procesos biológicos determinantes para la supervivencia de diversas formas de vida animal. La integridad de la vida depende de distintas funciones biológicas. Los animales tienen que tomar alimentos, eliminar los desechos del metabolismo y, por otra parte, mantener el equilibrio adecuado del funcionamiento interno. Mediante la evolución se han generado gran variedad de sistemas biológicos para realizar estas tareas. Sin embargo, los procesos fisiológicos internos, perfectamente armonizados, a menudo no bastan para mantener la integridad de la vida. Los animales y las personas vivimos en ambientes que están en constante transformación por los cambios climáticos, cambios en las fuentes alimenticias, la llegada y partida de depredadores, etc. Los efectos adversos de estos cambios a menudo son minimizados por medio de ajustes conductuales. Los animales tienen que saber, por ejemplo, cómo encontrar y obtener comida, cómo evitar a los depredadores cuando están en su territorio, y cómo encontrar nuevas guaridas cuando los cambios climáticos lo demandan. Lograr estas tareas requiere, obviamente, movimientos motores tales como andar y manipular objetos. Dichas tareas también requieren la capacidad de predecir acontecimientos importantes del entorno ambiental, como la disponibilidad de la comida en un lugar y en un momento concretos. La adquisición de una conducta motora nueva y de nuevas reacciones anticipatorias implica aprendizaje. Por tanto, los animales aprenden a ir a una fuente nueva de alimentación cuando en la vieja se terminó el alimento, y aparecen conductas anticipatorias nuevas cuando aparecen nuevas fuentes de peligro. Estos reajustes aprendidos en el ambiente no son menos importantes para la supervivencia que los procesos fisiológicos internos, como la respiración o la digestión.

La mayoría de la gente asocia automáticamente el aprendizaje con la adquisición de una nueva conducta. Es decir el aprendizaje se identifica por la aparición de una nueva respuesta en el repertorio del organismo. Sin embargo, el cambio de conducta implicado en el aprendizaje puede también consistir en la disminución o pérdida de una conducta del repertorio del

Aprendizaje y Desarrollo de Preferencias

organismo. Aprender a contener respuestas es tan importante como aprender a dar respuestas, si no más.

A menudo pensamos en el aprendizaje como un proceso complejo que requiere una práctica especializada, y que da lugar a formas de conducta sofisticadas y llamativas. Pero hay sistemas de respuestas mucho más simples en los que el aprendizaje también está involucrado. Los investigadores del aprendizaje han dedicado gran parte de su esfuerzo, a estudiar los mecanismos de aprendizaje de sistemas de respuesta simples, con la esperanza de que el conocimiento obtenido de tales investigaciones les lleve a la formulación de principios generales de aprendizaje. Pero también se espera que la investigación de sistemas de respuestas simples proporcione la información básica necesaria para el estudio de formas más complejas de aprendizaje.

No existe una definición del aprendizaje aceptada universalmente. Sin embargo, no habría mucho inconveniente en aceptar la siguiente frase que contiene aspectos esenciales del concepto de aprendizaje:

El aprendizaje es un cambio duradero en los mecanismos de conducta, resultado de la experiencia con los acontecimientos ambientales, (Domjan y Bukhard, 1986).

¿Cuáles son los aspectos importantes de esta definición? Primero, se dice que el aprendizaje es un cambio en los mecanismos de la conducta, no un cambio en la conducta misma. Se define el aprendizaje como un cambio en los mecanismos de la conducta, dado que esta última está determinada por muchos factores además del aprendizaje. En este sentido es importante ilustrar la distinción entre aprendizaje y la actuación.

Estudios de Elección

Por actuación nos referimos a las acciones de un organismo en un momento concreto. Que un organismo haga una cosa u otra (su actuación) depende de muchas cosas. Incluso la

Aprendizaje y Desarrollo de Preferencias

ejecución de una respuesta simple, está determinada por multitud de factores. Así la actuación está determinada por la oportunidad, la motivación y las capacidades sensoriales y motoras, además del aprendizaje. Por tanto, no puede considerarse automáticamente que un cambio en la actuación refleje aprendizaje.

La definición anteriormente enunciada postula al aprendizaje como un cambio en los mecanismos de conducta para hacer hincapié en la distinción entre aprendizaje y actuación. Sin embargo, los investigadores no pueden observar directamente estos mecanismos. Lo que ocurre es que a partir de cambios en la conducta se infiere un cambio en los mecanismos de la conducta².

De tal manera, la conducta de un organismo (su actuación) se utiliza para proporcionar pruebas de que existe aprendizaje. No obstante, ya que la actuación está determinada por muchos factores aparte de aprendizaje, se debe tener mucho cuidado al decidir si un aspecto concreto de la actuación, refleja o no aprendizaje.

Por otro lado, como habíamos mencionado, la conducta animal y humana implica mucho más que solo aprendizaje y la simple repetición de respuestas individuales. Incluso en una situación simple como la caja de Skinner, los organismos toman parte en una gran variedad de actividades y continuamente están haciendo elecciones entre las distintas respuestas que son capaces de realizar. Además, de la aparición de una respuesta concreta depende mucho más de la disponibilidad de otras alternativas y de que los resultados de procedimientos de elección proporcionan un índice alternativo de las dimensiones subyacentes de la fuerza de respuesta. Como argumenta Herrnstein (1970), toda la conducta es elección, en el sentido de que siempre hay otras alternativas que la respuesta medida por el experimentador, esto es, los sujetos tienen que estar “decidiendo” constantemente que hacer, cual respuesta ejecutar.

2. Las propiedades de un sistema bajo estudio dan valor o sentido a los datos observables aun cuando estas propiedades en si mismas no son observables. Entonces inferencias entrelazadas permiten construir la validación, suposiciones reales u objetivas acerca de las propiedades no observables. Esta técnica es algunas veces llamada validación convergente (Garner, Hake y Eriksen, citado en Lachman, Lachman y Butterfield; 1979); dado que datos de diferentes clases convergen sobre una conclusión, la conclusión entonces es convergentemente validada.

Aprendizaje y Desarrollo de Preferencias

Ahora bien, supongase medidas de “fuerza absoluta” de cierta respuesta estudiada en aislado (por ejemplo, la velocidad de carrera en un callejón en T) son en efecto el resultado de las dinámicas de la elección (Williams, 1994a); entender los mecanismos de la elección de respuestas es también fundamental para la comprensión de la conducta, dado que las elecciones que los organismos hacen determinan la aparición de respuestas individuales. Las elecciones que están al alcance de los animales y el hombre son muy complejas. Analizar todos los factores que controlan la conducta individual sería una tarea formidable si no es que imposible. Por lo cual, los psicólogos comenzaron las investigaciones experimentales de los mecanismos de la elección estudiando situaciones más simples. La situación de elección más simple es aquella en que el sujeto tiene dos opciones de respuesta y cada respuesta va seguida de un reforzador de acuerdo con un patrón de reforzamiento.

Enfoques relativamente recientes del estudio de la elección utilizan cajas de Skinner equipadas con dos manipulandos, como por ejemplo, dos palancas o dos teclas. En el experimento típico, las respuestas en cada manipulando se refuerzan con un patrón de reforzamiento distinto. Los dos programas actúan simultáneamente, y el sujeto es libre de pasar de un manipulando a otro. Esta clase de procedimiento se llama programa concurrente de reforzamiento. Dichos programas permiten medidas continuas de elección porque el sujeto es libre de cambiar de una opción de respuesta a la otra. La preferencia se mide por la tasa de respuesta en cada manipulando o por el tiempo que el sujeto pasa respondiendo en cada una de ellas.

De tal modo elección ha sido un tema importante en investigación operante en los pasados treinta años. Pichones, ratas y seres humanos han sido estudiados sobre una variedad de procedimientos en los cuales respuestas repetitivas simples, tales como picar una tecla o presionar una palanca, son intermitentemente reforzadas de acuerdo con varios programas. La distribución de conducta entre alternativas de elección, comunmente teclas o palancas izquierdas o derechas, han sido medidas como una función del tipo de programa, parametros del programa, y la tasa de reforzamiento obtenido. (Ver Williams, 1988; y varios capítulos de Honig y Staddon, 1977; para una revisión).

Aprendizaje y Desarrollo de Preferencias

La mayor parte de este período de 30 años, el énfasis teórico ha sido sobre principios de equilibrio molar, estático y reversible de entre los cuales ha destacado la ley de igualación (Davison y McCarthy, 1988; Herrnstein, 1961). Igualación es una relación empírica demostrada más claramente en elección entre dos o más programas de reforzamiento intervalo variable (IV) (por ejemplo, programas en los cuales la primera respuesta después de un tiempo variable desde el reforzamiento anterior es reforzada). El hallazgo aquí es que bajo condiciones apropiadas de estado estable la razón de respuestas, x/y es aproximadamente igual a la razón de tasas de reforzamiento obtenido $R(x)/R(y)$, ya que, $X/(x+y) = R(x)/[R(x)+R(y)]$. Esto es importante, ya que, a pesar de que igualación es usualmente expresada como igualdad de proporciones de repuestas y reforzadores, esto también puede ser expresado como igualdad de probabilidades de reforzamiento : $R(x)/x = R(y)/y$ La Ley de igualación es reversible en el sentido de que igualación es usualmente un estado estable que finaliza bajo ciertas condiciones de reforzamiento, ampliamente independiente de la historia previa del organismo. Igualación es molar porque las tasas x , y , $R(x)$, y $R(y)$ son medidas sobre un extenso período comunmente varias horas (por ejemplo, el promedio de las últimas cinco sesiones experimentales bajo una cierto grupo de programas). Resumiendo, igualación es un principio de equilibrio (más que una ley causal) ya que relaciona dos cantidades, tasas de respuesta y tasas de reforzamiento, que son mutuamente dependientes. Las relación estática de igualación es compatible con varios procesos dinámicos de elección (Hinson y Staddon, 1983). No obstante, una comprensión de los procesos específicos o procesos que subyacen a elección ofrecería varias ventajas. Nos ayudaría a explicar no solo igualación sino quizás también desviaciones sistematicas de ella. Permitiría decir algo acerca de elección sobre una base momento a momento -idealmente en tiempo real, pero sino, al menos sobre una base de respuesta a respuesta- abriría la manera para entender que la conducta depende de la historia pasada remota, es decir, sobre condiciones experimentales precedentes a la actual. También diríamos algo acerca de las propiedades transitivas de una historia de entrenamiento particular, tales como efectos sobre conducta en extinción o sobre el aprendizaje de alguna nueva tarea.

El objetivo general es explicar el efecto de historias complejas de conducta de elección en sujetos individuales. La estrategia es comparar modelos computables simples, esperando aprender de los éxitos y fracasos de modelos particulares, qué propiedades del modelo son

Aprendizaje y Desarrollo de Preferencias

críticas para un entendimiento mayor de la conducta de elección. Aparte de exponer las propiedades críticas en varias combinaciones a la vez, esperamos llegar a modelos cada vez más comprensivos de la elección.

El esfuerzo para explicar historias complejas es de alguna manera más ambiciosa que la típica tarea utilizada por las teorías de elección: explicar relaciones de estados estables reversibles o a lo más, los efectos de transiciones simples, tales como funciones de adquisición o extinción, o bien la transición entre una discriminación y sus inversos. La cuestión es que estaremos satisfechos en esta etapa con hacer predicciones cualitativas correctas de una amplia serie de datos, más que predicciones cuantitativas exactas de una serie de datos más restringida.

Aprendizaje y Desarrollo de Preferencias

ANEXO 2

Niveles en el sistema nervioso.

Las discusiones concernientes a la naturaleza del fenómeno psicológico y sus bases neurobiológicas invariablemente hacen mención a la noción de niveles. Tratando de ser mas preciso acerca de lo que entendemos por “nivel”, se han encontrado tres diferentes ideas acerca de los ellos en la literatura: niveles de análisis, de organización, y de procesamiento. Estrictamente hablando, las distinciones están dirigidas a través de las siguientes líneas: **i)** niveles de organización son esencialmente anatómicos y se refieren a la jerarquía de componentes y a las estructuras que forman estos componentes, **ii)** niveles de procesamiento son fisiológicos, y se refieren a la localización de un proceso relativo a los transductores y músculos y **iii)** los niveles de análisis son conceptuales y se refieren a las diferentes clases de cuestionamientos acerca de como el cerebro ejecuta una tarea: dentro de que subareas el cerebro divide las tareas, que pasos de procesamiento ejecuta una subarea, y que estructuras físicas llevan o realizan dichos pasos? Nosotros nos concentraremos solo en los niveles de análisis, dado que nuestro interés es meramente en torno a los mecanismos psicológicos y no neurobiológicos.

Niveles de Análisis

Un marco para una teoría de niveles, articulado por Marr †(1982), provee un importante e influyente antecedente para pensar acerca de los niveles en el contexto de computación para estructuras nerviosas³. Esta estructura plantea la concepción de niveles para las ciencias de la computación, y de acuerdo con Marr esta estructura esta caracterizada en tres:

(1) El nivel computacional abstracto del análisis del problema, descomponer la tarea (ejemplo, determinar la profundidad de los objetos en 3era. Dimensión del patrón de la retina en 2a. Dimensión.) dentro de sus constituyentes principales; **(2)** El nivel del algoritmo, se especifica un procedimiento formal para ejecutar la tarea ademas de dar una entrada (input), los resultados de

3. La concepción original de niveles de análisis puede ser encontrada en Marr & Poggio (1976, 1977). Mientras que Marr (1982) enfatiza la importancia de el nivel computacional, la noción de jerarquía de niveles proviene de un trabajo anterior de Reichardt & Poggio (1976) sobre el control visual de orientación en la mosca. En este sentido, la visión actual sobre la interacción entre niveles no esta mucho mas alejada de los anteriores puntos de vista como un retorno a la practica previamente establecida por Reichardt, Poggio, y aun por Marr mismo, quien publico una serie de escritos sobre modelos de redes neuronales de la corteza cerebral y la corteza del cerebelo (ver, por ejemplo, Marr (1969) y (1970), citado en Churchland & Sejnowski (1992)). El énfasis sobre el nivel computacional, sin embargo ha sido una importante influencia sobre los problemas y resultados que conciernen a la generación actual de modelos neuronales y conexionistas (Sejnowski et al. 1988).

Aprendizaje y Desarrollo de Preferencias

salida (Output) correctos; y **(3)** El nivel de implementación física, se construye un dispositivo de trabajo usando una tecnología particular. Esta división corresponde realmente a tres diferentes tipos de cuestiones que pueden ser sobresalientes acerca del fenómeno: **(1)** ¿Cómo separo el problema en partes?, **(2)** ¿Qué principios gobiernan a las partes que interactúan para resolver el problema?, y **(3)** ¿Cuál es el material cuyas interacciones causales implementan los principios?

Un importante elemento en la perspectiva de Marr fue que la solución de un nivel alto fue ampliamente independiente de los niveles bajos, y por tanto el problema computacional del nivel alto puede ser analizado independientemente del entendimiento o comprensión del algoritmo que ejecuta la computación. Similarmente, el problema del algoritmo del segundo nivel fue pensado para ser resuelto independientemente del entendimiento o comprensión de la implementación física. De este modo su estrategia preferida fue de arriba-abajo más que de abajo-arriba. Al menos esta fue la doctrina oficial de pensamiento, en la práctica, en un vistazo hacia atrás, figuran significativamente en los esfuerzos de Marr por encontrar un análisis de problemas y soluciones algorítmicas. Irónicamente, dada su vocación de la estrategia de arriba hacia abajo. El trabajo de Marr fue en si mismo altamente influenciado por consideraciones neurobiológicas, y los actos de implementación restringieron su elección del problema y la naturaleza de sus insights computacional y algorítmico.

Desafortunadamente, dos aspectos muy diferentes fueron confundidos en la doctrina de independencia. Uno se refiere a , -el cómo una *materia de descubrimiento*, se puede configurar fuera el algoritmo relevante del análisis del problema, independientemente de los hechos acerca de la implementación. El otro concierne a , -cómo una "materia de teoría formal", tal como un algoritmo que es aproximadamente conocido para ejecutar una tarea en cierta maquina (por ejemplo, el cerebro) puede ser implementado en cualquier otra maquina la cual tiene una arquitectura diferente. Hasta ahora en lo concerniente a esto ultimo, a cuál o a qué teoría computacional llamamos algoritmo tal que pueda ser corrido en diferentes maquinas, y en que sentido y solo en que sentido, el algoritmo es independiente de la implementación. El punto

Aprendizaje y Desarrollo de Preferencias

central directamente es: ya que un algoritmo es formal, no especifica los parámetros físicos (Por ejemplo, tubos al vacío, Ca^{2+}) son parte del algoritmo.

Es decir, es importante ver que el punto meramente formal no puede hablar del aspecto de cómo descubrir en efecto el mejor algoritmo usado por cierta máquina, o como arribar mejor neurobiológicamente a la tarea de análisis adecuada. Ciertamente esto no puede decirnos que el descubrimiento de los algoritmos relevantes a las funciones cognitivas sean independientes de un detallado entendimiento del sistema nervioso. Además, eso no nos dice que alguna implementación es mejor que otra. Y eso lo hace mejor o no, ya que diferentes implementaciones exhiben enormes diferencias en rapidez, tamaño, eficacia, elegancia, etc. La independencia formal del algoritmo en relación a la arquitectura es algo que podemos explotar para construir computacionalmente máquinas equivalentes una vez que conozcamos como trabaja el cerebro, sin embargo esto no es una guía para descubrir, si no sabemos como trabaja el cerebro.

El resultado de independencia de niveles marca una principal diferencia conceptual entre Marr (1982) y la actual generación de investigadores estudiando modelos neuronales y conexionistas (Sejnowski *et. al.* 1988; McClelland, J. y Rumelhart, D. 1988; Bailey, A. 1997). En contraste, para la doctrina de independencia, investigación actual sugiere que consideraciones de implementación juegan un papel vital en la clase de algoritmos que son planeados y la clase de insights computacionales disponibles para el científico. El conocimiento de la arquitectura cerebral, lejos de ser irrelevante para el proyecto, puede ser la base esencial y un invaluable catalizador para proyectar y planear probables y poderosos algoritmos -que tengan una razonable conjetura en explicar como en efecto trabajan las neuronas.

Aprendizaje y Desarrollo de Preferencias

ANEXO 3

¿Cómo se asocian los estímulos condicionados e incondicionados? : Un acercamiento al Modelo de, Rescorla & Wagner.

Un problema esencial a tratar con detalle dentro del condicionamiento clásico, y alrededor del cual ha girado la literatura tanto teórica como experimental en los últimos 30 años, ha sido la forma en que se asocian los estímulos condicionados e incondicionados.

¿Cuáles son los mecanismos del aprendizaje por asociación, los procesos subyacentes que se activan intensamente con los procedimientos de condicionamiento que producen un aprendizaje rápido, y que se activan débilmente con procedimientos menos eficaces para producir aprendizaje? Varios experimentos han demostrado que el condicionamiento clásico no se produce simplemente por las presentaciones contiguas del EC y el EI. Más bien, la relación de señal o contingencia, entre el EC y el EI, es lo esencial para que se produzca aprendizaje. ¿Cómo detecta el organismo esas relaciones de señal y cómo influyen en el aprendizaje de asociaciones? Estas cuestiones y otras parecidas han sido tema de estudio intenso, como señalamos anteriormente, durante los últimos 30 años. De este esfuerzo han surgido muchas ideas nuevas, y como resultado nuestra concepción del condicionamiento clásico ha cambiado radicalmente. Aun hoy la evolución de las teorías de condicionamiento clásico continúa. El esfuerzo de los investigadores es formular explicaciones, sino generales, si amplias de los mecanismos del aprendizaje por asociación que puedan abarcar todos los resultados emanados de las distintas investigaciones.

Dos interpretaciones del proceso de adquisición.

Uno de los hechos fundamentales del Condicionamiento Clásico es que no todos los emparejamientos del EC con el EI producen el mismo aumento en la ejecución de la respuesta condicionada. La figura A.1. muestra una curva hipotética de aprendizaje en el condicionamiento clásico. Los primeros emparejamientos o ensayos del EC y el EI producen grandes aumentos en la ejecución de la respuesta condicionada. Así en un principio, cada punto de la curva de aprendizaje es más alto que el anterior. Estos cambios en la respuesta condicionada vienen representados por ΔRC en la figura A,1 (el símbolo Δ , delta se usa para representar el

Aprendizaje y Desarrollo de Preferencias

incremento.) A medida que el condicionamiento avanza, el aumento en la ejecución de la respuesta condicionada se hace mucho mas pequeño de cada emparejamiento EC-EI el valor de ΔRC va disminuyendo cada vez más. Después de un entrenamiento suficiente, se alcanza un nivel estable de respuesta, y el ΔRC tiende a cero. Esta parte de la curva de aprendizaje se llama asíntota.

¿ Por qué en los primeros ensayos se producen aumentos mucho mayores en la ejecución de la respuesta condicionada que en los ensayos posteriores ? No se puede atribuir este efecto a la fatiga o al descenso de la motivación, puesto que se obtienen los mismos resultados cuando se realiza un ensayo al día. Existen dos explicaciones plausibles. Una de ellas analiza la curva de aprendizaje de acuerdo con los cambios en la capacidad de l estímulo condicionado para asociarse con el EI o aumentar la fuerza asociativa. Según esta concepción en el transcurso del condicionamiento el EC va perdiendo gradualmente su capacidad para asociarse con el EI. Se supone que, inicialmente, el EC tienen una gran capacidad para asociarse con el EI. Sin embargo, cuanto más se condiciona el EC, más difícil le resulta aumentar su asociación con el EI. Esta idea puede denominarse **hipótesis de la reducción del EC** . A medida que el condicionamiento avanza, disminuye la capacidad del EC de aumentar la fuerza asociativa. Como la capacidad asociativa del EC se reduce gradualmente, los aumentos de la respuesta condicionada se hacen progresivamente menores hasta que se alcanza la asíntota y el aprendizaje se acaba.

Un segundo enfoque en el análisis de la adquisición sostiene que existen cambios en la capacidad del estímulo incondicionado para producir aprendizaje. Según esta idea, el estímulo incondicionado es totalmente eficaz durante los primeros ensayos del condicionamiento, y por tanto produce grandes aumentos de la respuesta condicionada. Sin embargo, conforme progresa el entrenamiento, el EI pierde gradualmente su capacidad de producir un condicionamiento de la situación, y los aumentos de la respuesta condicionada cada vez son menores. Cuando llega a la asíntota se supone que el EI es totalmente incapaz de producir más condicionamiento, y que ΔRC es cero. Esta concepción puede denominarse hipótesis de la reducción del EI. A medida que avanza el condicionamiento, la eficacia del EI se reduce gradualmente, y cada vez se produce menos condicionamiento.

Aprendizaje y Desarrollo de Preferencias

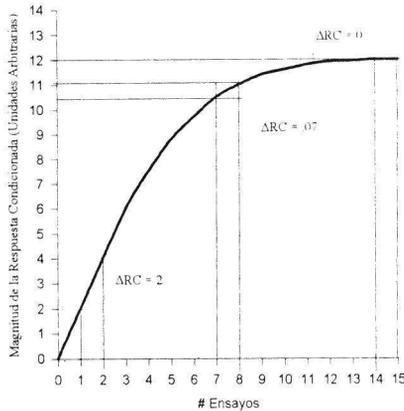


Fig. A1. Curva ideal de aprendizaje. La magnitud de la respuesta condicionada se incrementa 2 unidades ($\Delta RC = 2$) entre el primer y segundo ensayo de condicionamiento; 0.7 unidades ($\Delta RC = 0.7$) entre los ensayos séptimo y octavo y 0 unidades ($\Delta RC = 0$) entre los ensayos décimo cuarto y décimo quinto.

Las teorías contemporáneas sobre aprendizaje difieren en su concepción del proceso de adquisición. Unas teorías hacen hincapié en los cambios del EC en el transcurso del condicionamiento clásico, mientras que otras resaltan los cambios en el EI. Vamos a empezar nuestra discusión sobre las teorías del aprendizaje por asociación explorando la idea de que la capacidad del EI para producir un nuevo aprendizaje se altera como resultado del condicionamiento. Esta idea la fomentaron los investigadores del efecto de bloqueo en el condicionamiento pavloviano. Dado el papel central que el efecto de bloqueo ha tenido en los desarrollos teóricos contemporáneos, vamos a tratar sobre el a continuación.

El efecto de bloqueo.

Los animales condicionados fuera del laboratorio habitualmente se encuentran con numerosos estímulos. Los hechos que señalan peligro, por ejemplo, pueden contener múltiples estímulos visuales, auditivos y olfativos. Para el organismo, lo más ventajoso es considerar sólo aquellas claves que señalan el peligro de manera más eficaz. Es preferible olvidarse de otras claves que, o bien no lo señalan convenientemente, o son redundantes. De otra forma, el

Aprendizaje y Desarrollo de Preferencias

organismo podría estar en un constante estado de alerta o de tensión, preparado para huir o luchar innecesariamente. La hipótesis de la reducción del EI sugiere la manera de evitar que se condicionen los estímulos redundantes o que no informan adecuadamente. Condicionando un EC con un EI concreto hasta llegar a la asíntota se reduce la capacidad del EI para producir más condicionamiento en esa situación. Por tanto, si se añade un nuevo EC a la situación, se producirá muy poco condicionamiento del nuevo estímulo, si es que se produce alguno. Se ha llegado a esta conclusión gracias a una serie de experimentos sobre lo que se ha dado en llamar **efecto de bloqueo**. Estos experimentos, iniciados por Leon Kamin, describen los casos específicos en los que los procedimientos de condicionamiento no condicionan.

El efecto de bloqueo se ha investigado muy extensamente utilizando el procedimiento de supresión condicionada con ratas (Kamin 1969). El procedimiento comprendía tres fases (véase figura A. 2.). En la primera fase el grupo experimental recibía emparejamientos repetidos del estímulo condicionado A con el estímulo incondicionado. Esta fase del experimento continuaba hasta que el estímulo A se condicionaba al llegar a la asíntota, es decir hasta que el animal suprimía completamente sus respuestas de presión de la palanca al presentarle el estímulo A. El condicionamiento del estímulo A reducía presumiblemente la eficacia del estímulo incondicionado. Para evaluar la pérdida de eficacia del EI, se hizo un intento de condicionar un estímulo B, con el mismo EI. En la segunda fase del procedimiento, los estímulos A y B se presentaron simultáneamente y se emparejaron con el EI.

GRUPO	FASE 1	FASE 2	PRUEBA
Experimental	A--->EI	[A + B*] ---> EI	B
Control		[A + B*] ---> EI	B

Figura A. 2. Diagrama del procedimiento de bloqueo. Durante la fase 1, el estímulo A es condicionado con el EI en el grupo experimental, mientras que en el grupo de control no recibe ensayos de condicionamiento. Durante la fase 2 ambos grupos, experimental y de control, reciben ensayos de condicionamiento en los cuales el estímulo A y el estímulo B se presentan simultáneamente y emparejados con el EI. Una prueba posterior de la respuesta al estímulo B aislado muestra que se obtiene un menor condicionamiento al estímulo B en el grupo experimental que en el de control.

Después de varios de estos ensayos de condicionamiento, el estímulo B se presentó solo en un ensayo de prueba para ver hasta qué grado los animales aprendían a suprimir su conducta

Aprendizaje y Desarrollo de Preferencias

en presencia del estímulo B. El grupo de control recibió el mismo entrenamiento en la segunda fase con el estímulo B que los sujetos experimentales, pero a ellos no se les condicionó el estímulo A en la primera fase. Por lo tanto, el EI no iba a reducir su eficacia en la segunda fase. La hipótesis de la reducción del EI predice un menor condicionamiento del estímulo B en el grupo experimental que en el grupo de control. La prueba con el estímulo B presentado solo, al final del experimento, confirmó estos resultados. En muchas repeticiones del experimento, el estímulo B producía invariablemente menos supresión condicionada de la conducta en el grupo experimental que en el de control.

El efecto de bloqueo es curioso porque las nociones informales, o de sentido común, sobre condicionamiento clásico no podían predecirlo. Por ejemplo, el efecto de bloqueo indica claramente que el condicionamiento clásico no tiene lugar simplemente porque se presente el estímulo condicionado junto con el estímulo incondicionado. Durante la segunda fase del experimento de bloqueo, el estímulo B se empareja con el EI de forma idéntica en el grupo experimental que en el de control. Si el emparejamiento del EC y el EI bastara para que se diera el condicionamiento, el estímulo B se condicionaría en ambos grupos. El hecho de que el estímulo B se condicione sólo en el grupo de control es una prueba fehaciente de que el emparejamiento del EC y el EI no basta para producir aprendizaje. ¿Qué otros factores son necesarios? constituye una interrogante fundamental del condicionamiento clásico.

Además de demostrar el efecto de bloqueo, Kamin (1968, 1969) realizó muchos experimentos para averiguar qué aspectos de este procedimiento causaban la interferencia con el condicionamiento del estímulo B en el grupo experimental. Estos experimentos y otros han mostrado que el condicionamiento del estímulo B se bloquea si es redundante, es decir, si B no añade ninguna información acerca del EI. Para que se cumpla este requisito existen dos aspectos del procedimiento de bloqueo que son esenciales. En primer lugar, el estímulo A debe presentarse junto con estímulo B. Segundo, el estímulo A tiene que predecir adecuadamente el EI durante los ensayos de condicionamiento del estímulo B. Estas características aseguran que el estímulo A es suficiente por sí solo para señalar el EI, y el B es redundante (innecesario). Si no se cumplen las condiciones que hacen redundante el estímulo B, el bloqueo no se producirá. Por ejemplo, el estímulo A no bloquea el condicionamiento del estímulo B si el estímulo A no se

Aprendizaje y Desarrollo de Preferencias

presenta durante la segunda fase. El bloqueo tampoco se produce si el estímulo A no se condiciona con el EI durante la primera fase, o si las propiedades condicionadas del estímulo A se extinguen entre las fases primera y segunda. En ambos casos el resultado es que el estímulo A no señala las apariciones del EI durante la segunda fase, y por tanto, el estímulo B no es redundante. Estos hallazgos indican que el condicionamiento del estímulo A hasta la asíntota no reduce la eficacia del estímulo incondicionado en cualquier circunstancia. Más bien lo que ocurre es que la eficacia del EI se reduce sólo en las situaciones en que el estímulo A señala el estímulo incondicionado. Cuando el estímulo A está ausente, o se ha extinguido y ya no señala más el EI, la eficacia del estímulo incondicionado vuelve al máximo.

¿ Por qué el EI no condiciona a un estímulo redundante ? La presencia del estímulo A durante los ensayos de condicionamiento del estímulo B, en la segunda fase del procedimiento de bloqueo, hace que el EI sea totalmente predecible. Así pues, el EI no produce sorpresa en la segunda fase. Estas consideraciones le sugirieron a Kamin que el estímulo incondicionado tenía que ser "sorprendente" para producir condicionamiento. Si el estímulo incondicionado no es sorprendente, el animal no se sobresalta y no se estimula el "esfuerzo mental" necesario para la formación de una asociación. Los hechos esperados no exigen ajustes del organismo y, por tanto, no estimulan un nuevo aprendizaje. Por definición, los hechos inesperados son estímulos para los que el organismo no está ajustado. Por tanto, es más probable que los hechos inesperados creen un nuevo aprendizaje.

EL MODELO DE CONDICIONAMIENTO DE RESCORLA-WAGNER

La idea de que el estímulo incondicionado tiene que sorprender para proveer el Condicionamiento Clásico es un concepto central en las teorías contemporáneas sobre condicionamiento. Robert Rescorla y Allan Wagner (Rescorla & Wagner, 1972; Wagner & Rescorla, 1972) fueron de los primeros en desarrollar esta idea de una forma sistemática dentro de una teoría más amplia sobre condicionamiento. Estos investigadores formularon un modelo matemático del concepto de "sorpresividad" del EI. El modelo matemático tiene dos ventajas importantes sobre las anteriores descripciones verbales de la idea. En primer lugar supone un tratamiento más preciso del concepto de sorpresividad del EI. En segundo lugar, con el uso de

Aprendizaje y Desarrollo de Preferencias

derivadas matemáticas y simulaciones por computadora, las implicaciones del concepto de sorpresividad del EI pueden extenderse a una amplia variedad de fenómenos de condicionamiento. Consecuentemente, el modelo de Rescorla-Wagner dominó la investigación sobre condicionamiento hasta aproximadamente 20 años después de su formulación. a continuación describiremos con detalle el tratamiento matemático, conceptual y las implicaciones de la teoría.

COMPETENCIA ENTRE ASOCIACIONES POR LA CAPACIDAD PARA ELICITAR RESPUESTAS CONDICIONADAS.

En cualquier momento dado, hay cientos de estímulos y elementos ambientales que pueden ser procesados por un organismo, y el número de asociaciones estímulos-eventos es igualmente enorme. Hay numerosas razones para que haya o para que tengan que existir límites el número de asociaciones que son adquiridas durante un episodio particular. Nos sentiríamos abrumados si tuviéramos que aprender una asociación de todo con todo. Durante dos mil años, los filósofos han estado preocupados por las leyes que nos señalan qué estímulos serán los que se asociarán con otros. Como dijimos anteriormente, la contigüidad ha sido hecha responsable de la adquisición de asociaciones.

Muchos hallazgos, la mayoría de ellos recientes, han encontrado que la contigüidad no es una condición ni necesaria ni suficiente para el aprendizaje (ver Papini y Bitterman, 1990; Plaud y Volgelantz, 1991, para otra visión del tema). Un viejo descubrimiento que mostraba la insuficiencia de la contigüidad como soporte del aprendizaje asociativo es el ensombrecimiento. Se entiende por ensombrecimiento un efecto descubierto por Pavlov - la aparición de una pobre RC ante EC principal cuando éste es apareado con el EI y en presencia de un segundo EC. La RC débil surge en el grupo de ensombrecimiento en el que el EC principal es apareado con un EI como parte de estímulo compuesto asociado con el EC principal, pero ocurrirá una buena RC cuando solo el EC, sea apareado con el EI. La presencia de un EC accesorio disminuye la cantidad de RC que se obtiene ante el EC principal. El ensombrecimiento es más fácilmente conseguido cuando el EC secundario es muy saliente, dado que se dará un mayor aprendizaje para el EC más saliente del compuesto, a costa de la menor saliencia del EC principal.

Aprendizaje y Desarrollo de Preferencias

El ensombrecimiento produce una débil RC, incluso aunque el EC se presente en una excelente relación de contigüidad en el EI; esto es, a pesar de una buena contigüidad, una débil asociación EC-EI va a tener lugar. Otro fenómeno que muestra que la contigüidad no es suficiente para el aprendizaje como ya mencionamos es el bloqueo.

Hubo un enorme interés en el condicionamiento clásico de los estímulos compuestos al final de la década de los sesenta y durante los años setenta, incluyendo los trabajos sobre ensombrecimiento, bloqueo e inhibición condicionada. Esta investigación contribuyó al desarrollo del modelo de condicionamiento de Rescorla y Wagner (1972). Su modelo de "competencia asociativa" establece que los estímulos condicionados compiten por una limitada cantidad de fuerza asociativa que puede ser soportada por el EI. La aplicación de este modelo al ensombrecimiento es obvia, ya que el apareamiento de dos EC con un EI en un ensayo de condicionamiento, facilitará una ocasión típica para la competencia entre estímulos condicionados. El modelo estipula que el EC de mayor saliencia ganará más fuerza asociativa que el EC más débil. Esta es una previsión consistente con los resultados del ensombrecimiento vistos anteriormente. También el bloqueo es fácilmente predecible por el modelo de Rescorla y Wagner. El EC previamente condicionado ya ha adquirido la totalidad de la fuerza asociativa que el EI puede soportar y de esta forma hay muy poca fuerza asociativa disponible para ser adquirida por el EC nuevo.

EXPLICACION DEL MODELO DE RESCORLA Y WAGNER.

La intuición básica que sustenta el modelo de Rescorla y Wagner es expresar que los reforzadores esperados son menos efectivos que los reforzadores sorprendidos. El mismísimo EI puede ser diferencialmente efectivo al producir condicionamiento, dependiendo de la fuerza para la cual esta ocurrencia sea bien predicha por el estímulo antecedente disponible. Hull (1943) sugirió (aunque en diferente lenguaje) para que la fuerza de un EC fuera bien asociado con un EI, cada subsecuente apareo EC-EI produciría un pequeño incremento en el valor condicionado (produciendo una curva de aprendizaje negativamente acelerada), originalmente Hull había propuesto un incremento similar para la fuerza de hábito (H_i) que en su sistema era la variable

Aprendizaje y Desarrollo de Preferencias

intermedia responsable del aprendizaje y consistente con la asociación E-R propiamente dicho. La fuerza de hábito según Hull, determinaba la conducta en combinación multiplicativa con otras variables como impulso (De esta teoría fue en apreciable medida derivado el modelo de Bush y Mosteller, 1955). La modificación importante de esta idea, sugerida por Wagner (1969) e implementada en el modelo de Rescorla y Wagner, fue que cambios en el valor condicionado de un EC, x , resultaban de un apareo x -EI, son no solo debido a la extensión del aprendizaje anterior acerca de x , sino son también debido a la extensión del aprendizaje anterior acerca de la complejidad entera de estímulos en la cual x , fue incorporada. Si el EI es esperado sobre la base de que ha sido aprendido acerca de x , o sobre la base de que ha sido aprendido algo acerca de los otros estímulos ocurridos simultáneamente con x , las consecuencias para un aprendizaje nuevo o adicional acerca de x , son las mismas.

Más formalmente, el modelo propone que cambios en el valor asociativo de un estímulo x , pueden ser representado como:

$$\Delta V_x = \alpha_x \beta [\lambda - V]$$

Donde V se representa el total de fuerza asociativa de todos los estímulos presentes. Esto es calculado por una simple suma de las fuerzas de los componentes, de modo que la fuerza total de un componente x , sería $V_x + V_x * \lambda$ representa el nivel asintótico de la fuerza asociativa que el EI podría soportar con $\lambda = 0$ representando no reforzamiento. En la fórmula se puede ver como que esta no es la fuerza absoluta de el EI λ que determina cambios en V_x , sino más bien la discrepancia entre el EI $(\lambda - V)$ obtenido y anticipado, que determina dichos cambios. Alfa (α) y Beta (β) son parámetros que influyen en la tasa de cambio, y pueden tomar valores entre 1 y 0. Alfa representa la saliencia de un componente individual EC (en la ecuación mencionada, la saliencia de x). Así, estímulos diferentes adquirirán fuerzas asociativas a diferentes tasas a pesar de reforzamientos iguales si ellos difieren en alfa. Beta representa la saliencia de el EI, permitiendo así, que la tasa de aprendizaje sea influenciada por la naturaleza del EI.

Aprendizaje y Desarrollo de Preferencias

Nótese que V puede tomar valores negativos así como positivos. Los positivos corresponden a una excitación condicionada mientras que valores negativos corresponden a inhibición condicionada. Notesé también que V es en principio ilimitada; sin embargo, en la práctica, aplicaciones repetidas de la discrepancia $(\lambda - V)$ tienden a forzar V hacia λ .

En resumen, la fuerza asociativa adquirida por un estímulo en un ensayo n es igual al producto de la constante de velocidad de aprendizaje por la diferencia entre la asíntota del EI y la fuerza asociativa ya poseída por el estímulo. Aplicando la fórmula puede observarse que la cantidad de fuerza asociativa adquirida por el EC va reduciéndose a medida que se suceden los ensayos de condicionamiento. Esta reducción progresiva de la adquisición de fuerza asociativa, predicha por la teoría de Rescorla y Wagner, es coherente con la observación usual de que las curvas que representan la adquisición de la RC muestran una aceleración negativa. Por otra parte, tal reducción es el reflejo de la disminución progresiva de la diferencia $\lambda - V$, es decir la disminución de la sorpresividad del EI durante sus emparejamientos con el EC. Como se ha dicho, el postulado fundamental de la teoría de Rescorla y Wagner es que la cantidad de fuerza asociativa adquirida por un estímulo, debido a su emparejamiento con un EI, es directamente proporcional a la sorpresividad del EI.

De tal manera el modelo de Rescorla y Wagner es un modelo de adquisición basado en la contigüidad, que ha mostrado una importante influencia para examinar los cambios en la fuerza asociativa del EC. Ha servido para organizar un abundante cuerpo de investigación experimental del aprendizaje animal y es una referencia obligada en esta importante área de estudio.

El problema que puede encontrarse en este modelo de adquisición es que hasta ahora ha mostrado poca generalidad más allá del condicionamiento animal (Gluck y Bower, 1988; Wasserman, 1990) y anteriormente Thomas (1983) ha señalado que el modelo de Rescorla y Wagner es un modelo que exige un análisis molecular del condicionamiento, puesto que describe los cambios en la fuerza asociativa del EC ensayo a ensayo. Como consecuencia, tiene problemas de aplicación en las situaciones en que el aprendizaje no ocurre ensayo a ensayo y requieren de un análisis más molar. (por ejemplo, la relación entre la tasa de respuestas y de reforzamiento en condicionamiento operante). Quizás esa sea una de las razones de los pocos estudios que se han

Aprendizaje y Desarrollo de Preferencias

realizado bajo el modelo de Rescorla y Wagner en el aprendizaje humano (Dickinson y Shanks, 1985; Wasserman, 1990).

PROBLEMAS CON EL MODELO DE RESCORLA-WAGNER

La teoría de Rescorla-Wagner explica una amplia variedad de datos experimentales, pero hay algunos de ellos que no explica. Un fenómeno que resulta difícil para la teoría, tiene que ver con el efecto de pre-exposición de estímulos. Si un organismo es expuesto a el EC un número de veces antes de comenzar un experimento y entonces se dan los apareamientos EC-EI en el experimento, la tasa de condicionamiento se ve afectada. (Reiss y Wagner, 1972). El fenómeno de pre-exponer levemente el estímulo condicionado es llamado **inhibición latente**. El organismo ha sido expuesto a un número de ocurrencias sin que suceda algo y tiene un sesgo a mantener una creencia en la ineficiencia del EC después de que los ensayos de condicionamiento han comenzado. Sin embargo, el modelo de Rescorla y Wagner no hace predicciones para dicho efecto. La fuerza de EC-EI inicia en cero, y con ninguna presentación de EI en los ensayos pre-exposición, éste último permanece en cero. De este modo la teoría de Rescorla y Wagner predice que condicionamiento seguiría como si no hubiera habido ensayos de pre-exposición. Y se ha sugerido, que la tasa de aprendizaje α , refleja la saliencia del EC y que el efecto de pre-exposición al EC sería a hacerlo menos saliente. Es decir supongamos que la saliencia de un EC puede cambiar como un resultado de la experiencia del sujeto con el EC. Recordemos que el parámetro α representa la saliencia del EC. En su descripción inicial del modelo, Rescorla y Wagner (op. cit.) está implícito que el valor de α es una propiedad fija de un EC, así como de las capacidades sensitivas del sujeto. Sin embargo, el efecto de inhibición latente sugiere que alguna alteración de estas suposiciones, es en el orden: Asumiendo que la estructura general del Modelo de Rescorla y Wagner es correcta, los datos sugieren que α decrementa cuando cuando un EC es presentado repetidamente sin consecuencia. Se puede cambiar fácilmente el valor de α en una ecuación matemática, pero ¿qué significa hacer dicho cambio desde una perspectiva psicológica? Algunos teóricos del aprendizaje han propuesto que este cambio consiste de un decremento en la atención del organismo al estímulo i . Ellos proponen que durante los ensayos de pre-exposición al EC el sujeto aprende que es en vano poner atención a ese estímulo, de tal modo

Aprendizaje y Desarrollo de Preferencias

que cuando el estímulo es apareado más tarde con un EI le toma más tiempo para desarrollar una asociación.

Otros modelos recientes de condicionamiento clásico explican adecuadamente los fenómenos del condicionamiento de estímulos condicionados compuestos, como los anteriormente descritos. Dichos modelos, incluyendo el de Pearce y Hall y el de Mackintosh son también modelos de "competencia asociativa", pero centran su interés en la efectividad del EC (la atención al EC o la asociabilidad de los estímulos compuestos) para formar parte de una asociación, más que en la mera limitación de la cantidad de aprendizaje que soporta el EI (Mackintosh, 1975; Pearce y Hall, 1980).

Comprensión Neuronal: La Regla Delta

La teoría de Rescorla-Wagner corresponde a una idea actual acerca de como el aprendizaje tomó lugar a un nivel neuronal, lo cual ha jugado un importante papel, particularmente en una teoría de procesamiento neuronal llamada **conexionismo**. Esta teoría enfatizó la importancia de las conexiones sinápticas entre neuronas. Así una activación neuronal se da a través de un grupo de neuronas de entrada, cada una de las cuales hace sinapsis con un grupo de neuronas de salida. El patrón de conexiones se presenta en una situación más compleja que la encontrada en la *Aplysia*, donde una neurona (neuromotora) se hace más activa cuando otra neurona (una neurona sensorial) esta en actividad. La tarea de aprendizaje para cierta red neuronal es agrupar las fuerzas de asociación entre las neuronas de tal manera que cuando un patrón particular de activación ocurra sobre las neuronas de entrada, un patrón de activación deseado ocurra en las neuronas de salida. Hay una regla conexionista de aprendizaje, llamada **la regla delta**, la cual esta basada en la ecuación de Rescorla-Wagner, que explica como se da este aprendizaje neuronal.

En el contexto de modelamiento neuronal, la ecuación de Rescorla-Wagner, o la regla delta, es tomada como una regla para ajustar la fuerza de la conexión sináptica entre la neurona de entrada i y una neurona de salida j . La regla delta es formulada como:

Aprendizaje y Desarrollo de Preferencias

$$\Delta A_{ij} = \alpha A_i (T_j - A_j)$$

Donde ΔA_{ij} es el cambio en la fuerza de la conexión sináptica entre la entrada i y la salida j ; A_i es el nivel de activación de una neurona de entrada i ; A_j es el nivel de activación de la neurona de salida j ; y T_j es la actividad deseada o marcada de j . Esta ecuación puede ser comparada con la ecuación de Rescorla y Wagner.

$$\Delta V_x = \alpha_x \beta [\lambda - V]$$

Donde ΔV_x corresponde a ΔA_{ij} ; $\alpha_x \beta$ corresponde a αA_i ; λ corresponde a T_j ; y V corresponde a A_j . Como en la teoría de Rescorla y Wagner, el aprendizaje es proporcional a la diferencia $T_j - A_j$. De acuerdo a la regla delta, el aprendizaje es también proporcional a A_i , el nivel de activación de la neurona de salida. Recordemos que un propósito para la extensión de la teoría de Rescorla y Wagner fue hacer proporcional el aprendizaje a la saliencia del estímulo.

La regla delta se ha convertido en un importante constructo en teoría de aprendizaje neuronal. Ha sido usada en una amplia variedad de modelos de procesamiento neuronal y ha sido importada por las ciencias de la computación para construir modelos de máquinas de aprendizaje. También ha sido usada para predecir aprendizaje humano complejo Gluck y Bower (1988).

Aprendizaje y Desarrollo de Preferencias

REFERENCIAS

- Bailey, J.T., y Mazur, E.J. (1990). Choice behavior in transition development of preference for the higher probability of reinforcement. *Journal of the Experimental Analysis of Behavior*, **3**, 409-422.
- Baum, W. M., y Rachlin, H.C. (1969). Choice as time allocation. *Journal of the Experimental Analysis of Behavior*, **12**, 861-874.
- Baum, W.M. (1974). On two types of deviation for the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior*, **22**, 231-241.
- Baum, W. M. (1979). Matchingn, undermatching and overmatching in studies choice *Journal of the Experimental Analysis of Behavior*, **32**, 269-281.
- Belke. T.W. (1992). Stimulus preference and the transitivity of preference. *Animal Learning y Behavior* **20**, 401-406.
- Bouzas, A., Morán, C., y Vázquez, F. (1992). Modelos de equilibrio de la asignación de respuestas: Elección Intertemporal. En V. Colotla (Ed.) *La Investigación del Comportamiento en México*. México, D.F. CONACYT-UNAM.
- Bush, R.R. y Mosteller, F. (1955). *Stochastic Models for learning*. New York: Wiley.
- Catania, A.C. (1969). Concurrent performances: Inhibition of one response by reinforcement of another. *Journal of the Experimental Analysis of Behavior*, **12**, 731-744.
- Coobs, H.C., Dawes, M.R. y Tversky, A. (1981) *Introducción a la Psicología Matemática*, Madrid España, Aliana Editorial. pp. 321-383.

Aprendizaje y Desarrollo de Preferencias

- Couvillon, P.A., y Bitterman, M.E. (1985). Analysis of choice in honeybees. *Animal Learning y Behavior* **13**, 246-252.
- Couvillon, P.A., y Bitterman, M.E. (1988). Compound-component and conditional discrimination of colors and odors by honeybees: Further tests of a continuity model. *Animal Learning y Behavior* **16**, 67-74.
- Couvillon, P.A., y Bitterman, M.E. (1991). How honeybees make choices. En L. J. Goodman y R. C. Fisher (Eds.), *The behaviour and physiology of bees* (pp. 116-130). Wallingford, U.K.: CAB International.
- Churchland, P.S. & Sejnowski, T.J.(1992). *The Computacional Brain*. The MIT Press, Cambridge, Massachusetts. London England.
- Daly, H.B. (1985). Observing response acquisition: Preference for unpredictable appetitive rewards obtained under conditions predicted by DMOD. *Journal of Experimental Psychology: Animal Behavior Processes*, **11**, 294-316.
- Davison, M., y McCarthy, D. (1988). *The Matching Law: A research review*. Hillsdale, NJ:Erlbaum.
- Davis, D.G.S., Staddon, J.E.R., Machado, A., y Palmer, R.G. (1993). The Process of recurrent choice. *Psychological Review*, **100**, 320-341.
- Dickinson, A. y Shank, D. (1985) Animal conditioning and human causality judgement. En L.G. Nilsson y T. Archer (Eds.), *Perspectives on Learning and Memory* (pp. 167.191) London: Lawrence Erlbaum Associates.
- Dow, S.M., y Lea, S.E.G. (1987). Foraging in a changing enviroment: Simulation in the operant laboratory. En M.M.L. Commons, A. Kacelnik, y S.J. Shettleworth (Eds.), *Quantitative Analyses of Behavior VI: Foraging* (pp. 90-113). Hillsdale, NJ:Erlbaum.

Aprendizaje y Desarrollo de Preferencias

- Domjan, M. y Burkhard, B. (1990). *Principios de aprendizaje y Conducta*. Madrid. Ed. Debate.
- Dreyfus, L.R. (1991). Local shifts in relative reinforcement rate and time allocation on concurrent schedules. *Journal of Experimental Psychology: Animal Behavior Processes*, **17**, 486-502.
- Fleshler, M. y Hoffman, H.S. (1962) A progression for generating variable-interval schedules. *Journal of the Experimental Analysis of Behavior*, **5**, 529-530.
- Gallistel, C.R. (1990). *The Organization of Learning*. Cambridge: MIT Press.
- Gibbon, J., Church, R., Fairhurst, S., y Kacelnik, A. (1988). Scalar expectancy theory and choice between delayed rewards. *Psychological Review*, **95**, 102-114.
- Gibbon, J. (1995). Dynamics of time matching: Arousal makes better seem worse. *Psychonomic Bulletin & Review*, **2(2)**, 208-215.
- Gluck, M.A., y Bower, G.H. (1988). From conditioning to category learning: An adaptive Network model. *Journal of Experimental Psychology: General*, **117**, 227-247.
- Herrnstein, R.J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, **4**, 267-272.
- Herrnstein, R.J., y Hines, P.N. (1966) Negative reinforcement as shock frequency reduction. *Journal of the Experimental Analysis of Behavior*, **9**, 421-430.
- Herrnstein, R.J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, **13**, 243-266.
- Herrnstein, R.J., y Loveland, D. (1975). Maximizing and matching on concurrent ratio schedules. *Journal of the Experimental Analysis of Behavior*, **24**, 243-266.

Aprendizaje y Desarrollo de Preferencias

- Herrnstein, R.J., y Vaughan, W.J. (1980). Melioration and behavioral allocation. En J.E.R. Staddon (Ed.) *Limits to action: The allocation of individual behavior* (pp. 143-176). New York: Academic Press.
- Hinson, J.M. y Staddon, J.E.R. (1983). Matching, maximizing, and hill climbing. *Journal of the Experimental Analysis of Behavior*, **40**, 321-331.
- Honig, W.K. y Staddon, J.E.R. (1977). *Handbook of operant behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- Horner, J.M., y Staddon, J.E.R. (1987). Probabilistic choice: A simple invariance. *Behavioral Processes*, **15**, 59-92.
- Hull, C. (1943). *Principles of Behavior*. New York, Appleton-Century-Crofts.
- Kacelnik, A., Krebs, J., y Ens, B. (1987). Foraging in a changing environment: An experiment with starlings. En M.M.L. Commons, A. Kacelnik, y S.J. Shettleworth (Eds.), *Quantitative Analyses of Behavior VI: Foraging* (pp. 63-87). Hillsdale, NJ: Erlbaum.
- Kagel, J.H. (1987). Economics according to the rats (and pigeons too): What have we learned and what can we hope to learn? En Roth (Ed.), *Laboratory experimentation in economics*. Cambridge University Press.
- Kamin, L.J. (1969). Predictability, surprise, attention, and conditioning. En B.A. Campbell y R.M. Church (Eds.), *Punishment and aversive behavior*. New York: Appleton-Century-Crofts.
- Killen, P. (1972). A yoked-chamber comparison of concurrent and multiple schedules. *Journal of the Experimental Analysis of Behavior*, **18**, 13-22.
- Killen, P. (1992). Mechanics of the animate. *Journal of the Experimental Analysis of Behavior*, **57**, 429-463.

Aprendizaje y Desarrollo de Preferencias

- Killen, P. (1994). Mathematical principles of reinforcement. *Behavioral and Brain Sciences*, **17**, 105-172.
- Lachman, R., Lachman J.L. , y Butterfield, C. (1979) *Cognitive Psychology and Information Processing: An Introduction*. New Jersey. Lawrence Erlbaum Associates, Publishers, Hillsdale,
- Mackintosh, N.J. (1975). A Theory of attention. *Psychological Review*. **82**, 276-298.
- Mackintosh, N.J. (1977). *The Psychology of animal learning*. London: Academic Press.
- Mark, T.A., y Gallistel, C.R. (1994). Kinetics of Matching. *Journal of Experimental Psychology: Animal Behavior Processes*, **20**, 79-95.
- Marr, D. (1969). A theory of cerebellar cortex. *Journal of Physiology* (London) **202**, 437-470
Reprinted in Vaina (1991) 11-46.
- Marr, D. (1970). A theory for neocerebellar cortex. Proceedings of the Royal Society of London B. **176**, 161-234. Reprinted in Vaina (1991) 59-117.
- Marr, D.(1982). *Vision*. New York: Freeman.
- Marr, D. & T. Poggio. (1976). Co-operative computation of stereo disparity. *Science* **194**: 283-287. Reprinted in Vaina (1991) 239-244.
- Marr, D. & T. Poggio. (1977). From understanding computation todo understanding neural circuitry. *Neuroscience Research Program Bulletin* **15**: 470-488.
- Maxwell S.E. y Delaney H.D. (1990) *Designing Experiments and Analyzing Data: A model Comparison Perspective*. California, U.S.A. Wadsworth Publishing Company. pp. 612-691.

Aprendizaje y Desarrollo de Preferencias

- Mazur, J.E. (1996). Past Experience, recency, and spontaneous recovery in choice behavior. *Animal Learning & Behavior* **24(1)** , 1-10.
- Mazur, J.E. (1995). Development of preference and spontaneous recovery in choice behavior with concurrent variable-interval schedules. *Animal Learning & Behavior* **24(1)** , pp. 93-103
- Mazur, J.E. (1992). Choice behavior in transition: Development of preference with ratio and interval schedules. *Journal of Experimental Psychology: Animal Behavior Processes*, **18** , 364-378.
- Mazur, J.E., y Ratti, T.A. (1991). Choice behavior in transition: Development of preference in a free-operant procedure. *Animal Learning y Behavior*, **19**, 241-248.
- McClelland, J. y Rumelhart, D. (1988). Explorations in Parallel Distributed Processing. Cambridge, MA: MIT Press.
- McLean, A.P. y White, K.G. (1981). Undermatching and contrast within components of multiple schedules. *Journal of the Experimental Analysis of Behavior*, **39**, 405-426.
- Myerson, J., y Miezin, F.M. (1980). The kinetics of choice: An operant systems analysis. *Psychological Review*, **87**, 160-174.
- Myerson, J., y Hale, S. (1988). Choice in transition: A comparison of melioration and the kinetic model. *Journal of the Experimental Analysis of Behavior*, **49**, 291-302.
- Papini, M.R. y Bitterman, M.E. (1990) The role of contingency in classical conditioning. *Psychological Review*, **97**, 396-403.
- Plaud, J.J., y Vogeltanz, N. (1991) Behavior therapy: Lost ties to animal research?. *The Behavior Therapist*, **1**, 89-93.

Aprendizaje y Desarrollo de Preferencias

- Pearce, J.M. y Hall, G. (1980). A model for Pavlovian learning. *Psychological Review*, **87**, 532-552.
- Rachlin, H. (1973). Contrast and Matching. *Psychological Review*, **80**, 217-234.
- Reichardt, W., y Poggio, T. (1976) Visual control of orientation behavior in the fly. Part I. A quantitative analysis. *Quarterly Review of Biophysics*, **9**, 311-375.
- Reiss, S. y Wagner, A.R. (1972). CS habituation produces a "latent inhibition" effect but no active conditioned inhibition. *Learning and Motivation*, **15**, 237-245.
- Rescorla, R.A. (1972). Informational variables in Pavlovian conditioning. En G. Bower (Ed.), *The Psychology of Learning and motivation*. New York: Academic Press.
- Rescorla, R.A., y Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. En A.Black y W. Prokasy (Eds.), *Classical conditioning II*. New York: Appleton
- Reynolds, S. G. (1963). On some determinants of choice in pigeons. *Journal of the Experimental Analysis of Behavior*, **6**, 53-59.
- Sejnowski, T. J., Koch, C., y Churchland, P.S. (1988) Computational neuroscience. *Science*, **241**, 1299-1306.
- Silberberg, A., Hamilton, B., Zirix, J., y Casey, J. (1978). The structure of choice. *Journal of Experimental Psychology: Animal Behavior Processes*, **4**, 368-398.
- Staddon, J.E.R. (1982). Behavioral competition, contrast and matching. En: M.L. Commons, R.J. Herrnstein, y H. Rachlin (Eds.) *Quantitative analysis of Operant Behavior. Vol. 2. Matching and maximizing accounts*. Cambridge, M. A. Bellinger.

Aprendizaje y Desarrollo de Preferencias

- Staddon, J.E.R. (1988). Quasi-dynamic choice models: Melioration and ratio invariance. *Journal of the Experimental Analysis of Behavior*, **49**, 303-320.
- Staddon, J.E.R., y Horner, J.M. (1989). Stochastic choice models: A comparison between Bush-Mosteller and a source-independent reward-following model. *Journal of the Experimental Analysis of Behavior*, **52**, 57-64.
- Stenberg S. (1963). Stochastic Learning Theory. En R. D. Luce, R. R. Bush y E. Galanter (Eds.) *Handbook of Mathematical Psychology Vol II cap 9*. (pp. 1-120) Editado por. U.S.A. John Wiley and Sons Inc.
- Stephens, D.W., y Krebs, J.R. (1986). *Foraging theory*. Princeton University Press.
- Thomas, G. V. (1983). Contiguity and Contingency in Instrumental conditioning. *Learning and Motivation*, **14**, 513-526.
- Vaughan, W., Jr. (1985). Choice: A local analysis. *Journal of the Experimental Analysis of Behavior*, **43**, 383-405.
- Vaughan, W., Jr., y Herrnstein, R.J. (1987). Stability, melioration, and natural selection. En L.Green y J.H. Kagel (Eds.), *Advances in behavioral economics* (Vol. 1, pp. 185-215). Norwood, NJ: Ablex.
- Vázquez, F. (1992). *Distribución de Respuestas en un programa múltiple estocástico IV-IV: Igualación y Contraste*. Tesis de Maestría en Análisis Experimental de la Conducta. Facultad de Psicología. UNAM.
- Wasserman, E.A. (1990). Attribution of causality to common and distinctive elements of compound stimuli. *Psychological Science*, **1**, 298-302.

Aprendizaje y Desarrollo de Preferencias

- Williams, B.A. (1983). Another look at contrast in múltiple schedules. *Journal of the Experimental Analysis of Behavior*, **39**, 345-384.
- Williams, B.A. (1988). Reinforcement, choice, and response strength. En: *Steven's handbook of experimental psychology. Vol. 2: Learning and Cognition*, 2nd ed., Eds. R.C. Atkinson, R.J. Herrnstein, G. Lindzey y R.D. Luce. Wiley.
- Williams, B.A. (1994a). Reinforcement and Choice. En *Animal Learning and Cognition.*, Ed. N. J. Mackintosh. Academic Press.
- Williams, B.A. (1994b). The Role of probability of reinforcement in models of choice. *Psychological Review*, **101**, 704-707.