

218
2 es.



**UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO**

FACULTAD DE PSICOLOGIA

**DINAMICA DE LA ADAPTACION A
PROGRAMAS MULTIPLES**

T E S I S
QUE PARA OBTENER EL TITULO DE:
LICENCIADA EN PSICOLOGIA
P R E S E N T A
DIANA POSADAS SANCHEZ

DIRECTOR DE TESIS

DR. ARTURO BOUZAS RIAÑO



MEXICO, D. F.

1998

**TESIS CON
FALLA DE ORIGEN**

259634



Universidad Nacional
Autónoma de México



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.

A ti Señor,
Por todas las experiencias que han conformado mi vida y que han labrado lo que soy ahora. Por estar siempre a mi lado, y brindarme más de lo que había imaginado. Por la capacidad de emocionarme ante los acontecimientos sencillos de la vida.
Por tu maravilloso regalo, y por permitirme iniciar el día de hoy este camino que realmente amo. ¡Muchas gracias!

A mi padre,
Porque me enseñaste que solo triunfa aquél que hace de su vida una constante lucha.
Por heredarme esa inquietud de conocer y explicar el mundo.
Porque me has enseñado a no cometer tus errores, invitándome a seguir tu camino de aciertos.

A mi madre,
Por tu alegría y tus cuidados. Por ser ejemplo vivo de compromiso, responsabilidad y entrega.
Porque me enseñaste a aceptar mis derrotas con la cabeza en alto, y levantarme una y otra vez sin perder de vista el camino.

A Rossy,
Porque desde pequeñas me enseñaste a hacer todo cada vez mejor.
Por los ideales, el entusiasmo y cada uno de los momentos compartidos.
Porque siempre supe que tendría abiertas las puertas de tu corazón.

Yo también los amo...

Agradezco a la Universidad Nacional Autónoma de México, por el privilegio y la responsabilidad de convertirme en universitaria. Porque sé que el compromiso que adquirí ha de reflejarse en mi participación decidida en las actividades que se desarrollan en su interior y que ejercen un impacto en el mundo exterior.

Por todos los recursos que estuvieron a mi alcance; por mi formación científica y mi desarrollo como persona, porque en su escenario viví las experiencias más enriquecedoras, importantes e inolvidables de mi vida. ¡Muchas gracias!

Agradezco a la Facultad de Psicología, por todas las oportunidades que tuve durante mi camino por la Universidad; por los maestros que afectaron mi vida académica y que me transmitieron ese conocimiento que traspasa las aulas.

Al Dr. Arturo Bouzas,

Porque nunca existieron límites para el aprendizaje y siempre hubo nuevos retos que superar. Porque los consejos y las críticas a mi trabajo me han ayudado a crecer y enfrentar nuevas situaciones, ya que me proporcionaron herramientas invaluable que constituyen la base de mi vida académica y que hoy me abren las puertas. Por el apoyo incondicional, la confianza que siempre he recibido, las constantes enseñanzas y ese gran ejemplo a seguir.

Por estar tan involucrado en cada idea, en cada paso, en cada resultado, al contemplar con gran entusiasmo los movimientos del contador de respuestas. Gracias por enseñarme el extraordinario y fascinante camino del Análisis Experimental de la Conducta, y recorrer conmigo parte de él.

Al grupo de Comportamiento y Adaptación,

Por lo que cada uno de ustedes me regaló durante la convivencia diaria en nuestro laboratorio.

Al Dr. John Staddon,

Muchas gracias por brindarme la oportunidad de incorporarme a su grupo de trabajo y hacer uno de mis sueños realidad. Por su tiempo y por intervenir en forma tan importante en mi desarrollo en el mundo de la Investigación. Mi más profundo agradecimiento, respeto y admiración.

Al grupo de Aprendizaje y Conducta,
Por las vigorosas discusiones a las versiones previas de este trabajo;
especialmente a José Bentes, Valentin Dragoi y Silvano Zanutto, por aquellas
interminables discusiones en las que la ciencia y el mundo de la Psicología
siempre fueron protagonistas. Por su amistad, consejos y ayuda durante mi
estancia en Duke.

¡Obrigado, Multȃmesc, Gracias!

Gracias al Mtro. Gustavo Bacha, al Dr. Florente López, al Dr. David
Velázquez y al Mtro. Fernando Vázquez, por la revisión de este trabajo. Por
compartir sus conocimientos y su tiempo.

A cada uno de ustedes,
Que con su granito de arena hicieron posible que este proyecto se convirtiera
en el principio de una hermosa realidad. Por su trabajo, ayuda, paciencia,
dedicación y amistad.

Quiero que sepan que tengo perfectamente bien establecidas las conexiones
en mi memoria, y que aunque el espacio sea pequeño, atesoro y agradezco lo
que cada uno hizo por mí durante este proceso. ¡Muchas gracias!

A Constanza Alcaráz, Alicia Lara, Sheyla Mejía y Karina Simón, por todas
las aventuras que compartimos en la Facultad de Psicología y que forjaron
una hermosa amistad. Porque a pesar de elegir diferentes caminos, siempre
estuvieron presentes en cada acontecimiento importante.

A Gina Ricks,
Porque siempre estuviste dispuesta a dar ese paso conmigo. Porque hasta en
los pequeños detalles estabas ahí, moviendo mar y tierra para ayudarme,
dando siempre lo mejor de ti. Mil gracias.

A Brenda Mendoza,
Por todos los momentos tan divertidos que vivimos. Por estar siempre cerca,
a pesar de la distancia. Por los obstáculos que siempre logramos superar.
Gracias por compartir en forma plena miedos, problemas, alegrías y sueños.

A Vladimir Orduña,
Por los momentos de trabajo compartidos, porque entre registros y
experimentos creció una gran amistad. Gracias por tu confianza y por
permitirme conocerte.

A Oscar Zamora,
Por tu valiosa ayuda, porque resolviste las 958347563974515874968 dudas que tuve. Por ceder parte de tu tiempo, ya que siempre conté contigo cuando mandaba alguna señal de S.O.S. ¡Muchas gracias!

A mis abuelos Enrique y Roberto,
Por transmitirme el deseo entusiasta de vivir.

A mi querida Lidia,
Por tus oraciones, tu amor y tu gran ejemplo de mujer.

A mi Abis,
Gracias por tu amor e inmensa confianza. Por todo lo que hemos compartido, por tus apapachos y consentimientos. Por tu compañía y tu interés en cada uno de mis logros.

A mis tíos y primos,
Por las sonrisas, ayuda, consejos y cariño que recibí de cada uno.

A mi tía Rebe,
Por tu gran cariño y tu hermosa presencia en mi vida.

A mi tío Jorge,
Por tu gran interés en mis actividades. Porque siempre supe que podía contar contigo, ya que tu ayuda nunca se dejó esperar. Por tu cariño. Gracias.

A mi tía Betty,
Porque con tu ejemplo siempre demostraste lo hermoso que es ayudar a otros. Gracias por tu confianza.

A Paco, Moni, Gabriel y Güere,
Por una infancia llena de aventuras, travesuras y demás locuras. Por todos los momentos tan divertidos que hemos compartido. Por ser cómplices en todo y siempre recibirme con los brazos abiertos.

¿Cuándo un descubrimiento científico ha hecho las cosas más fáciles? Puede clarificar una confusión previa o simplificar una dificultad anterior, pero siempre deja nuevos problemas, los cuales son más oscuros, más difíciles y ¡más interesantes! Use su imaginación, hombre. Mire lo que falta por hacer.

B. F. Skinner
Walden Dos. 1948

Presidente	Dr. Florente López Rodríguez
Vocal	Dr. Arturo Bouzas Riaño
Secretario	Dr. David N. Velázquez Martínez
1er. Suplente	Mtro. Gustavo Bacha Méndez
2do. Suplente	Mtro. Fernando Vázquez Pineda

Sitio donde se desarrolló el tema: Laboratorio de Comportamiento y Adaptación
Facultad de Psicología
Universidad Nacional Autónoma de México.

Learning and Adaptive Behavior
Laboratory
Department of Psychology: Experimental
Duke University.

Estoy en deuda con el Dr. Arturo Bouzas, por su insustituible guía durante la preparación y el desarrollo de este trabajo; con el Mtro. Oscar Zamora, por su gran apoyo en la programación de los experimentos y numerosos detalles operativos; con Vladimir Orduña y Fusae Nakazawa, por su ayuda en la conducción y recolección de datos del experimento 3, así como sus comentarios durante el desarrollo del mismo.

El presente trabajo se realizó con el apoyo del Consejo Nacional de Ciencia y Tecnología, proyecto 5-4676H.

ÍNDICE

Resumen	1
----------------	---

Capítulo I. Antecedentes

• Elección y la Ley de Igualación	4
• Estudios de Elección en Periodos de Transición	7

Capítulo II. Principios Teóricos del Aprendizaje Operante y la Elección

• Teorías de Fortalecimiento	12
Modelo de Operador Lineal	14
Modelo de Rendimientos Promedio	18
Modelo de Rendimientos Ponderados	23
• Teorías Representacionales	35
Representación de Demoras de Reforzamiento	36
Representación de Tasas de Ocurrencia de Sucesos	46

Capítulo III. Adaptaciones Conductuales en Programas de Reforzamiento Múltiples

• Programas de Reforzamiento Múltiples y la Ley de Igualación:	
Interacciones a Distancia	54
• Contraste Conductual	57
• Teoría de Competencia Conductual	59
• Efectos Generales del Contexto de Reforzamiento	61
• Dinámica de la Elección Sucesiva	64

Capítulo IV. Trabajo Experimental

• Método	67
Sujetos	67
Aparato	67
Procedimiento General	68

Experimento 1.

Cambio Alternativo Diario y Dentro de una Misma Sesión Experimental.

Procedimiento	70
Resultados	72
Discusión	80

Experimento 2.

Cambio Aleatorio Diario y a la Segunda Mitad de la Sesión Experimental.

Procedimiento	82
Resultados	83
Discusión	88

Experimento 3.

Transición Probabilística Intrasesión.

Procedimiento	90
Resultados	91
Discusión	96

Discusión General	97
-----------------------------	----

Referencias	103
-----------------------	-----

Apéndice A.	115
---------------------	-----

RESUMEN

Treinta años de investigación en situaciones de elección simultánea y sucesiva han generado un conjunto de regularidades capturadas con éxito por variantes de la ley cuantitativa del efecto. Un resultado fundamental que se ha encontrado en la literatura contemporánea operante, es que las tasas relativas de respuesta igualan las tasas relativas de reforzamiento (Herrnstein, 1970). Una pregunta aún sin respuesta firme es cuál es la dinámica que sigue este ajuste, en particular cuando las tasas de reforzamiento ocurren en diferentes periodos. Hasta muy recientemente poco se conocía acerca de los mecanismos dinámicos que podrían generar tales equilibrios; para estudiarlos, es necesario conocer la trayectoria que sigue la tasa relativa de respuesta ante cambios en la tasa relativa de reforzamiento.

El presente trabajo se concentró en estudiar las adaptaciones conductuales a entornos temporalmente heterogéneos, y tuvo como objetivo evaluar cómo se ajustan las tasas de respuesta a cambios en las condiciones de refuerzo en diferentes periodos, utilizando programas de reforzamiento múltiples. Las respuestas de picar una tecla en uno de los componentes se refuerzan de acuerdo a un programa de razón variable 30, mientras que las respuestas de picar la tecla en el otro componente no reciben reforzamiento (Múlt. RV30-EXT).

Se estudió la dinámica de los ajustes a dos tipos de cambios en programas de reforzamiento. El ajuste a cambios diarios alternos y dentro de una misma sesión experimental, y el ajuste a cambios aleatorios diarios, a la segunda mitad de la sesión, y desde el inicio de la sesión experimental.

Se encontró que en situaciones de elección intertemporal, la conducta de los animales se adapta rápidamente a alternaciones diarias en las condiciones de reforzamiento y a cambios diarios a la mitad de la sesión experimental. Sin embargo, la adaptación es más lenta cuando las condiciones asociadas al refuerzo son totalmente impredecibles. Los resultados se discuten en términos de los cambios en la tasa relativa de respuesta para cada una de las condiciones experimentales.

Los organismos en su ambiente natural con frecuencia dependen de fuentes de comida que son transitorias, y siempre deben de responder a cambios en la calidad y distribución de la comida, la conducta de los depredadores, la localización de conespecíficos, y otros atributos que contribuyen a la calidad del sitio de alimentación. En un extremo, el comportamiento esta completamente determinado por la estimulación actual; en el otro, refleja toda la experiencia del organismo.

Estos cambios pueden suceder en periodos que van desde unos cuantos meses a horas o minutos, y el animal debe ser sensible a dos aspectos de los cambios que son importantes: Si los cambios en el sitio de alimentación ocurren con el paso del tiempo, independientemente de las acciones del animal; o si ocurren independientemente del tiempo, a una tasa determinada por las acciones del animal.

De esta forma, los depredadores deben ajustar su comportamiento ante cambios en su entorno empleando señales medioambientales que predigan la llegada de comida a un sitio particular; pero si no hay señales bien correlacionadas, el depredador puede seguir los cambios muestreando diferentes partes de su hábitat en forma regular. Por ejemplo, para dejar un sitio de alimentación cuya fuente se esta acabando, el animal debe continuar evaluando de alguna forma, su tasa de consumo en el sitio de alimentación actual, y comparar ésta con el resto del medio ambiente. El animal debe explotar el mejor sitio de alimentación sin invertir más tiempo del necesario en recolectar información acerca de la calidad de las alternativas. La información necesaria acerca del resto del medio ambiente incluye recordar tiempos de

viaje, tasas de ganancia en otros sitios de alimentación y probabilidades de encuentro de diferentes tipos de presa.

Asimismo, pocos animales pueden buscar alimento continuamente. El forrajeo puede ser interrumpido por diferentes agentes, tales como la llegada de rivales o depredadores. Entonces, el tiempo disponible en el cual los animales pueden, basados en su experiencia, tomar decisiones acerca de aceptar o rechazar las presas actuales en comparación con presas esperadas en un futuro cercano (el horizonte temporal), es comúnmente limitado.

El problema de cómo los organismos aprenden y se adaptan a cambios rápidos en su entorno, es el objetivo de este estudio.

CAPÍTULO I

ANTECEDENTES

Elección y la Ley de Igualación

Una adaptación fundamental de todos los organismos es la asignación que hacen de sus tiempos y respuestas para obtener, bajo restricciones, los recursos necesarios para vivir. La casi totalidad de los experimentos de elección propuestos asume la presencia de al menos dos teclas (derecha e izquierda) y una variedad de procedimientos en los cuales respuestas repetitivas son reforzadas intermitentemente de acuerdo a programas específicos que han sido diseñados para enfatizar sus efectos adaptativos. El modelo experimental más comúnmente empleado consiste en el estudio de la distribución de dos respuestas, cada una reforzada de acuerdo a programas de reforzamiento independientes. La investigación con este procedimiento, conocido como programa concurrente de reforzamiento, ha obtenido un resultado de gran generalidad: La razón de tasas de respuesta, P_1/P_2 , es aproximadamente igual a la razón de tasas de reforzamiento, R_1/R_2 . A esta relación empírica se le llama Ley de Igualación (Herrnstein, 1961, 1970; Davison & McCarthy, 1988; Williams, 1988).

La ley de igualación ha servido como un marco empírico fértil para el desarrollo de las diferentes teorías de elección. Durante este desarrollo, dos cuestiones interrelacionadas han estado compitiendo continuamente: La regla de elección por la cual la conducta es distribuida entre las alternativas, y los

determinantes de la fuerza de cada alternativa. En principio, estas cuestiones están separadas, pero en la práctica han estado totalmente entrelazadas (Williams, 1994).

Estas cuestiones teóricas surgen porque la ley de igualación es definida en términos de medidas molares de la distribución conductual, típicamente el resultado total de respuestas en una sesión completa. Sin embargo, la sesión es sólo un nivel posible en un continuo de niveles, en los cuales la conducta y los refuerzos pueden ser acumulados para computar el grado en el cual la razón de distribución de tiempo del animal, iguala la razón de refuerzos recibidos (Real & Dreyfus, 1985).

La ley de igualación no plantea nada directamente acerca de las respuestas individuales que constituyen las medidas molares. Dado que la elección obedece a la ley de igualación en conjunto, es necesario determinar qué es lo que controla dichas respuestas.

Se han ofrecido explicaciones económicas (véase Rachlin & Burkhard, 1978, para una revisión) y mecanicistas (Davison & McCarthy, 1988; Staddon, 1988; Williams, 1988, para revisiones) para explicar la ley de igualación y los datos obtenidos en programas de reforzamiento. Las aproximaciones mecanicistas son todas derivadas de la ley del efecto. Como todos los modelos, especifican el estado del sistema y permiten que el estado en un instante determine la respuesta (y el estado) en el siguiente instante. La mayoría de los modelos para la elección recurrente comparte una propiedad simplificada, sin embargo, iguala el estado del sistema programado del sujeto con variables conductuales y ambientales (tales como las tasas de respuesta y reforzamiento o probabilidades) medidas en un breve intervalo (en el orden de minutos). Staddon (1992) llama a esta propiedad *localidad*.

La aproximación más común ha sido asumir que el mejor estimado de la probabilidad local de respuesta son sus probabilidades molares. Igualación es entonces el resultado automático de cada respuesta ocurriendo en proporción a su propia fuerza, sin mayores especificaciones (e.g., Herrnstein, 1970; Myerson & Miezin, 1980).

Una segunda aproximación asume que cada respuesta individual está determinada por las contingencias operando en ese momento, de tal forma que igualación en conjunto es el resultado de ajustes a cambios en las contingencias variables. Es entonces la naturaleza de estos ajustes, en vez de igualación por sí misma, lo que revela los principios fundamentales del comportamiento (e.g., Herrnstein & Vaughan, 1980; Shimp, 1966).

Una tercera aproximación se deriva de la teoría de optimización, la cual considera que igualación es la distribución conductual que maximiza el total de refuerzo en la situación (Rachlin, Battalio, Kagel & Green, 1981). Esta aproximación no hace el intento de explicar los procesos moleculares que generan la maximización molar, sino que proporciona una descripción normativa del patrón óptimo de conducta dado algún grupo de valores para las varias actividades disponibles en el medio ambiente.

La amplitud del marco de referencia temporal o intervalo en el cual se integra el refuerzo es una cuestión fundamental. Las teorías de equilibrio implican que la conducta es controlada por refuerzos agregados en un periodo relativamente largo (e.g., una sesión completa). La suposición común de que las variables que controlan el comportamiento operan en un marco de referencia temporal amplio, está reflejado en la manera en la que los datos son recolectados y agrupados. La relación de igualación (Herrnstein, 1961) ha sido aplicada a diferentes situaciones de elección (véase de Villiers, 1977, para una

revisión), asumiendo el lugar de una ley del comportamiento. Aunque hay un acuerdo general de que algunas formas de la relación de igualación describen adecuadamente la ejecución en programas concurrentes, hay un considerable desacuerdo concerniente a los orígenes de este fenómeno empírico. En el punto de vista molar, la distribución de elecciones está determinada por relaciones a gran escala entre la conducta y sus consecuencias. Respuestas y reforzadores son agrupados en periodos relativamente prolongados, tales como una sesión experimental completa, y es a este nivel que la tasa relativa de reforzamiento es presumiblemente la variable que controla el comportamiento. En contraste, el punto de vista molecular considera estas relaciones a gran escala entre la elección y sus consecuencias, derivadas de propiedades locales de los programas y plantea que igualación es una consecuencia de procesos moleculares más fundamentales. Las posiciones molares y moleculares son con frecuencia entendidas como diferentes niveles de análisis, aunque se ha sugerido que ya que ambas posiciones se centran en respuestas y reforzadores, éstas deberían ser caracterizadas en términos de niveles de agregados, en lugar de niveles de análisis. Implícito en ambas posiciones está la suposición de que las variables que intervienen en el proceso de elección hacen contacto con el responder y operan a niveles particulares de agregación.

Estudios de Elección en Periodos de Transición

De los numerosos estudios de elección que han sido conducidos en las décadas pasadas (estudios usando procedimientos tales como programas concurrentes, concurrentes encadenados y procedimientos de ensayos discretos), la vasta mayoría se ha centrado en la conducta en estado estable. En un experimento típico, un par de programas de reforzamiento pueden

permanecer por 20 ó 30 sesiones, hasta que las elecciones de los sujetos no muestren cambios sistemáticos de sesión a sesión. El resultado de las últimas cinco sesiones, usualmente ha sido considerado como la medida de la ejecución de los sujetos en estado estable con este par de programas de reforzamiento.

Sin embargo, experimentos recientes conducidos en su totalidad en programas concurrentes, han empezado a estudiar relaciones entre la conducta de elección y sus consecuencias en periodos más breves, esto es, en periodos de transición (aquellos periodos en los que al sujeto se le presenta un nuevo par de programas de reforzamiento y su conducta no ha sido estabilizada). En éstos periodos, se analiza el curso que sigue el comportamiento al momento de un cambio en los parámetros de reforzamiento, es decir, el ajuste que sigue el comportamiento antes de llegar a equilibrio (estado estable). Por ejemplo, experimentos en los cuales la tasa relativa de refuerzo se cambió diariamente (Hunter & Davison, 1985) o cada seis días (Davison & Hunter, 1979) muestran que gran parte de los cambios últimos en el comportamiento ocurren dentro de las primeras sesiones siguiendo un cambio en la tasa relativa de refuerzo. Estos resultados pueden sugerir la posibilidad de que la distribución del comportamiento sea sensible a cambios más locales en la distribución de los reforzadores.

Dos han sido las principales transiciones que se han estudiado: 1) Alteraciones en los valores de los programas de refuerzo, y 2) Cambios a extinción.

Un número de evidencia creciente (e.g., Kacelnik, Krebs & Ens, 1987; Mazur, 1992; Davis, Staddon, Machado & Palmer, 1993) sugiere que para entender el proceso específico que subyace la adquisición y extinción en aprendizaje operante, se debe de tener un conocimiento completo de los

patrones dinámicos de la conducta en transición. Este conocimiento permitiría explicar la conducta en una base momento a momento o en una base respuesta a respuesta. De esta forma, sería posible entender la influencia de la historia remota (las condiciones experimentales precedentes a la actual) en conducta operante.

Aunque cada experimento sobre elección debe incluir un periodo de adquisición antes de alcanzar la ejecución en estado estable, la mayoría de los estudios publicados no ha incluido información acerca de esta fase transicional. Excepciones a esta regla incluye estudios que han empleado una variedad de especies y procedimientos experimentales (e.g., Catania, 1969; Couvillon & Bitterman, 1985, 1986; Dreyfus, 1991; Kacelnik et al., 1987; Mark & Gallistel, 1994; Myerson & Hale, 1988).

En contraste con la escasez del trabajo empírico, ha habido un considerable interés en teorías dinámicas y cuasidinámicas de elección, que se refieren al tópico de adquisición. Por varias décadas, los modelos de operadores lineales (Bush & Mosteller, 1955) han sido ampliamente utilizados para describir curvas de aprendizaje, aunque la adecuación de estos modelos ha sido discutida (Mazur & Hastie, 1978; Staddon & Horner, 1989). Muchas variantes del modelo de operador lineal han sido empleadas en modelos de forrajeo óptimo, designados para predecir la adquisición de una preferencia en situaciones de elección (véase Kacelnik et al., 1987, para una revisión; Couvillon & Bitterman, 1985). Otras teorías de adquisición en situaciones de elección incluyen las teorías de mejoramiento (Herrnstein & Vaughan, 1980), el modelo cinético (Myerson & Míezin, 1980), el modelo de invarianza de razón (Staddon, 1988; Staddon & Horner, 1989); y finalmente, modelos desarrollados en la literatura de condicionamiento clásico, tales como el modelo de

modificación de H. B. Daly y J. T. Daly al modelo de Rescorla-Wagner (Daly & Daly, 1982), han sido también aplicados a la adquisición de la conducta de elección, y se discuten en el capítulo 2.

Una dicotomía que emerge en el tópico de elección es entre modelos locales y globales. A simple vista, puede parecer que la discusión sobre esta cuestión es simple. La tendencia actual es hacia modelos globales, que consideran la historia completa de entrenamiento como el determinante de la dinámica del aprendizaje (e.g., el modelo de efectos acumulativos de Davis et al., 1993). De la misma forma, varios modelos de redes neuronales, tales como la aproximación de Schmajuk, Lam y Gray (1996), son de hecho modelos globales, que eventualmente consideran ventanas temporales grandes en las cuales las asociaciones son fortalecidas y el comportamiento actual es distribuido en base a esas asociaciones.

Otra dicotomía es aquella entre teorías de fortalecimiento (Staddon, 1993) y teorías representacionales (Williams, 1994). El primer tipo de teorías (i.e., el modelo de operador lineal, el modelo cinético, y el modelo de efectos acumulativos, entre otros) asume un número de variables (e.g., la probabilidad de reforzamiento, la probabilidad de respuesta, o los reforzadores actuales y las respuestas actuales), que están correlacionadas en una ecuación simple (o quizás dos) cuyo resultado representa un valor de P , y su fuerza es asociada con una respuesta particular. Las suposiciones que conciben la ecuación con el valor de P son en general muy simples e intuitivas, y evitan cualquier estado interno.

El segundo grupo de teorías, las representacionales (Gibbon, 1995; Gibbon, Church, Fairhurst & Kacelnik, 1988; Williams, 1994), no asume un proceso de incremento/decremento, sino que el animal va representando ya sea intervalos entre-reforzadores (Gibbon, 1995; Gibbon et al., 1988), o tasas de

ocurrencia de sucesos (Gallistel, 1990, 1994), y los combinan con alguna regla de decisión; también asume un número de principios neuroconductuales que están ligados por una arquitectura de redes neuronales (Schmajuk et al., 1996). El comportamiento generado por tales representaciones es más complejo y parece explicar mejor los datos, con un número creciente de hipótesis acerca de formaciones y procesos neuronales (e.g., diferentes tipos de memoria, diferentes tipos de circuitos neuronales, etc.) que cooperan para controlar el comportamiento.

CAPÍTULO II

PRINCIPIOS TEÓRICOS DEL APRENDIZAJE OPERANTE Y LA ELECCIÓN

Hay una aleatoriedad intrínseca en el medio ambiente. Varios eventos biológicos importantes ocurren irregularmente, lo que obliga a los animales a poseer mecanismos que les permitan enfrentarse con la ocurrencia irregular de eventos; un mecanismo que haga desaparecer el patrón de irregularidad, y que permita la predicción de la ocurrencia de eventos futuros.

La ley cuantitativa del efecto predice que el comportamiento en programas de reforzamiento será una función de las tasas de reforzamiento obtenidas en los programas (Herrnstein, 1961, 1970). Una pregunta fundamental acerca del comportamiento animal es cómo la información del pasado inmediato es combinada con información de un pasado más distante para explorar cambios medio ambientales; en otras palabras, cómo una serie de eventos es transformada en un estimado de la densidad de reforzamiento disponible en un programa. Todas las explicaciones populares tanto para forrajeo como para el comportamiento en programas de reforzamiento, hace uso de los resultados de dicha integración.

Teorías de Fortalecimiento

Como ya se discutió, una de las familias de modelos de elección esta representada por las teorías de fortalecimiento, que asumen que la fuerza de

una respuesta depende de los episodios de reforzamiento y extinción a los cuales ha sido expuesta.

La literatura en aprendizaje contiene varios intentos de resolver el problema fundamental de la integración en el tiempo (e.g., Horner, Staddon & Lozano, 1997). Lea y Dow (1984) le llaman a la propuesta unánime “el modelo común” de integración e información, que se caracteriza por los siguientes elementos:

1. *Un grupo de alternativas.* Cada alternativa es descrita por una distribución de probabilidad con rendimientos, condicional a su elección.
2. *Un organismo que aprende.* Cada organismo es descrito por un vector de probabilidad inicial en las alternativas y, después de cada ensayo de aprendizaje, por un vector de probabilidad corregido.
3. *Una regla de decisión.* En cada secuencia de ensayos, el organismo escoge una alternativa. La elección esta determinada probabilísticamente por el vector de probabilidad de ese organismo.
4. *Una regla de resultados.* La elección de una alternativa particular se extrae de la distribución de probabilidad de rendimientos para esa alternativa.
5. *Una regla de aprendizaje.* Después de cada ensayo, el organismo modifica el vector de probabilidad. Si $p_i(t)$ es una probabilidad específica del organismo de escoger la alternativa i en el ensayo t , una regla de aprendizaje describe cómo $p_i(t)$ es transformada a $p_i(t+1)$.

Varios modelos se ajustan a esta descripción y difieren solo en las reglas de aprendizaje que asumen. De acuerdo a March (1996), hay tres reglas de aprendizaje que deben ser consideradas. La primera define los modelos de

operador lineal. Esta es una variante de los modelos estocásticos de aprendizaje de las tradiciones Bush-Mosteller-Estes (Bush & Mosteller, 1955; Estes, 1950). La segunda regla de aprendizaje define un modelo de rendimiento promedio, una variante de los modelos de igualación de Herrnstein (Herrnstein, 1961, 1970; Staddon, 1977). La tercera regla de aprendizaje define al modelo de rendimiento ponderado, una versión de la actualización exponencial (Killeen, 1981, 1994).

Aunque las posiciones molares y moleculares pueden ser similares en los datos que emplean, los procesos conductuales que cada posición asume son muy diferentes. El control por relaciones molares implica integración y peso diferencial a eventos discretos en periodos relativamente extensos, mientras que el control por relaciones moleculares requiere exploración momento a momento de los cambios locales en la probabilidad de reforzamiento (véase Shimp, 1966).

Modelo de Operador Lineal

En el modelo de operador lineal, $p_i(t)$ cambia como resultado de una acción realizada y de los resultados observados en $t-1$. Si los resultados son favorables, entonces la probabilidad asociada con la acción realizada incrementa. Si los resultados son desfavorables, la probabilidad asociada con la acción realizada decrementa. En el caso de un refuerzo después de escoger la i -ésima alternativa,

$$p_i(t+1) = p_i(t) + \alpha(1-p_i(t)) \quad (1)$$

$$(0 \leq \alpha \leq 1)$$

Si la i -ésima alternativa es escogida pero no reforzada,

$$p_i(t+1) = (1-\alpha) p_i(t) \quad (2)$$

$$(0 \leq \alpha \leq 1)$$

Mientras mayor sea el valor del parámetro de aprendizaje (α), más rápida es la adaptación. Para permitir refuerzos variables, la regla de aprendizaje se modifica para proporcionar el incremento dependiendo de la magnitud del refuerzo. Si la alternativa i es escogida en el tiempo t y se realiza un refuerzo positivo de k , entonces $p_i(t+1)$ incrementa aplicando el operador de refuerzo k veces a $p_i(t)$. Entonces,

$$p_i(t+1) = 1 - [(1-\alpha)^k (1-p_i(t))] \quad (3)$$

$$(0 \leq \alpha \leq 1)$$

Si se realiza un refuerzo negativo de k , entonces $p_i(t+1)$ decrementa aplicando el operador de no refuerzo k veces a $p_i(t)$. Entonces,

$$p_i(t+1) = p_i(t) (1-\alpha)^k \quad (4)$$

$$(0 \leq \alpha \leq 1)$$

Las probabilidades iniciales pueden asumir cualquier valor que satisfaga los requerimientos que se encuentren entre 0 y 1, y sumen 1 sobre todas las i , pero cada una de las $p_i(0)$ podrían ser normalmente igual a $1/n$, donde n es el número de alternativas. En el presente caso, con dos alternativas, $p_2(t) = 1 - p_1(t)$ y $p_1(0) = 0.5$.

El modelo de operador lineal descrito, se deriva de la tradición Bush-Mosteller-Estes, pero esto no significa que capture todas las variaciones en esta tradición. Por ejemplo, el considerar el mismo parámetro de aprendizaje para refuerzo que para no refuerzo es una simplificación (March, 1996).

Bailey y Mazur (1990) condujeron un experimento con pichones, cuyo objetivo era examinar algunos factores que pudieran afectar la tasa de adquisición de una preferencia cuando dos alternativas proporcionan reforzadores con diferentes probabilidades. Se utilizó un procedimiento de ensayos discretos de dos elecciones (en este tipo de procedimiento, el sujeto emite una sola respuesta por ensayo, y esta respuesta puede o no ser reforzada). En cada una de diez condiciones experimentales, las dos teclas de respuesta tenían inicialmente probabilidades iguales de reforzamiento. Después se cambiaban las probabilidades de reforzamiento de tal forma 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 particular. Estos programas fueron los ensayos discretos análogos a los programas concurrentes de razón aleatoria, en los cuales comúnmente se encuentra la preferencia casi exclusiva para la alternativa con la más alta probabilidad de reforzamiento, una vez alcanzado el estado estable (Herrnstein & Loveland, 1975). Entre la sesión de una probabilidad y las sesiones de dos probabilidades, había una sesión de transición, la cual empezaba con el programa de una probabilidad y cambiaba al programa de dos probabilidades después de 100 ensayos. El propósito de cada condición fue empezar con los sujetos respondiendo aproximadamente igual en las dos teclas, y después observar el cambio en responder que ocurrió cuando era implementado el programa de dos probabilidades. Los valores de la

alta probabilidad de reforzamiento y la baja probabilidad de reforzamiento (ϕ_1 y ϕ_2 , respectivamente) se variaron a través de las condiciones, para determinar cómo afectaban la tasa de adquisición.

Los resultados indicaron que 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 era más alta. Además observaron que los animales eran más sensibles a las razones y no a las diferencias en las probabilidades de reforzamiento.

En un análisis ensayo por ensayo de las respuestas individuales y los reforzadores, se encontraron dos efectos sobre la elección. El efecto a corto plazo fue un incremento en la probabilidad de regresar a la misma tecla en el primero o segundo ensayos seguidos de un reforzador. El efecto de largo plazo fue un incremento gradual en la proporción de respuestas en la tecla con alta probabilidad de reforzamiento, un incremento que continuaba generalmente después de varios cientos de ensayos. Estos hallazgos parecen apuntar hacia un tipo de efecto de la ley de Weber en la adquisición de preferencias.

Los resultados obtenidos por Bailey y Mazur (1990) y Mazur y Ratti (1991) se resumen de la siguiente forma: Con programas concurrentes RV RV la adquisición de una preferencia ocurre más rápidamente con razones más grandes de $\phi_1:\phi_2$ aun cuando la diferencia entre las dos probabilidades de reforzamiento ($\phi_1-\phi_2$) se mantiene constante. Este resultado fue obtenido en procedimientos de ensayos discretos (Bailey & Mazur, 1990) y de operante libre (Mazur & Ratti, 1991). Segundo, cuando la razón $\phi_1:\phi_2$ se mantiene constante, la adquisición de una preferencia es más rápida con probabilidades más grandes de reforzamiento (Mazur, 1992 Experimento 1). El experimento 1

proporciona mayor evidencia acerca de la idea de que la tasa de transición depende de la diferencia ($\phi_1 - \phi_2$), pero también proporciona evidencia en contra de la hipótesis de Bailey y Mazur de que depende de la razón $\phi_1 : \phi_2$. Esta hipótesis también fue inconsistente con los patrones de adquisición en el experimento 2, porque no había evidencia para una aproximación más rápida a asíntota con proporciones de reforzamiento más extremas.

La habilidad de los sujetos para discriminar y responder a estas propiedades locales de los programas es interesante por sí misma y es consistente con una gran cantidad de investigación previa. Sin embargo, las ecuaciones básicas que intentan describir los efectos de reforzamiento y extinción en el modelo de operador lineal y sus diferentes versiones (el modelo de invarianza de razón, Staddon, 1988; Staddon & Horner, 1989; el modelo cinético, Myerson & Míezin, 1980; y el modelo de Couvillon y Bitterman, 1985) no están diseñadas para explicar estos efectos.

Modelo de Rendimientos Promedio

En el modelo de rendimientos promedio, $p_i(t)$ cambia con la historia de rendimientos que han sido reforzados en las diferentes alternativas. Se considera que el organismo posee una memoria del promedio de rendimientos para cada alternativa. Entonces, si $S_i(t)$ es la suma de los refuerzos recibidos cuando la alternativa i fue escogida en cualquier periodo anterior a t , y $m_i(t)$ es el número total de veces que la alternativa i fue escogida anterior a t , entonces:

$$A_i(t) = S_i(t) / m_i(t) \quad (5)$$

La probabilidad de escoger la alternativa i en el periodo t es proporcional a los promedios de resultados pasados. Donde las huellas de memoria son todas positivas ($j=1$ a n),

$$p_i(t) = A_i(t) / \sum A_j(t) \quad (6)$$

Entonces, en donde hay dos alternativas, la razón de las dos probabilidades es igual a la razón de su promedio de rendimiento a la situación actual. Donde las huellas de memoria son todas negativas ($j=1$ a n),

$$1-p_i(t) = A_i(t) / \sum A_j(t) \quad (7)$$

Los valores iniciales para $S_i(0)$ y $m_i(0)$ pueden estar determinados en cualquier valor, pero la suposición más simple es que $m_i(0) = 1$ y que $S_i(0) =$ al valor esperado en la i -ésima alternativa, entonces $p_i(0) = 1/n$ para todas las n alternativas, si las alternativas tienen valores esperados iguales.

El aspecto central del modelo de rendimientos promedio se deriva de varios modelos de aprendizaje que pueden ser agrupados bajo el rubro general de “modelos de igualación” (Herrnstein, 1961). El presente modelo es una simplificación. Por ejemplo, la mayoría de los modelos de igualación, no asume una memoria indefinidamente extensa de resultados (c.f., el modelo de efectos acumulativos de Davis et al., 1993), sino que normalmente asume una memoria de alguna duración fija. Sin embargo, la desviación más grande aquí, es relacionar el modelo con elecciones secuenciales entre alternativas con diferentes probabilidades de refuerzo basadas en ensayos, en vez de relacionarlo con tasas de respuesta en diferentes cantidades y demoras de refuerzo, aspectos que son centrales en la mayoría de los modelos en la

tradición de igualación (Williams, 1988). La relación es básicamente aquella entre modelos probabilísticos de tiempo discreto y un modelo de tiempo continuo.

Davis y Staddon (1990) condujeron un experimento en el cual un grupo de pichones fueron expuestos a la discriminación de cambios diarios sucesivos y cambios en bloques de dos y cuatro días, en la situación del bandido de dos brazos (concurrente razón aleatoria). Los sujetos escogían entre dos teclas de respuesta, derecha (D) e izquierda (I), cada una reforzada de acuerdo a programas probabilísticos (1/8). La tecla “rica” variaba de día a día en forma regular (IDID, IIDD, ó IIIDDDD). Los resultados obtenidos mostraron que los pichones cambiaron más rápidamente sus respuestas después de ser expuestos a cambios diarios que a cambios en bloques de dos días; y cambiaron más rápidamente después de la exposición a cambios en bloques de dos días que a cambios en bloques de cuatro días. Se observó una mejoría en la ejecución en los ensayos subsecuentes, principalmente en la condición de alternación diaria. En esta condición, los animales requirieron sólo de un reforzador para determinar la localización de la tecla rica cada día. Se observó que las primeras 50 respuestas al inicio de cada sesión, estaban determinadas por el programa de reforzamiento en el día anterior; y para las últimas 50 respuestas, el programa de reforzamiento actual era muy importante. De esto se puede decir que el modelo lineal es más molar que molecular.

Davis y Staddon (1990) asumen que la diferencia de patrones entre sesiones puede ser modelada por un modelo estocástico de aprendizaje como el de Bush y Mosteller (1955) o el de invarianza de razón de Staddon y Horner (Staddon, 1988; Staddon & Horner, 1989), y el supuesto que se maneja es que cada reforzamiento incrementa la tendencia del animal a responder en la opción

reforzada, y la ausencia de reforzamiento incrementa su tendencia a responder en la otra alternativa.

Una interpretación de estos resultados es que en este tipo de arreglos experimentales operan dos procesos aditivos. Uno es un proceso de memoria de largo plazo, que afecta la preferencia inicial de cada día. El segundo es un proceso de memoria de corto plazo, que permite que el animal se ajuste a las condiciones de reforzamiento dentro de cada sesión diaria. Davis y Staddon plantean que estos procesos son cualitativamente similares puesto que hay una escala de tiempo grande (24 horas para el proceso entre sesiones, y en el orden de segundos o minutos para el proceso dentro de una sesión). Esto significa que ambos procesos pueden ser el mismo, siendo la única diferencia la amplitud de la escala temporal.

Sin embargo, recientemente se ha demostrado que estos modelos de equilibrio fracasan cuando el procedimiento se modifica para evaluar efectos históricos y de acarreo (Davis et al., 1993). Si en lugar de cambiar diariamente en forma alterna cuál de las dos respuestas es reforzada, se les refuerza en dos bloques de varias sesiones y luego se extinguen, es frecuente encontrar un fenómeno de *regresión*, al preferir la respuesta reforzada en el primer bloque. De igual forma, los modelos citados no pueden dar cuenta de la mejoría en ejecución en procedimientos de reversión diaria sucesiva. Para explicar estos y otros resultados, Davis et al. (1993) han propuesto un modelo alternativo, el modelo de efectos acumulativos, que considera la historia lejana de reforzamiento.

Por otro lado, los estudios de aprendizaje de cambios incluyen exponer a los sujetos a diversos estímulos y arreglar diferentes programas de reforzamiento para las respuestas a cada estímulo. El estímulo discriminativo

puede ser presentado simultáneamente en un ensayo simple, o sucesivamente en ensayos consecutivos. Discriminaciones simultáneas y sucesivas pueden ser arregladas tanto en experimentos de ensayos discretos, como en experimentos en operante libre (Mackintosh, 1974).

En discriminaciones sucesivas, las presentaciones de S+ y S- son alternadas y únicamente las respuestas a S+ son reforzadas. En este tipo de discriminaciones, solo una respuesta esta disponible al sujeto y esa respuesta es reforzada o no reforzada, dependiendo del estímulo mostrado en ese ensayo particular. Hay evidencia que sugiere que los sujetos solucionan problemas sucesivos aprendiendo a seleccionar respuestas particulares (Couvillon, Brandon, Woodard & Bitterman, 1980; Couvillon & Bitterman, 1985, 1986; Davis & Staddon, 1990; Mackintosh, 1974; Staddon & Frank, 1974; Sutherland & Mackintosh, 1971; Theios, 1965; Williams, 1976a; Woodward & Bitterman, 1976).

En el aprendizaje de cambios, cuando se alternan las contingencias de reforzamiento entre dos elecciones, después de un cierto número de ensayos, los sujetos muestran una mejoría progresiva en la ejecución (Beale, 1970; Colwill & Delamater, 1995). El patrón asintótico en aprendizaje de cambios cuando los pichones son entrenados en condiciones en las cuales estímulos positivos y negativos cambian entre sesiones, es relativamente bien conocido: Los animales responden inmediatamente a ambos estímulos al inicio de cada sesión y después dejan de responder al estímulo negativo (Beale, 1970; Mackintosh, 1974; Staddon & Frank, 1974). El patrón es explicado fácilmente bajo la suposición de que la excitación permanece cerca de su asíntota mientras que la inhibición decae en el intervalo entre sesiones. Cuando el cambio sucede en una misma sesión, no es fácil de entender en términos de decaimiento

temporal, por qué se observa un incremento precipitado en el nivel de responder al primer estímulo negativo, ya que el intervalo entre ensayos permite que haya menos oportunidad para el decaimiento.

En varios procedimientos de aprendizaje de cambios (e.g., Woodward & Bitterman, 1976; Couvillon et al., 1980), se ha encontrado que el refuerzo y el no refuerzo por una respuesta en un ensayo particular, es utilizado por los sujetos como una señal para responder en un ensayo sucesivo. Ligado a esto, hay una variedad de fenómenos tales como el lento responder después de refuerzo que después de extinción (cuando refuerzo y extinción son regularmente alternados); decrementar la resistencia a la extinción en el curso de adquisición y extinción repetidas, y el efecto de reforzamiento parcial.

Modelo de Rendimientos Ponderados

En el modelo de rendimientos ponderados, así como en el modelo de rendimientos promedio, $p_i(t)$ cambia con la historia de rendimientos que han sido reforzados en varias alternativas. Se asume que el organismo actualiza la memoria de rendimientos después de cada elección de la alternativa i , computando un promedio móvil de la memoria previa y de la experiencia más reciente. Entonces, $A_i(t)$ permanece sin cambio en cualquier periodo en el cual la alternativa i no es escogida. Si la i -ésima alternativa es escogida al tiempo t , $A_i(t+1)$ es una mezcla entre $A_i(t)$ y el resultado realizado en el periodo t , $O_i(t)$, con el peso asignado a ambos, $(1-\alpha)$ y α respectivamente. Entonces,

$$A_i(t+1) = \alpha O_i(t) + (1-\alpha) A_i(t) \quad (8)$$

$$(0 \leq \alpha \leq 1)$$

Experiencias más recientes en cualquier alternativa particular se ponderan más que experiencias más remotas en esa alternativa, con la tasa relativa de aprendizaje u olvido que es indicada por α . El valor de $p_i(t)$ se deriva de $A_i(t)$ de la misma forma como en el modelo rendimientos promedio (ecuaciones 6 y 7).

Comúnmente, el estimado de reforzadores permanece constante de momento a momento, pero en ciertos periodos en el tiempo, el sujeto lo actualiza. Es este proceso de actualización lo que constituye la idea común del modelo, que incluye multiplicar el viejo estimado por un factor ponderado menor que la unidad (α), y agregarle a éste una cantidad que dependa de si un reforzador acaba de ser recibido o no, multiplicado por el complemento del peso distribuido al viejo estimado ($1-\alpha$). Frecuentemente, esta cantidad actual de obtención de reforzador será simplemente 0 ó 1, pero podría ser el recíproco del número de respuestas requeridas para obtener el último reforzador, o el tiempo invertido en realizar esto. Para completar la descripción del modelo, es necesario especificar la regla por la cual la densidad estimada de reforzadores se transforma en comportamiento.

Entre más cercano a la unidad este el factor ponderado α , el sujeto está más influenciado por el pasado que por el presente. Otra forma de describir esto es asumiendo que valores grandes de α implican una amplia memoria. De hecho, la cantidad $1/(1-\alpha)$ es un estimado de qué tan remota en el pasado debe de ser la obtención previa de un reforzador antes de que ésta no tenga influencia en el comportamiento actual, es decir, la amplitud temporal de la ventana de memoria.

El modelo de rendimientos ponderados puede ser considerado como un modelo que captura algunos elementos de una versión limitada de memoria del modelo de rendimientos promedio. Es equivalente a mantener una memoria de los promedios móviles ponderados exponencialmente de los rendimientos para la elección de cada alternativa (Killeen, 1981, 1994). Los modelos de promedios móviles ponderados exponencialmente son formas aceptables de balancear la necesidad de muestras grandes con las posibilidades de que las observaciones más recientes contengan información más relevante para estimar las condiciones actuales.

De esta manera, hay una posibilidad de cambiar el peso de la memoria. Killeen (1981, 1994) incluye la idea de que el factor de memoria puede ser variable y dependiente de la tasa de obtención del reforzador. Su modelo asume que los sujetos dan relativamente más peso a la experiencia inmediata cuando la tasa de refuerzo es alta. Este cambio implica incluir un parámetro adicional, responsable de la tasa de ajuste del factor de memoria. También se postula que el factor de memoria (α), cambia como un promedio móvil ponderado exponencialmente, y que el modelo emplea formas de igualación como la regla de decisión.

Por otra parte, los modelos de rendimientos ponderados varían enormemente en el énfasis de sus suposiciones básicas. Lea y Dow (1984) establecen las formas principales en las cuales pueden variar:

1. *La naturaleza del parámetro que es estimado.* Las principales posibilidades son: a) Densidad de reforzamiento por respuesta, y b) densidad de reforzamiento por unidad de tiempo. En el último caso, se puede tener un estimado local o uno global i.e., el reforzador obtenido en la alternativa por

el tiempo invertido en esa alternativa, o el reforzador obtenido en esa alternativa por el tiempo total invertido en la sesión completa.

2. *El evento que desencadena la actualización de la densidad estimada.* Las principales posibilidades son: a) El paso de cierto periodo, b) la ocurrencia de una respuesta, y c) la obtención del reforzador. También hay posibilidades compuestas e.g., la actualización puede ocurrir después de la obtención de un reforzador, o cuando un cierto periodo haya transcurrido sin obtener algo.
3. *El valor del parámetro que pondera el pasado contra el presente (α), y si es una propiedad fija o variable del organismo.* Si el peso es variable, éste puede cambiar de día a día, pero tener el mismo valor para cada alternativa en un día específico; o puede tener diferentes valores para cada una de las alternativas.
4. *La regla que transforma la densidad estimada de refuerzo en comportamiento.* Por ejemplo, si la transformación de la densidad de refuerzo en comportamiento es absoluta o estocástica, en otras palabras, si el animal escoge siempre la opción con la densidad estimada actual más alta. La alternativa que parece más evidente es que la distribución de tiempo sea una función continua de la densidad estimada, de tal forma que todas las alternativas sean aprovechadas, aunque aquella con la densidad estimada más alta sea la que más se aproveche.
5. *El valor asignado a la densidad estimada de reforzamiento antes del inicio de una sesión experimental.* Las posibilidades principales para estimados iniciales de densidad de reforzamiento incluyen: a) El valor de cero, b) el promedio encontrado en alternativas similares en el pasado, y c) el valor

necesario para asegurar el consumo diario básico dentro del tiempo disponible de una sesión.

Killeen (1981, 1994) plantea que el responder exclusivamente a la experiencia muy reciente dejaría al animal muy vulnerable a sobrerreaccionar ante breves fluctuaciones azarosas en el medio ambiente. Además sugiere que la forma de explorar cambios medio ambientales con una carga mínima de memoria es usar alguna forma de promedios móviles ponderados exponencialmente (véase Dow & Lea, 1987; Killeen, 1982), esto es, que hay una amplitud de memoria óptima que depende de la estabilidad del medio ambiente, y debe ser ajustada ya sea a través de la evolución o a través de la experiencia individual (Shettleworth & Plowright, 1992). El estimado actual del parámetro medio ambiental en cuestión debe ser un promedio ponderado de la experiencia más reciente y del promedio previo.

En un ambiente que es constante en promedio, las fluctuaciones locales aleatorias deben ser ignoradas, pero en un medio ambiente que cambia frecuentemente (pero no demasiado frecuente) la ventana de memoria debe ser lo suficientemente corta para permitir una rápida exploración de un cambio genuino.

Se ha estudiado en el contexto del forrajeo, cómo los organismos se adaptan a cambios constantes en el ambiente (Dow & Lea, 1987; Kacelnik et al., 1987; Kacelnik & Todd, 1992; Lea & Dow, 1984; Shettleworth, 1988; Shettleworth, Krebs, Stephens & Gibbon, 1988; Shettleworth & Plowright, 1989, 1992). Experimentos recientes han estudiado las decisiones de partida de starlings (*Sturnus vulgaris*) de los sitios de alimentación, en los cuales se varió el tiempo de viaje entre éstos (Cuthill, Kacelnik, Krebs, Haccou & Iwasa,

1990; Kacelnik & Todd, 1992; Bateson & Kacelnik, 1995; Brunner, Kacelnik & Gibbon, 1996). Los resultados mostraron que los animales no siempre promedian la experiencia en la forma óptima para una situación particular, pero en cambio, pueden sobrestimar el pasado reciente. Cuthill et al. (1990), evaluaron a starlings en un ambiente operante con un sitio de alimentación sencillo (programa de razón progresiva). Los resultados indicaron que los animales se quedaban por tiempo más prolongado después de un tiempo de viaje largo, que después de un tiempo de viaje corto, como si estuvieran actualizando su estimado del tiempo de viaje en una ventana de memoria muy corta. Los viajes que precedieron a aquellos más recientes, parecen no haber tenido efecto sobre el tiempo de partida del sitio de alimentación actual.

De esta forma, responder al promedio medio ambiental es óptimo, sin embargo, sólo si el medio ambiente es verdaderamente estable en un largo plazo y el tiempo disponible para buscar alimento (horizonte de tiempo) es largo. Cuthill et al. mostraron que, al menos para el problema de tiempo de viaje que ellos estudiaron, explorar es óptimo si el ambiente, el sitio de alimentación y el horizonte temporal son cortos; sugiriendo que los animales deben de tener en la situación experimental, una regla de promedio que sea adaptativa para los diferentes tipos de ambiente.

Shettleworth y Plowright (1992, Experimento 1) entrenaron a un grupo de pichones en programas operantes simulando encuentros sucesivos con presas. Cuando las presas eran encontradas de acuerdo con programas de intervalo variable, los animales aceptaron con mayor facilidad una presa pobre (demora de comida larga) mientras más tiempo habían buscado, como si estuvieran promediando la densidad de presa en una ventana de memoria corta. En el segundo experimento, cuando una búsqueda corta predecía una búsqueda

que la última presa fue encontrada. Esto es consistente con muchos datos de elección en estado estable en programas de reforzamiento, donde la demora promedio para el reforzamiento parece ser un parámetro más útil que el promedio de la frecuencia de reforzamiento (Killeen, 1968).

Dreyfus (1991) condujo un experimento cuyo propósito fue determinar el grado en el cual la distribución de conducta en un programa concurrente cambiaba como una función de los cambios a corto plazo en la tasa relativa de refuerzo. En este experimento, los pichones fueron expuestos a programas concurrentes mixtos IV IV los cuales estaban caracterizados por uno o más cambios señalados en la tasa relativa de refuerzo durante una sesión. En los dos experimentos que condujo, se varió el número de cambios en la distribución de reforzadores dentro de una sesión, la duración del componente y la tasa total de refuerzo. Para el análisis de resultados, Dreyfus tabuló la razón de distribución de tiempo y la razón de los tiempos dedicados a cada programa, en intervalos sucesivos cortos (ventanas) antes y después de los cambios en la tasa relativa de refuerzo.

Los resultados indicaron que las razones de distribución de tiempo cambiaron muy rápidamente, después de los cambios en la tasa relativa de refuerzo. Dreyfus observó que la distribución de tiempo relativo se ajustó a los cambios en la distribución de los reforzadores más estrechamente en las siguientes condiciones: a) Cuando había pocos cambios dentro de una sesión, b) cuando los componentes eran largos, y c) cuando las tasas totales de refuerzo eran más altas.

Estos resultados sugieren que la experiencia global del animal en los programas de refuerzo (su experiencia en las sesiones) puede afectar fuertemente su respuesta a cambios en la tasa relativa de refuerzo dentro de una

sesión. En otras palabras, la magnitud del cambio en la distribución de tiempo dentro de la duración de un intervalo entre-reforzadores del componente pobre después del cambio programado, dependió del tipo de sesiones que el animal había experimentado recientemente. Si esas sesiones tuvieron ensayos que duraron varios intervalos entre-reforzadores esperados en el programa pobre, entonces los animales alteraban su distribución de tiempo muy rápidamente después de que la tecla rica se convertía en la tecla pobre. Pero si las sesiones habían tenido ensayos que duraban solo la duración del intervalo entre-reforzadores esperado en el componente pobre, entonces los animales no alternaban su distribución de tiempo cuando la tecla rica se convertía en la tecla pobre. En este tipo de sesiones, dividieron su tiempo de igual forma entre las teclas, independientemente de la tasa relativa de refuerzo durante un ensayo determinado dentro de la sesión. De esta forma, los datos reportados por Dreyfus (1991) demuestran que la tasa relativa de refuerzo o sus componentes, pueden controlar el comportamiento en un periodo más breve que aquel implicado por la mayoría de las teorías molares y que el periodo en el cual los reforzadores afectan el comportamiento subsecuente depende de la tasa total de refuerzo (Dreyfus, 1991).

Por otra parte, Kacelnik et al. (1987) condujeron un experimento donde starlings escogían entre dos sitios de alimentación, en los cuales la comida estaba disponible de acuerdo a un programa de razón aleatoria. En la alternativa “estable” (asignada aleatoriamente para cada sesión) la probabilidad permanecía en 0.08 por toda la sesión, mientras que en la otra opción (“cambiante”) iniciaba con una alta probabilidad (0.25 ó 0.75) y eventualmente cambiaba a cero, pero esta diferencia no fue señalada para los sujetos.

Los resultados indicaron que los starlings rápidamente aprendieron a discriminar entre las dos alternativas y respondieron casi exclusivamente en aquella que ofrecía la probabilidad más alta. Todos los animales mostraron la misma tendencia, pero los starlings en el grupo 0.75 desarrollaron una preferencia más rápida comparada con el grupo 0.25. En otras palabras, ante un cambio no señalado en la probabilidad de refuerzo, los animales cambiaron de alternativa más rápidamente cuando el cambio a cero ocurrió en la alternativa con la probabilidad más alta que cuando el cambio ocurrió en la alternativa con probabilidad de refuerzo baja.

Asimismo, varios experimentos acerca de la dinámica del proceso de adaptación inicial a programas temporales, indican que los pichones son capaces de reaccionar rápidamente a cambios en el patrón temporal de entrega de comida, al menos a los intervalos entre-reforzadores cortos empleados en estos experimentos (Higa, 1996a, 1996b; Higa, Wynne & Staddon, 1991; Wynne, Staddon & Delius, 1996). También demostraron que los pichones siguieron diferentes series de cambios rápidamente en los intervalos entre-reforzadores que variaron aleatoriamente de día a día. Los cambios rápidos en tiempo de espera parecen ser similares a los resultados recientes de estudios de elección (Higa, 1996b).

Por otro lado, se ha observado en experimentos empleando diversos programas de reforzamiento, la reaparición de una respuesta preferida previamente extinguida. Estos cambios en la ejecución pueden ocurrir en un intervalo cuando los sujetos no son expuestos a contingencias de reforzamiento, y aun cuando los sujetos no son expuestos a la situación experimental (efectos del tiempo entre sesión). A este fenómeno se le denomina *Recuperación espontánea*, y el término ha sido ocasionalmente aplicado a la conducta en

situaciones de elección (véase Beale, 1970). Hay varias teorías que se han propuesto para explicar este cambio “espontáneo” (e.g., Robins, 1990; Devenport, Hill, Wilson & Ogden, 1997) y que se derivan de las teorías de extinción (véase Mackintosh, 1974, para una revisión).

Mazur (1995, 1996) propuso que el comportamiento al inicio de una sesión refleja alguna forma de promedio ponderado de los eventos de varias sesiones pasadas. Los experimentos de Mazur (1995), además de otros que demuestran recuperación espontánea en elección (Mark & Gallistel, 1993; Mazur, 1996; Devenport et al., 1997), proporcionan algunas explicaciones acerca de cómo las experiencias pasadas son almacenadas en la memoria del animal y cómo estas experiencias se combinan para afectar la conducta actual. Una de las formas más simples para representar las experiencias pasadas en una situación de elección, sería asignar un número a cada alternativa que represente la fuerza o el valor de esa alternativa (tales como $p_i(t)$ y $p_i(t+1)$ en las ecuaciones 1-4) y después continuamente actualizar ese número conforme se hacen nuevas elecciones. Esta estrategia es incorporada ya sea explícita o implícitamente, en los modelos matemáticos que usan promedios móviles ponderados exponencialmente (e.g., Couvillon & Bitterman, 1985; Killeen, 1982; Staddon, 1988). Este método de actualizar preferencias en el tiempo es matemáticamente conveniente, pero desafortunadamente no puede explicar recuperación espontánea u otros efectos del paso del tiempo.

Es muy fácil explicar estos cambios entre sesión si asumimos que un animal almacena representaciones separadas para cada alternativa en cada sesión, y que las representaciones de varias sesiones pasadas son promediadas al inicio de cada nueva sesión.

Varios modelos asumen que el comportamiento al inicio de la sesión refleja la influencia combinada de diferentes medidas de fuerza de la respuesta (estimados separados de $p_i(t)$, según el modelo de promedios). Quizá esta sea la suposición más básica que se pueda derivar de la evidencia de recuperación espontánea en situaciones de elección. Sin embargo, postular una representación simple para cada alternativa de respuesta (ya sea V , $p_i(t)$ o cualquier otra medida) no es suficiente. Esta aproximación no puede explicar por qué las proporciones de respuesta al inicio de una nueva sesión se mueven hacia indiferencia en algunas situaciones y lejos de indiferencia en otras, dependiendo de las contingencias de reforzamiento que prevalecieron en sesiones previas.

Devenport et al. (1997) proponen como alternativa a los modelos de promedios móviles ponderados exponencialmente, una regla de peso temporal que permite explicar los datos obtenidos en estudios de forrajeo en ambientes variables y recuperación espontánea, y cuya significancia adaptativa es muy clara. A diferencia del descuento exponencial, la regla de peso temporal descuenta la experiencia reciente a un ritmo más rápido que la experiencia más lejana. A simple vista, la tarea de implementar la regla de peso temporal parece difícil, pero no lo es. Los animales comparan intervalos como si compararan razones, con el resultado de que los eventos que están lejanos en el pasado parecen más y más contemporáneos. La regla de peso temporal ajusta los pesos como una función hiperbólica de tiempo, es decir, los pesos temporales simplemente se vuelven más parecidos conforme se agrega una cantidad constante de tiempo.

El modelo propuesto por Devenport et al. (1997) parece ser hasta el momento, la aproximación que mejor explica los datos obtenidos sobre

recuperación espontánea en la conducta de elección (Mark & Gallistel, 1993; Mazur, 1995, 1996). Sin embargo, indudablemente será necesario trabajo adicional, antes de que se afirme con mayor precisión cómo los eventos de varias sesiones pasadas se combinan para afectar el responder al inicio de una nueva sesión.

Por último, los tres modelos descritos pueden ser evaluados en situaciones que incluyan escoger entre dos alternativas, cada una caracterizada por una magnitud y una probabilidad de refuerzo fijas. Como se esperaría, cada modelo presenta distribuciones de $p_i(t)$ en series de ensayos de aprendizaje que son consistentes con el aprendizaje animal en situaciones estándar.

Teorías Representacionales

Los modelos anteriores son versiones de la ley del efecto, que asumen que las respuestas individuales ganan y pierden fuerza por episodios individuales de reforzamiento/no reforzamiento. La segunda familia de modelos, no asume un proceso de incremento/decremento, sino que el animal tiene una representación de su medio ambiente, en términos de intervalos entre-reforzadores y tasas de ocurrencia de sucesos.

Cuáles son las variables que determinan el patrón de respuesta observado experimentalmente en situaciones de elección operante, es una pregunta extremadamente difícil y probablemente tiene más de una respuesta. Una de las disputas más populares sobre los determinantes de la elección está centrada alrededor de dos variables: La probabilidad de reforzamiento programada y la

probabilidad de reforzamiento obtenida (que emerge de las contingencias respuesta-reforzamiento).

Aparentemente, ni la probabilidad local de reforzamiento ni la probabilidad programada pueden describir la ejecución de elección en todas las condiciones. Williams (1994) concluye que cualquier modelo que considere la probabilidad de reforzamiento como la variable que controla el comportamiento, está abierto a una seria pregunta. De hecho, esta conclusión apunta hacia otro tipo de explicación: Quizá no podamos encontrar una respuesta satisfactoria mientras se evalúen variables individuales que parecen determinar la elección. Probablemente se necesita diseñar experimentos que investiguen correlaciones entre sus múltiples determinantes, una vez que sea claro que la probabilidad de reforzamiento esta involucrada en algunas condiciones. Para entender estas condiciones, necesitamos evaluar mecanismos en lugar de variables que influyan directamente en la elección. A este respecto, Williams (1994) propuso las teorías representacionales como una alternativa para una variable determinante fundamental.

Representación de Demoras de Reforzamiento

Innumerables estudios de comportamiento operante y forrajeo muestran que los animales responden a la tasa promedio de disponibilidad de comida en el pasado relativamente reciente. Por ejemplo, la ley de igualación (Herrnstein, 1961) describe relaciones ordenadas entre medidas promedio del comportamiento y el reforzamiento dentro de las condiciones experimentales.

Una teoría de elección explícitamente basada en la suposición de que el sujeto adquiere un conocimiento de la frecuencia de reforzamiento asociada con las diferentes alternativas, y con la base de alguna regla distribuye su

comportamiento, ha sido propuesta por Gibbon, Church, Fairhurst y Kacelnik (1988). Esta teoría constituye una teoría representacional de igualación (en la cual el concepto tradicional de fuerza de la respuesta es irrelevante), que explica cómo los animales obtienen los estimados de las tasas de reforzamiento que deben subyacer estos datos. Gibbon et al. (1988) sugirieron que cuando la comida es distribuida en tiempo, los animales acumulan una memoria exhaustiva de las esperas de comida asociada con cada alternativa disponible y una vez que se forma la representación de memoria, el valor del intervalo o cantidad original puede ser recordado tomando una muestra simple aleatoria de su distribución de memoria asociada, cada vez que hacen una elección. Se asume que los sujetos escogen entre dos opciones recuperando una muestra de memoria para cada opción, comparando estas muestras y prefiriendo la opción que ofrece el mejor valor (tamaño grande de refuerzo o demora corta). La memoria para una demora determinada tiene la propiedad escalar, i.e., la varianza en memoria es proporcional a la media de la demora. Específicamente, las demoras son recordadas con una varianza que es proporcional al cuadrado de la media.

Aunque el modelo fue desarrollado para experimentos en los cuales los refuerzos ocurren sobre la base del tiempo, también puede operar en situaciones en las cuales los refuerzos son programados en términos de probabilidad por respuesta. La aproximación de Gibbon et al. (1988), ha sido aplicada exitosamente a otras simulaciones operantes de forrajeo y a programas operantes convencionales (Kacelnik & Todd, 1992; Shettleworth et al., 1988). Esta idea es similar a mejoramiento y maximización momentánea en que la elección es gobernada por demoras en el refuerzo, pero difiere al asumir que la elección es gobernada por la distribución total recordada de las demoras de

refuerzo, no por eventos a corto plazo (como el tiempo desde la última elección de una de las dos alternativas).

Varias teorías de igualación, incluyendo maximización momentánea (Shimp, 1966), mejoramiento (Herrnstein & Vaughan, 1980), y el modelo reciente de efectos acumulativos de Davis et al. (1993), comparten la suposición de que la variable determinante es la probabilidad obtenida de reforzamiento para las diferentes alternativas. Esto es, igualación no es la regla de elección *per se*, sino es el resultado de que el sujeto siempre está escogiendo la alternativa con la más alta probabilidad de reforzamiento en ese momento (véase Williams, 1988, para una discusión). Evidencia reciente ha puesto a prueba esta clase general de modelos, mostrando que la probabilidad obtenida de reforzamiento no es de hecho la variable crítica (Belke, 1992; Williams, 1993, 1994; Williams & Royalty, 1989).

Williams y Royalty (1989) diseñaron tres experimentos para evaluar la hipótesis de mejoramiento. En el experimento 3, los pichones fueron expuestos a dos componentes alternos de un programa múltiple de ensayos discretos. El primer componente presentaba la elección entre dos alternativas asociadas con programas IV interdependientes, reforzados con probabilidades de 0.2 y 0.05. El segundo componente era un IV simple con probabilidad de 0.1. Dado que los pichones igualaban durante el primer componente, las probabilidades de reforzamiento obtenidas eran iguales para ambas alternativas (0.24). Para el segundo componente, la probabilidad de reforzamiento obtenida fue igual al valor programado (0.1). En una prueba que se hizo al final del entrenamiento, los sujetos tenían que escoger entre la alternativa con probabilidad 0.2 y la alternativa con probabilidad 0.1, o entre la alternativa con probabilidad 0.05 y la alternativa con probabilidad 0.1. Los resultados mostraron que los pichones

prefirieron la alternativa con probabilidad 0.2 sobre la alternativa con probabilidad 0.1, mientras la alternativa 0.1 fue preferida sobre la alternativa 0.05, lo que sugiere que la probabilidad de reforzamiento programada, no la probabilidad obtenida, era la variable determinante.

El experimento 1 fue similar en procedimiento al experimento 3. Los pichones fueron entrenados simultáneamente en dos programas concurrentes IV IV. En una condición, los animales tenían que escoger entre un IV20 seg y un IV120 seg, y en la condición alterna tenían que escoger entre un IV60 seg y un IV80 seg. De acuerdo a igualación, los pichones igualaron las probabilidades de reforzamiento en ambas condiciones de entrenamiento, excepto que la probabilidad absoluta de reforzamiento en la cual ocurrió igualación fue más alta en la condición IV20 seg IV120 seg (aproximadamente 210 reforzadores/hora) que en la condición IV60 seg IV80 seg (aproximadamente 105 reforzadores/hora), puesto que la probabilidad de igualación en equilibrio esta limitada por el programa IV más rico. La condición de prueba consistió en presentar el estímulo IV120 seg con el estímulo IV60 seg en extinción. Mientras la tasa de reforzamiento programada (y el número absoluto de reforzadores distribuidos en cada alternativa) favoreció a la alternativa IV60 seg, la probabilidad local obtenida de reforzamiento favoreció a la alternativa IV 120 seg.

Los pichones violaron nuevamente la hipótesis de mejoramiento y escogieron sobre la base de tasas de reforzamiento programadas, prefiriendo el IV60 seg sobre el IV120 seg. De esta forma, los experimentos de Williams y Royalty (1989) muestran las limitaciones de los modelos de probabilidad local.

En el experimento de Williams (1993), un grupo de pichones fue entrenado en un programa de reforzamiento múltiple, en el cual dos

componentes separados alternaron periódicamente (cada uno consistía de una serie de ensayos discretos en los cuales una sola respuesta era permitida). Durante el componente de elección del programa múltiple, dos respuestas alternativas estaban disponibles, una asociada con una probabilidad de reforzamiento programada de .20 y la otra con una probabilidad de reforzamiento programada de .05. La probabilidad de un reforzador disponible fue determinada independientemente para ambos programas al inicio de cada ensayo. Este procedimiento de ensayos discretos simulaba las contingencias de un programa de reforzamiento concurrente IV IV en operante libre, en el que la probabilidad de un reforzador disponible para una alternativa específica, incrementaba con el número de ensayos que intervenían desde la elección precedente de esa alternativa.

El resultado en el componente de elección fue aproximadamente igualación. Esto es, en alrededor del 80% de los ensayos la respuesta fue dirigida a la alternativa correlacionada con el programa que tenía la probabilidad de reforzamiento de .20, correspondiendo a la razón de 4:1 en las tasas de reforzamiento obtenidas para las dos alternativas. Además, la razón de reforzadores obtenida de las dos alternativas fue de 4:1, y la razón de respuestas emitidas fue también de 4:1.

En el segundo componente del programa múltiple, una alternativa simple estuvo disponible con una probabilidad programada de reforzamiento de .10. La probabilidad obtenida de reforzamiento también fue de .10. La situación de prueba consistió en colocar en forma separada cada una de las dos alternativas con la alternativa de .10 que había sido presentada sola. Puesto que las probabilidades de reforzamiento obtenidas para las alternativas fueron aproximadamente iguales, la predicción de la teoría de mejoramiento y el

modelo de efectos acumulativos de Davis et al. (1993) es que cada una de las alternativas debería haber sido preferida en un grado similar sobre la alternativa .10. Los resultados reales fueron que la alternativa .20 del componente de elección fue preferida sobre el estímulo .10, mientras que el estímulo .10 fue preferido sobre la alternativa .05.

Una posible explicación de este hallazgo y los hallazgos reportados por Williams (1994), es que la elección es controlada por la distribución de intervalos entre-reforzadores asociados con cada estímulo. En otras palabras, el valor de una alternativa no es calculado con respecto a la probabilidad de reforzamiento recibida por responder a esa alternativa; en lugar de ello, los sujetos registran de alguna forma el conjunto completo de los tiempos entre-reforzadores para cada estímulo.

Los hallazgos de Williams (1993) apoyan esta posibilidad. Sin embargo, un estudio reciente de Belke (1992) representa el reto más grande a las explicaciones existentes sobre igualación. Los resultados de Belke hacen muy clara la necesidad de entender la dinámica de la elección concurrente, y muestran las limitaciones de los modelos que explican elección usando la suposición libre de contexto.

En el estudio de Belke (1992), los pichones escogieron entre dos pares de programas concurrentes IV en una cámara de cuatro teclas. Dos teclas en un lado fueron dedicadas a un par de IV; las dos teclas restantes fueron dedicadas al otro par. El entrenamiento en línea base consistió en concurrentes IV20 seg IV40 seg asociados con las teclas blanca y roja en la izquierda, y concurrentes IV40 seg IV80 seg asociados con la tecla verde y amarilla en la derecha. Los sujetos fueron expuestos a cada par de programas por un minuto, seguido por 10 segundos de tiempo fuera, y por una exposición al par alterno. Belke

encontró igualación a las tasas relativas de reforzamiento, esto es, una preferencia de 2/3 en la distribución de tiempo para la tecla blanca sobre la roja (el doble del tiempo invertido en la blanca) y una preferencia de 2/3 para la verde sobre la amarilla, como se esperaría. La preferencia fue expresada en tasas de respuesta, pero las tasas locales fueron aproximadamente constantes en las dos situaciones; en consecuencia, igualación de respuestas fue el resultado de igualación de tiempo.

Durante la situación de prueba, se incluyeron periodos cortos de 1 minuto, durante los cuales las teclas roja y verde fueron iluminadas (en lugar del par rojo-blanco o el par verde-amarillo), y no hubo reforzamiento. En este periodo, los pichones mostraron una preferencia de 4:1 para la tecla verde. Distribuyeron el 79% de sus respuestas y el 82% de su tiempo a ésta. La preferencia de los pichones por la tecla verde sobre la tecla roja fue significativamente más fuerte, como su preferencia por la tecla verde sobre la tecla amarilla. Esto es sobresaliente, porque las teclas verde y roja estaban asociadas con el mismo programa IV40 seg, mientras que la tecla amarilla estuvo asociada con el programa IV80 seg.

Sin embargo, las situaciones de prueba no reforzadas entre rojo y verde mostraron una preferencia de 4:1 en la distribución de tiempo para la tecla verde. Como las probabilidades locales y programadas de reforzamiento eran iguales para las dos teclas, se esperaría indiferencia entre ellas de acuerdo a las explicaciones molares de igualación. Si la probabilidad local (obtenida) de reforzamiento fuera el factor principal que determina la elección, entonces los pichones deben de haber preferido el estímulo IV40 seg (previamente asociado con el IV20 seg).

De acuerdo con la explicación de Gibbon (1995) y Gibbon et al. (1988), si los sujetos muestrean una memoria para el IV40 seg en el rojo y una memoria para el IV40 seg en el verde, estas dos muestras son en promedio igualmente parecidas y por lo tanto, el nivel de igualación debe ser indiferencia. Por otro lado, como Williams (1994) y Mark y Gallistel (1994) han explicado, pudiera ser que esta probabilidad de cambio de un programa fuera la variable predominante para la preferencia en esta situación. Los sujetos pueden responder en la situación de prueba como si la alternativa fuera la alternativa original en el entrenamiento.

Una explicación cualitativa de los resultados de Belke puede ser atribuida a la generalización de preferencias. Los sujetos en la situación de prueba prefirieron el programa que había sido preferido en la situación de entrenamiento (prefirieron el verde sobre el rojo).

Si la probabilidad de reforzamiento obtenida, y la distribución de intervalos entre-reforzadores no controlan el grado de preferencia, es importante establecer cuáles podrían ser las variables que controlan el comportamiento. Dos explicaciones de los resultados de Belke (1992) que ofrecen Williams y Bell (1996) parecen ser posibles. La primera es que el control por los intervalos entre-reforzadores fue modulado por las mismas variables que producen el fenómeno de contraste conductual (véase Williams, 1988). Es decir, que el programa IV 40 seg apareado con la alternativa IV80 seg fue valorado más porque éste ocurrió en un contexto pobre de reforzamiento.

Una explicación diferente es que la variable crítica que determina el patrón de preferencia en las situaciones de prueba fue el patrón de comportamiento establecido durante el entrenamiento en línea base, que

precedió a las diferentes alternativas presentadas por primera vez juntas en los ensayos de prueba. En otras palabras, el entrenamiento inicial estableció un tiempo promedio operando en un programa determinado antes de cambiar al programa alternativo, y este promedio de tiempos de permanencia después fue transferido a los ensayos de prueba con la nueva combinación de alternativas. Entonces en el estudio de Belke (1992), el entrenamiento en el componente del programa múltiple con los programas concurrentes IV20 seg IV40 seg, estableció duraciones de permanencia relativamente cortas en la alternativa correlacionada con el programa IV40 seg; mientras que en el componente con el programa concurrente IV80 seg IV40 seg, se establecieron duraciones de permanencia relativamente largas al programa de IV40 seg. Durante la situación de prueba, los dos estímulos que habían sido correlacionados con las diferentes alternativas IV40 seg, fueron evaluados uno con el otro. El que había estado apareado con el programa IV80 seg fue fuertemente preferido porque el sujeto continuó con el patrón de permanecer más tiempo en ese estímulo antes de cambiar a la otra alternativa.

A un nivel más general, los resultados de Belke causan dificultad para la posición general de que igualación ocurre debido a la distribución automática de responder en proporción a las tasas de reforzamiento asociadas con diferentes alternativas. Si la preferencia es entendida como el resultado del patrón de comportamiento que ha sido establecido a un estímulo, entonces las explicaciones en términos de tiempos de permanencia y probabilidades de cambio, deben ser preferidas sobre aquellas que asumen que la probabilidad de reforzamiento o los intervalos entre-reforzadores controlan la distribución de elección directamente. Esto es, cualquier explicación teórica adecuada debe

sido apareada con el programa de más alto valor. Los cambios en el patrón de preferencia implican que el patrón de cambio del comportamiento establecido durante el entrenamiento, fue lo que determinó los patrones de preferencia obtenidos en los ensayos de prueba.

Los resultados de Williams y Bell (1996) sugieren que el principal determinante de la preferencia durante los ensayos de prueba, fué el patrón de los tiempos de permanencia que había sido establecido durante el entrenamiento en línea base. Esto es, los sujetos parecen haber transferido el patrón de tiempos de permanencia a un estímulo determinado de línea base a los ensayos de prueba, sin considerar la probabilidad o tasa de reforzamiento por responder a los diferentes estímulos.

Representación de Tasas de Ocurrencia de Sucesos

Existe una posición alternativa dentro de las teorías representacionales, que sugiere que los estimados de tasa que los animales realizan, se derivan de computaciones sobre representaciones independientes de números y tiempos de entrega de comida. Gallistel (1990) ha proporcionado un argumento fuerte y convincente de que el aprendizaje no consiste de la formación de relaciones asociativas, sino de representaciones de la estructura del ambiente. En otras palabras, el animal adquiere conocimiento de las propiedades espaciales y temporales de los eventos importantes, lo cual es medido sólo indirectamente, por cómo esos eventos tienen contacto con el comportamiento por una contingencia de reforzamiento. Por ejemplo, desde este punto de vista, la ley de igualación es el resultado de que el animal tenga conocimiento de diferentes distribuciones de intervalos entre-reforzadores contingentes para cada alternativa, y no es el resultado del número de apareamientos de la respuesta y

el reforzador, o de la probabilidad de que una respuesta sea seguida por un reforzador. Evidencia substancial (Gibbon et al., 1988; Mark & Gallistel, 1994; Williams, 1993) existe ahora en favor de dicho análisis.

Gallistel (1990, 1994) plantea que si se asume que el animal posee un aparato neural que pueda medir intervalos, contar ocurrencias, estimar magnitudes y desempeñar operaciones isomórficas a la división y multiplicación, en las variables internas resultantes (i.e., las representaciones neurales del intervalo temporal, el número de ocurrencias y las magnitudes), entonces es posible elaborar una explicación de los procesos internos que explican la proporcionalidad observada entre la tendencia del animal a cambiar a un sitio de alimentación determinado y la prevalencia relativa de comida en ese sitio de alimentación. Sin embargo, tradicionalmente ha habido gran rechazo a atribuir la capacidad representacional-computacional a los animales infrahumanos.

La precisa correspondencia entre la distribución de tiempo del animal y las tasas relativas de refuerzo es evidencia de que ha computado una representación de esas tasas de representaciones de número y tiempo, en esa medida se asume que el animal tiene una representación de número en el sentido rico del término (Gallistel, 1990).

Mark y Gallistel (1994) desarrollaron un experimento para medir los efectos de cambios ensayo a ensayo en las tasas relativas y en las magnitudes de refuerzo. En su experimento, las ratas respondían a un programa concurrente IV IV teniendo como reforzador estimulación cerebral, en sesiones de dos ensayos. En las condiciones experimentales, se cambió la tasa relativa de reforzamiento programada en razón de 4:1 en una palanca. Entre los dos ensayos hubo un cambio señalado y las tasas relativas de reforzamiento

cambiaron en razón de 4:1 en la otra opción. La transición a igualación se encontró acumulando los tiempos de cuando cada palanca era presionada y el número de refuerzos recibidos dentro de las ventanas temporales sucesivas (iguales en amplitud al intervalo esperado entre-reforzadores del programa pobre). Los resultados encontrados apoyan los datos de Dreyfus (1991), esto es, las ratas ajustaron la razón de distribución de tiempo ventana a ventana dentro de la primera ventana temporal del ensayo después de la transición (i.e., un intervalo igual a los dos intervalos entre-reforzadores en el programa pobre). Además, las variaciones ventana a ventana en la distribución de tiempo, siguieron las variaciones ventana a ventana no señaladas en la tasa de reforzamiento experimentada; resultados que son independientes de la escala temporal del experimento. Esto sugiere que, bajo las condiciones de dicho experimento en donde los sujetos esperaron cambios intrasesiones en la tasa relativa de refuerzo, no hay un periodo prolongado durante el cual los sujetos aprendan las nuevas contingencias de reforzamiento. Este resultado concuerda con otros hallazgos que también demuestran ajustes rápidos del comportamiento, en respuesta a cambios muy recientes en las contingencias de reforzamiento (Dreyfus, 1991; Kacelnik & Todd, 1992; Shettleworth et al., 1988; Shettleworth & Plowright, 1992). Estos ajustes rápidos implican que el proceso determinante de la elección que distribuye el tiempo del animal entre *manipulanda* (o sitios de forrajeo) es, bajo estas circunstancias, sensible primordialmente a un muestreo extremadamente local de las tasas de refuerzo más recientes. Probablemente, un aspecto importante de estas circunstancias es que los animales en el experimento de Mark y Gallistel (1994) y en los otros experimentos citados, habitualmente encuentran cambios intrasesión en la tasa relativa de refuerzo.

Los resultados obtenidos por Mark y Gallistel (1994) han sugerido que bajo algunas circunstancias, los ajustes extremadamente rápidos a fluctuaciones en la tasa de refuerzo ocurren durante elección concurrente. Esto sugiere que el almacén en la memoria puede ser sensible a las fluctuaciones de las ganancias de las tasas. Pudiera asumirse que para esta explicación de elección no se requiere almacenar mas que unos pocos intervalos entre-reforzadores (digamos 3 ó 5) para generar igualación. Bajo estas condiciones, la distribución local de tiempo sigue las fluctuaciones locales en la tasa de refuerzo. Parece probable que un ambiente fluctuante predisponga a los sujetos a recordar un número de demoras pequeño, mientras que ambientes estables pueden fomentar un almacén de memoria grande. Para entender completamente su mecanismo, necesitamos conocer la escala temporal en la cual opera la tasa de estimación (Mark & Gallistel, 1994).

Finalmente, todos los resultados experimentales discutidos sobre elección en periodos de transición, son inconsistentes con la mayoría de los modelos de equilibrio. Los modelos dinámicos, tales como el modelo de efectos acumulativos (Davis et al., 1993) o el modelo de modificación de H. B. Daly y J. T. Daly (1982), fallan en generar las predicciones correctas cuando se aplican a estos fenómenos.

Si la globalidad es una condición necesaria para explicar el aprendizaje, entonces ¿por qué Mark y Gallistel (1993) encontraron igualación en una escala temporal extremadamente local? En su experimento, la razón de distribución de tiempo siguió grandes variaciones aleatorias en la razón de refuerzo. Esta igualación no parece ser explicada por ningún proceso en el cual la distribución de tiempo previa del animal afecte recurrentemente su experiencia reciente, como los modelos globales sugerirían. De otra forma sería difícil de explicar

por qué los animales siguen el cambio en la tasa relativa de refuerzo, y además, por qué igualación resulta ser invariable en escala. Por otro lado, los modelos locales no ofrecen ninguna explicación para el papel de la historia de entrenamiento (experiencia) que parece estar claramente involucrada en el mecanismo de aprendizaje (véase por ejemplo la mejoría en la ejecución con el entrenamiento). Dragoi (1997) y Dragoi y Staddon (en prensa), sugieren como solución a este problema que el aprendizaje es determinado por procesos locales y globales, y proponen un modelo en tiempo real que comprende procesos en diferentes escalas temporales para explicar las propiedades estáticas y dinámicas del aprendizaje. El modelo asume que el organismo genera expectativas o predicciones de eventos futuros basados en el reforzamiento experimentado; estas expectativas son tanto de corto como de largo plazo, y están ligadas a través de un mecanismo integrativo que modula el proceso de decisión.

CAPÍTULO III

ADAPTACIONES CONDUCTUALES EN PROGRAMAS DE REFORZAMIENTO MÚLTIPLES

Cuando dos programas de reforzamiento intervalo variable (IV) son combinados, el responder mantenido por cada uno esta parcialmente determinado por los reforzadores distribuidos en el otro programa. Esta interacción ocurre tanto en programas concurrentes, en los cuales dos o más programas operan simultáneamente para respuestas en diferentes teclas; como en programas múltiples, en los cuáles dos o más programas simples son presentados sucesivamente en tiempo con un estímulo discriminativo asociado a cada uno (los periodos señalados por los estímulos se denominan componentes de un programa múltiple), y que operan para respuestas en una tecla simple. Sin embargo, esta similaridad entre programas concurrentes y múltiples no se extiende a la forma en la cual las respuestas son distribuídas entre cada componente de los programas que constituyen el programa total. En programas concurrentes IV IV, los animales igualan la distribución de respuestas entre teclas a la distribución de reforzadores entre teclas (Herrnstein, 1961; Catania, 1963). En programas múltiples IV IV estas distribuciones no son igualadas (Reynolds, 1963b).

Catania (1963) ha detallado una serie de ecuaciones que dan una descripción cuantitativa precisa de las relaciones entre tasas de respuesta y tasas de reforzamiento en programas concurrentes. Asimismo, existe una

descripción comparable de estas relaciones para programas múltiples (Herrnstein, 1970; Williams, 1988).

Las funciones empíricas que relacionan la tasa de respuesta mantenida por programas de intervalo variable son caracterizadas como funciones monótonicamente crecientes, que se aproximan a asíntota a tasas de reforzamiento relativamente bajas. Herrnstein (1970) demostró que los datos de programas IV, eran bien descritos por una función hiperbólica, como la que proporciona la ecuación 9, que surge como una descripción general del comportamiento obtenido en procedimientos de una sola respuesta, reconociendo la posibilidad de otras fuentes de reforzamiento:

$$P = k \frac{R}{R + R_o} \quad (9)$$

Donde P representa la tasa de respuesta, R la tasa de reforzamiento, R_o es el agregado de reforzadores desconocido para otras alternativas, y k es una constante determinada empíricamente.

Herrnstein (1970) argumentó que la función hiperbólica derivada de la ecuación 9, surge directamente de la ley de igualación, con la suposición de que la elección no es mas que un comportamiento colocado en el contexto de otro reforzamiento, y donde no hay forma de evitar dicho contexto para cualquier respuesta. El resultado es que el sujeto es confrontado con una elección continua entre dar la respuesta requerida y cualquier otra conducta que sea producida por el reforzamiento alternativo. Dada esta suposición, la ley de igualación debe aplicarse de tal forma que la ecuación de igualación puede ser escrita como:

$$\frac{P_1}{P_1 + P_2} = \frac{R_1}{R_1 + R_2} \quad (10)$$

Donde P es el número de respuestas, R es el número de reforzadores y los subíndices denotan las dos alternativas. Esta es la ecuación concurrente con dos respuestas registradas que resume el hallazgo de que las tasas relativas de respuesta igualan las frecuencias relativas de reforzamiento. Dicha ecuación asume que k y R_o son las mismas para las respuestas bajo observación, una suposición razonable mientras las respuestas sean equivalentes en forma y esfuerzo.

Por otra parte, si se hace la suposición adicional de que P_i y P_o agotan los comportamientos disponibles y por lo tanto son una constante, el resultado es:

$$\frac{P_1}{k} = \frac{R_1}{R_1 + R_o}$$

Arreglando los términos, produce entonces la ecuación 9. La interpretación de sus constantes determinadas empíricamente es que k se refiere a la cantidad total de respuestas que el sujeto puede emitir, y R_o a la tasa de reforzamiento disponible de otras fuentes.

Esta interpretación tiene profundas implicaciones para el concepto de fuerza de la respuesta, que asume que la fuerza de una respuesta individual nunca puede ser evaluada en aislamiento, porque cualquier evaluación empírica siempre refleja la fuerza relativa y es por lo tanto, dependiente del reforzamiento disponible para actividades alternativas. También asume que la cantidad total de conducta operante es una constante y es independiente de la

frecuencia de reforzamiento en la situación (esto no quiere decir que el valor de k será siempre el mismo, ya que éste varía con las unidades de medición empleadas). Por lo tanto, el concepto tradicional de fuerza de la respuesta es equivocado, puesto que el efecto de reforzamiento es solo seleccionar, no fortalecer en su sentido habitual (Williams, 1988).

Programas de Reforzamiento Múltiples y la Ley de Igualación: Interacciones a Distancia

Para programas concurrentes, los reforzadores en un programa ocurren en el contexto de todos los demás reforzadores, con igual peso (ecuación 11). En otras palabras, cuando todos los reforzadores están disponibles simultáneamente, la interacción debida a los reforzadores de otras fuentes es máxima.

$$P_1 = k \frac{R_1}{R_1 + R_2 + R_o} \quad (11)$$

Herrnstein (1970) extendió su modelo de igualación para explicar la ejecución en programas múltiples. Hizo las mismas suposiciones básicas como aquellas para su aproximación teórica a programas simples (ecuación 9) y programas concurrentes (ecuación 11), pero agregando un parámetro adicional. Este parámetro es m , que describe el grado de interacción entre los componentes de un programa múltiple. Su valor se encuentra entre 0 (no interacción entre componentes) y 1 (máxima interacción). En programas múltiples, las varias fuentes de reforzamiento no están operando simultáneamente. Conforme los componentes de un programa múltiple están

más separados, es probable que la interacción entre los componentes disminuya; en el caso límite, disminuye a no interacción.

Hay varias formas de traducir este continuo de interacciones en una expresión formal, pero una de las más simples es asumir que el reforzamiento en un componente afecta el responder en el otro componente por alguna fracción constante de su efecto total. Entonces, para los programas múltiples de dos componentes, la tasa de respuesta estaría dada por¹:

$$P_1 = k \frac{R_1}{R_1 + mR_2 + R_0} \quad (12)$$

La forma de la función depende conjuntamente del grado de interacción (m) y de R_0 . Si R_0 y el término de interacción (m) se van a cero, entonces la tasa relativa de respuesta durante un componente será insensible a los cambios en la tasa relativa de reforzamiento. Para cualquier valor de m , la función es más empinada con valores más grandes de R_0 . Cuando m es menor que 1.0, la función se aproxima asintóticamente a la línea de igualación conforme R_0 incrementa. El denominador de la ecuación 12 representa el contexto total del refuerzo en el cual un evento de reforzamiento ocurre, y los efectos de fortalecimiento de la respuesta de ese evento están inversamente relacionados con el valor de dicho contexto.

¹ Note que la base de tiempo para calcular la tasa de respuesta en un programa múltiple es la duración del componente durante el cual las respuestas pueden ocurrir. Para programas concurrentes, la base de tiempo es la sesión experimental completa. En general, la base de tiempo para calcular la tasa de respuesta debe ser el tiempo durante el cual las respuestas pueden ocurrir, lo cual difiere, como fue indicado, para programas múltiples y programas concurrentes. También note que la cantidad R_0 en la ecuación 12, representa un compuesto de los reforzadores externos durante un componente más los reforzadores externos en el otro componente.

En programas múltiples (ecuación 12), la influencia de la tasa de refuerzo en el componente alterno es menor que el máximo nivel ($m < 1$), porque esos reforzadores están ocurriendo en otros tiempos que se presume que son altamente discriminables. Asimismo, si el componente alterno esta temporalmente distante del componente en cuestión, m debería de ser cero, y la tasa de respuesta en ese componente sería parecida a aquella de un programa simple como lo describe la ecuación 9. De hecho, la ecuación 11 puede ser considerada como un caso especial de la ecuación 12 donde $m=1$.

La ecuación 12 no solo asume un simple mecanismo para gobernar el grado de interacción, sino que describe tácitamente el tipo más simétrico de programas múltiples, como se muestra por la ausencia de subíndices en los parámetros k , m , y R_o . Esto es equivalente a asumir que las alternativas tienen tasas asintóticas iguales (k), que la interacción en una dirección es la misma que la interacción en la otra (m), y que el reforzamiento no programado es el mismo durante los dos componentes (R_o). En otras palabras, para un programa múltiple con la misma forma de la respuesta en ambos componentes (k), con la misma regla de alternación de un componente a otro componente así como la alternación en sentido inverso (m), y con los factores extra-experimentales constantes (R_o).

Las tasas relativas de respuesta pueden ser derivadas de la ecuación 12, escribiendo la expresión correspondiente para ambos componentes del programa múltiple. Entonces,

$$\frac{P_1}{P_2} = \frac{\frac{kR_1}{R_1 + mR_2 + R_o}}{\frac{kR_2}{R_2 + mR_1 + R_o}} \quad (13)$$

La ecuación 13 hace explícito que la aproximación de igualación de las tasas relativas de respuesta a las tasas relativas de reforzamiento en un programa múltiple, es una función directa de la similaridad entre el contexto de reforzamiento para los dos componentes. La aproximación a igualación incrementará cuando m incremente o el valor de R_0 sea grande, relativo a R_1 y a R_2 . Evidencia para la primera de estas predicciones surge con los efectos de la duración del componente (Charman & Davison, 1982). Evidencia para la segunda predicción se obtiene de estudios como el de Nevin (1974), en el cual la tasa de reforzamiento adicional varió mientras se mantuvieron constantes las tasas de reforzamiento durante los componentes de los programas múltiples (e.g., presentando reforzamiento suplementario durante periodos de tiempo fuera, separando los componentes).

Contraste Conductual

El cambio en la tasa de respuesta de un pichón en un componente de un programa múltiple que resulta del cambio en un segundo componente, ha sido llamada una *interacción* (Herrnstein & Brady, 1958; Reynolds, 1961a). Reynolds (1961a) fue el primero en demostrar que la tasa de respuesta en el primer componente de un programa múltiple depende no solo de la frecuencia de reforzamiento en ese componente, sino también del reforzamiento en el segundo componente. Estas interacciones entre programas de reforzamiento han sido denominadas *Contraste Conductual* (Catania, 1961; McSweeney, Farmer, Dougan & Whipple, 1986; Williams & Wixted, 1986) y ocurren cuando los dos programas están simultáneamente disponibles (programas

concurrentes: Herrnstein, 1970) y cuando son alternados sucesivamente (programas múltiples: Reynolds, 1961a).

Reynolds (1961b, 1963b) demostró que el cambio en la tasa relativa de reforzamiento en cada componente del programa es el principal determinante del contraste estable. El efecto de contraste parece ser producido por la introducción de un estímulo correlacionado con extinción, o cambiando un componente de intervalo variable (IV) a extinción (EXT) o tiempo fuera (Reynolds, 1961a, 1961b; Nevin & Shettleworth, 1966). También puede ser producido por manipulaciones sistemáticas del valor de una razón fija (RF) (Reynolds, 1961c) o de un programa de intervalo variable (Reynolds, 1963b). Una característica importante de los efectos de contraste es que no dependen del programa con el cual el reforzamiento alternativo es presentado (Williams, 1988).

Los resultados de Lander e Irwin (1968), confirman los resultados de Reynolds (1963b) de que los pichones expuestos a programas múltiples IV IV no igualan el número relativo de respuestas en un componente, al número relativo de reforzadores en ese componente. Al mismo tiempo, sugieren que el número de reforzadores afecta sistemáticamente el número de respuestas emitidas en estos programas.

Existe evidencia (Williams, 1983) que apoya la idea de que la tasa relativa de reforzamiento es una variable fundamental, sin tomar en cuenta las preparaciones experimentales particulares. Entonces, las interacciones observadas en programas múltiples son similares a aquellas en programas concurrentes, a pesar de que las dos opciones de respuesta nunca están simultáneamente disponibles.

Teoría de Competencia Conductual

Una aproximación alternativa al contraste conductual, basada directamente en igualación en programas concurrentes, ha sido propuesta por Staddon (1982). Staddon plantea que el comportamiento durante cualquier componente de un programa múltiple puede ser entendido en términos de la ley de igualación, y que la respuesta medida (P_I) y el comportamiento alternativo que no es medido por el experimentador (P_o), constituyen la conducta total. Entonces, la ecuación 9 se aplica también para programas múltiples.

$$P_I = k \frac{R_I}{R_I + R_o}$$

Para que el reforzamiento en el componente alternativo tenga un efecto, debe ser mediado por cambios en el valor del reforzamiento alternativo, R_o . Staddon (1982) sugiere un mecanismo para esta mediación, en términos de saciedad transitoria y efectos de privación con respecto a las actividades que generan R_o . La noción básica es que contraste en programas múltiples se deriva de la competencia entre comportamiento interino y terminal, los cuales son producidos por programas de reforzamiento de intervalo. Esto es, durante ambos componentes del programa múltiple IV IV, tanto comportamientos interinos como terminales ocurren y por lo tanto compiten mutuamente. Cuando el programa se cambia a múltiple IV EXT, ambos tipos de comportamiento se eliminan en el componente EXT. El comportamiento interino que efectivamente ocurre en el IV, puede moverse sin cambios al componente EXT dejando más tiempo para las respuestas terminales en el IV que permanece constante, de tal forma que ocurre un incremento en la tasa de respuesta. Entonces, contraste en

programas múltiples como en programas concurrentes, es debido primordialmente a un cambio en la distribución de tiempo, con un mayor porcentaje de tiempo dedicado al comportamiento terminal (picoteo) en un programa múltiple IV EXT, que en un programa múltiple IV IV (para el componente IV que permanece constante).

Los modelos de contraste de competencia conductual son puestos a prueba por un problema fundamental. La noción de competencia conductual opone la respuesta P contra P_o , entonces el contraste debe depender de la frecuencia de respuesta en el componente variable, no de la frecuencia de reforzamiento. Pero el comportamiento actual en el componente variable parece irrelevante, si la variable determinante es la tasa relativa de reforzamiento. Esto es mejor ejemplificado cuando un programa múltiple IV TV, o un programa múltiple IV IV señalado, se cambia a un programa múltiple IV EXT. Ya que se requiere un comportamiento operante mínimo tanto para el programa TV como para el programa IV señalado, esto significa que debe de haber poca competencia de P_o durante la línea base. Por lo tanto, el cambio a extinción debe causar un cambio pequeño en la frecuencia de P_o durante el componente variable y de igual forma, debe ocurrir un cambio simple en el componente constante. El grado de contraste no varía con manipulaciones que deben afectar el grado de competencia de respuesta. La implicación es que la elección, conceptualizada como competencia conductual, no puede explicar las interacciones observadas en programas múltiples; por lo tanto la relación entre tales interacciones e igualación en programas concurrentes, permanece incierta (Williams, 1988).

Asimismo, la ejecución en programas múltiples es de gran importancia para la interpretación teórica de contraste, puesto que la competencia directa

entre alternativas de respuesta no puede ser el mediador de contraste cuando las alternativas nunca están simultáneamente disponibles. Consecuentemente, los efectos de contraste en programas múltiples deben reflejar el efecto relativo del reforzamiento a su nivel más fundamental (Williams, 1983).

Efectos Generales del Contexto de Reforzamiento

Debido a que el marco de la ley de igualación no puede explicar los efectos de contraste en programas múltiples, existen explicaciones alternativas que intentan resolver este problema. Una noción simple, es la suposición de que la tasa de respuesta esta inversamente relacionada al contexto de reforzamiento en el cual ésta ocurre. En programas concurrentes, la tasa de respuesta mantenida por un programa IV constante está inversamente relacionada a la frecuencia (o cantidad) del reforzamiento alternativo en la misma situación, independiente de si el reforzamiento adicional genere cualquier conducta operante por sí mismo. Estos procedimientos difieren de los programas múltiples solo en que el reforzamiento alternativo ocurre mientras la respuesta operante esta inmediatamente disponible; en programas múltiples el reforzamiento alternativo ocurre en un tiempo diferente y en la presencia de un estímulo diferente.

El contexto del reforzamiento no puede ser explicado simplemente en términos de la tasa de reforzamiento promedio en la situación, tal y como Herrnstein (1970) propuso. La razón es que el contexto del reforzamiento, al menos en programas múltiples, es temporalmente asimétrico en sus efectos. Más aun, la dirección de la asimetría es justo la opuesta a la que ha sido propuesta comúnmente. Varios experimentos han mostrado que variaciones en

la tasa de reforzamiento durante el componente sucesivo de un programa múltiple, produce efectos más fuertes que la variación de reforzamiento en el componente precedente (Williams, 1976b, 1979, 1981). Esto implica que el contexto del reforzamiento no puede ser definido simplemente en términos del reforzamiento total (o promedio) en la situación. En lugar de ello, contraste es considerado como un efecto anticipatorio en el que la tasa de respuesta en el componente señalado está inversamente relacionada a la tasa de reforzamiento que está próxima en el componente sucesivo.

La importancia del componente sucesivo en programas de reforzamiento múltiples ha sido conocida desde hace varios años, siendo estudiada sistemáticamente por primera vez por Wilton y Gay (1969), quienes demostraron un incremento en la tasa de respuesta durante una señal precediendo un cambio a extinción. En su experimento, ocurrió un incremento y un decremento en la frecuencia de reforzamiento. El programa IV 5 min en un componente se cambió a un programa IV 1 min, y el programa IV 5 min en el otro componente, se cambió a extinción. La tasa de respuesta en el componente IV 5 min que precedía el componente de extinción fue típicamente más alta que la tasa de respuesta en el componente IV 5 min que precedía el componente IV 1 min. El hallazgo de que la tasa de respuesta en un componente varía como una función del programa operando en el componente que le sigue inmediatamente, apoya los datos obtenidos en otros estudios (Williams, 1976b, 1979, 1988). Estos resultados son de particular relevancia para los estudios que han investigado contraste en programas múltiples de dos componentes (Reynolds, 1961, 1963). Ellos sugieren que el contraste puede variar como una función del ordenamiento de los componentes.

Williams (1981, Experimento 2) presentó a un grupo de pichones un programa de tres componentes en el cual la tasa de reforzamiento durante el componente de en medio varió, mientras que los programas durante los componentes primero y tercero fueron constantes e iguales. El hallazgo fundamental fue que la variación del reforzamiento durante el componente de en medio produjo efectos de contraste más fuertes durante el primer componente que durante el tercer componente.

Dado que las interacciones en estos programas son temporalmente asimétricas, esto sugiere que igualación no puede ocurrir en programas múltiples con más de dos componentes. Por ejemplo, en el programa de tres componentes empleado por Williams (1981), igualación predice que la tasa relativa de respuesta para un componente específico debe ser determinada por su tasa relativa de reforzamiento a la tasa total de reforzamiento, a lo largo de todos los componentes. Pero esta predicción es criticada por los fuertes efectos de orden temporal que ocurrieron.

Los resultados obtenidos por Williams (1976c) muestran que los efectos separados de cambios locales en el reforzamiento operaron simultáneamente. Esto es, las respuestas durante un componente incrementaron cuando el componente siguió a un periodo de extinción y cuando precedía el periodo de extinción.

Nevin y Shettleworth (1966) registraron la tasa de respuesta en subintervalos sucesivos de los componentes de programas múltiples, es decir, mostraron las tasas de respuesta a diferentes distancias temporales de la alternación del componente. Cuando los componentes del programa distribuyeron el reforzamiento a diferentes tasas, las tasas de respuesta en los componentes de valor más alto fueron inicialmente altas y decrecieron a un

nivel asintótico hacia el final del componente. En componentes de bajo valor, las tasas de respuesta fueron inicialmente bajas e incrementaron hacia asintota. Estos efectos han sido llamados *contraste local positivo y negativo* respectivamente, y han sido reportados por Menlove (1975) y Rachlin (1973), entre otros.

Los efectos asimétricos del contexto de reforzamiento sugieren la posibilidad de una explicación más molecular, pero los intentos previos por descubrirla no han sido exitosos (Williams, 1979 Experimentos 2 al 4).

Dinámica de la Elección Sucesiva

Uno de los problemas fundamentales en el estudio de la elección ha sido la relación entre procedimientos en los cuales la elección es simultánea y aquellos en que la elección es sucesiva (discriminación simultánea y sucesiva). Al igual que en un programa concurrente, en programas múltiples la conducta en un componente depende de lo que ocurre en el otro componente.

Dada la simultaneidad de las opciones de respuesta en los programas concurrentes, el tiempo dedicado a responder a una opción hace que se reste el tiempo dedicado a responder a otra opción. La competencia entre respuestas por el tiempo disponible, permite que el incremento en la fuerza de una respuesta dé lugar a decrementos en las tasas de otras respuestas. Para poder determinar si el refuerzo determina la fuerza de solo una respuesta sobre la cual es contingente, o por el contrario, el animal aprende un contexto global en el que compara cada refuerzo recibido (Williams, 1994), es necesario estudiar la adquisición del comportamiento en situaciones de elección sucesiva.

Para estudiar la posible interacción entre reforzadores, es preciso separar las respuestas de manera tal 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 maneja en el presente trabajo, es empleando programas múltiples. Con este tipo de programas, las respuestas a diferentes opciones se emiten en componentes temporalmente separados, de forma tal que no pueden competir por un tiempo común.

En el contexto de la elección recurrente en operante libre, no existe evidencia empírica de la trayectoria que sigue el comportamiento hacia el equilibrio en situaciones intertemporales, en otras palabras, se desconoce la dinámica de los ajustes a cambios rápidos en las distribuciones de refuerzo en situaciones de elección sucesiva. El presente trabajo es una primera aproximación para abordar este problema, y la propuesta fue estudiar este fenómeno de la misma forma como se ha estudiado el caso de la elección simultánea.

Se utilizó el procedimiento del bandido de dos brazos secuencial, en el cual las respuestas a un componente de un programa múltiple reciben reforzamiento de acuerdo a un programa de razón variable 30, mientras que las respuestas en el otro componente no reciben reforzamiento (Múlt. RV30-EXT).

En el primer experimento, se deseaba estudiar qué adaptación sería más sencilla para el animal, si adaptarse a cambios señalados en las condiciones de reforzamiento entre sesiones o dentro de una misma sesión experimental. Para ello, se inició con la condición más elemental, que correspondía en alternar qué componente estaría asociado con el refuerzo cada día.

A partir de los resultados obtenidos en el experimento 1, en el experimento 2 decidimos estudiar qué tan sencillas serían las adaptaciones a

cambios aleatorios entre sesiones y cambios aleatorios señalados en la segunda mitad de la sesión experimental.

En el tercer experimento se estudió el otro extremo del continuo, en donde los cambios en las condiciones de refuerzo eran completamente impredecibles. A diferencia del experimento 2, en el cual las condiciones al inicio de la sesión eran las mismas y solo variaba a la mitad de la sesión, en este experimento, los cambios aleatorios ocurrieron desde el inicio de la sesión y a la mitad de la sesión se decidió aleatoriamente si continuaría así por el resto de la sesión.

Se realizó un análisis global de la trayectoria del comportamiento ante los cambios en la tasa relativa de refuerzo, día a día y dentro de una sesión experimental. Un segundo análisis, consistió en el análisis de la trayectoria en los cambios componente a componente durante toda la sesión experimental en cada una de las condiciones experimentales.

CAPÍTULO IV

TRABAJO EXPERIMENTAL

Método

Sujetos

Los sujetos experimentales fueron 12 pichones macho adultos (obtenidos del bioterio del laboratorio de Comportamiento y Adaptación, de la Facultad de Psicología, UNAM), mantenidos al 80% de su peso corporal *ad libitum* y con acceso libre a agua y grit cárico. Los sujetos recibieron la mayor parte de la comida durante la sesión experimental, pero también se les proporcionaba comida adicional en caso necesario para mantenerlos en el peso criterio. Todos los sujetos contaban con historia previa de entrenamiento en programas múltiples.

Aparato

Se emplearon 6 cajas de Skinner con las siguientes dimensiones: 37 cm. de alto, 30 cm. de ancho y 35 cm. de fondo. Tenían una rejilla de 5 cm. de ancho colocada en el piso de la caja. En la pared derecha de la cámara, a 21 cm. del piso, se encontraban tres teclas de respuesta separadas por 7 cm; cada una con 2.7 cm. de diámetro.

Las teclas podían ser iluminadas por una luz roja o verde producidas por uno de dos focos de 2 Wts. fijados detrás de cada tecla. Se requirió de una

fuerza de .15N para operar cada tecla. Para el presente trabajo sólo se utilizó la tecla izquierda.

El comedero, que presentaba una mezcla de grano balanceado por 2.5 seg, estaba situado a 5.5 cm. del piso y a 10 cm. debajo de la tecla central. Cada vez que el grano era presentado, el comedero se iluminaba por una luz blanca proveniente de los dos focos de 2 Wts. fijados sobre éste. Asimismo, una luz general de 2 Wts. estaba ubicada al centro del techo de la caja. Cada cámara experimental estaba equipada con un extractor de aire que a su vez servía como generador de ruido blanco, para atenuar el sonido proveniente del exterior.

Las cámaras experimentales estaban localizadas en un cuarto separado del equipo de control y registro. El control de los estímulos y el registro de las respuestas se hicieron con una computadora personal programada en lenguaje MED-PC, y conectada a una interfase MED-PC.

Procedimiento General

Los sujetos fueron alojados en jaulas habitación individuales teniendo los primeros 30 días acceso libre tanto al agua como al alimento (mezcla de grano balanceado) y grit cárico. Durante este periodo se registró diariamente el peso de los animales, tomando la media de este registro como punto de comparación posterior. A continuación, se restringió paulatinamente la cantidad de alimento disponible hasta alcanzar el criterio del 80% (± 5 gr.) de peso con respecto al valor *ad libitum*.

Previo al inicio del experimento, los pichones fueron sometidos a un programa de automoldeamiento, utilizando las luces roja y verde de la tecla izquierda.

A cada grupo se le asignaron al azar seis sujetos y el orden de presentación de las fases experimentales se contrabalanceó entre los grupos. Todos los sujetos fueron expuestos a programas múltiples RV30-EXT. Los componentes del programa múltiple estaban señalados por la iluminación roja o verde de la tecla izquierda de respuesta. La duración de cada componente fue de 30 seg y las sesiones diarias terminaban después de 30 exposiciones a cada componente del programa múltiple.

Para todos los experimentos, al inicio de cada sesión, la luz general se encendía y la tecla izquierda de respuesta era iluminada. Cada día se asignaba aleatoriamente el color con que iniciaba la sesión experimental; y en la condición intrasesión, a la mitad de la sesión experimental, aleatoriamente se asignaba el color de inicio en la transición. La luz general permanecía encendida durante la sesión completa, excepto durante los periodos de reforzamiento. Cada picoteo efectivo en la tecla derecha producía 0.1 seg de oscuridad en ésta como retroalimentación. Si una respuesta tenía como consecuencia reforzamiento, la luz general y la luz de la tecla se apagaban y el comedero era iluminado y levantado para presentar el grano por 2.5 seg. Después de la entrega del reforzador, la luz general y la luz de la tecla eran nuevamente iluminadas. Al final de la sesión, todas las luces de la cámara se apagaban.

Experimento 1.

Cambio Alternativo Diario y Dentro de Una Misma Sesión Experimental

En este experimento, el propósito principal fue analizar qué adaptación sería más sencilla para los pichones en situaciones de elección sucesiva: Adaptarse a cambios señalados en las condiciones de reforzamiento entre sesiones, o dentro de una misma sesión experimental. Para ello, se inició con la condición más simple, que correspondía en alternar qué componente estaría asociado con el refuerzo cada día.

Se utilizará el término *componente rico* para designar al componente asociado al programa de razón, y *componente pobre* para señalar al componente asociado al programa de extinción.

Procedimiento

En el experimento 1, todos los sujetos fueron expuestos a programas múltiples RV30-EXT y el orden de presentación de las fases se contrabalanceó entre grupos. Durante el primer ciclo, para un grupo, el componente (color) asociado al programa de razón alternaba diariamente (véase Apéndice A, Experimento 1).

Para el segundo grupo, los cambios del componente asociado con el programa de razón alternaban dentro de una misma sesión (véase Apéndice A, Experimento 1). En esta condición, en la primera mitad de la sesión, los reforzadores eran programados para las respuestas al componente rojo. Al inicio de la segunda mitad de la sesión experimental hubo 3 segundos de tiempo fuera, que señalaban el cambio en la distribución de reforzadores, y los

reforzadores fueron programados para las respuestas efectivas en el componente verde, continuando así hasta el final de la sesión.

Una vez que la ejecución en ambos grupos experimentales fue estable, se inició la fase de extinción, en la cual las respuestas en ambos componentes del programa múltiple nunca conducían a reforzamiento (Múlt. EXT-EXT). Esta fase comenzó para la condición entre sesión, con el color del componente que ese día estaría asociado con reforzamiento (véase Apéndice A, Múlt. EXT-EXT Alternó Ciclo 1).

Después de la fase de extinción, todos los pichones fueron expuestos a programas Múlt. RV30-RV30 hasta que sus respuestas en los dos componentes no presentaron variaciones sistemáticas de sesión a sesión.

Al inicio del segundo ciclo del programa Múlt. RV30-EXT, para el grupo que en el ciclo 1 había experimentado cambios dentro de una sesión experimental, el componente (color) asociado al programa de razón alternó diariamente (véase Apéndice A, Experimento 1). Para el grupo que en el ciclo 1 había experimentado cambios entre sesión, los cambios del componente asociado con el programa de razón alternaron dentro de una misma sesión (véase Apéndice A, Experimento 1).

Cuando la ejecución en ambos grupos alcanzó niveles asintóticos, se dió inicio a la última fase del experimento, que constituía la fase de extinción (ésta fue exactamente igual a la fase de extinción presentada en el ciclo 1 (Múlt. EXT-EXT), y comenzó para condición entre sesión, con el color del componente que ese día estaría asociado con reforzamiento (véase Apéndice A, Múlt. EXT-EXT Alternó Ciclo 2).

En la Tabla 1 se describe el orden de presentación de las condiciones experimentales para cada grupo.

TABLA 1. Orden de presentación de las condiciones experimentales para cada grupo, en el experimento de cambio alterno diario y cambio alterno intrasesión.

FASES EXPERIMENTALES					
Grupo 1	Entre sesiones	Extinción	Línea Base	Intrasesiones	Extinción
	Múltiple	Múltiple	Múltiple	Múltiple	Múltiple
	RV30-EXT	EXT-EXT	RV30-RV30	RV30-EXT	EXT-EXT
	Ciclo 1			Ciclo 2	
Grupo 2	Intrasesiones	Extinción	Línea Base	Entre sesiones	Extinción
	Múltiple	Múltiple	Múltiple	Múltiple	Múltiple
	RV30-EXT	EXT-EXT	RV30-RV30	RV30-EXT	EXT-EXT
	Ciclo 1			Ciclo 2	

Resultados

Nuestro propósito principal en este experimento fue analizar la adaptación a cambios diarios alternos y dentro de una misma sesión experimental, en los parámetros de reforzamiento en programas múltiples RV30-EXT.

Los resultados principales se muestran en la Figura 1, donde las tasas absolutas de respuesta en cada componente del programa múltiple, son promediadas sobre los 6 sujetos para cada ciclo del cambio entre sesión. Cada punto graficado representa las respuestas diarias al componente rico y al componente pobre. El grupo de funciones promedio en la Figura 1 fue representativo de las funciones para los sujetos individuales.

El panel izquierdo presenta la ejecución en el Ciclo 1. Después de fluctuaciones iniciales en las cuales los sujetos responden a un solo color con cierta preferencia, las funciones se separan a partir de la sesión 13 y se establecen a un nivel asintótico que indica que los animales empezaron a detectar rápidamente cuál era la tecla rica en cada día.

Al inicio del experimento, no se corrieron sesiones de línea base como se hizo antes del inicio del segundo ciclo. En el Ciclo 2 (panel derecho), los cuatro primeros puntos graficados representan el promedio de la tasa de respuesta en ambos componentes en línea base (Múlt. RV30-RV30). Al inicio del entrenamiento se observó contraste conductual, es decir, la tasa de respuesta al componente rico incrementó cuando comenzó el segundo ciclo.

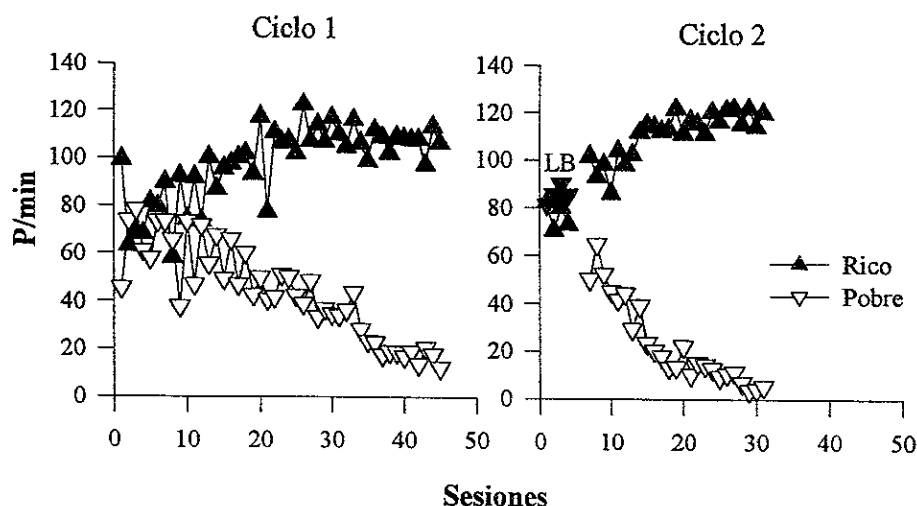


Fig.1. Tasas absolutas promedio en cada componente del programa múltiple, durante la condición de cambio alterno entre sesión. Ciclo 1 (panel izquierdo) y Ciclo 2 (panel derecho).

El grupo de pichones en el ciclo 2 rápidamente adquirió un patrón de alternación, respondiendo al componente rico a una tasa alta y al componente pobre a una tasa baja, resultados muy similares a aquellos encontrados en experimentos previos (Davis & Staddon, 1990). De esta forma, los ajustes del comportamiento de los pichones ante los cambios diarios subsecuentes, alcanzaron niveles asintóticos más rápidamente que en el ciclo-1.

Las tasas absolutas de respuesta promedio antes y después de la transición en la condición del cambio intrasesión en ambos ciclos, se presentan en la Figura 2.

En el ciclo 1, antes de la transición se observó un rápido ajuste una vez transcurridas las primeras 10 sesiones. Después de la transición, los animales tardaron aproximadamente 4 días en incorporar los cambios intrasesión en su distribución de comportamiento.

Los primeros cuatro puntos graficados en el ciclo 2 antes de la transición, representan el promedio de las tasas absolutas de respuesta en los últimos días en línea base. Asimismo, se observó un efecto de contraste conductual muy similar a aquel encontrado en la condición de cambio entre sesiones. Para efectos ilustrativos, se graficó en la segunda mitad de la sesión del ciclo 2, los 4 últimos días en línea base, para establecer la comparación entre las tasas base y las tasas después de la transición. La Figura 2 hace evidente que no hubo efectos de presentación de las fases experimentales, ya que la ejecución en ambos ciclos es muy similar. Además, es importante señalar que el ajuste en cada sesión fue más rápido en el cambio intrasesión que el observado en el cambio entre sesión, en ambos ciclos.

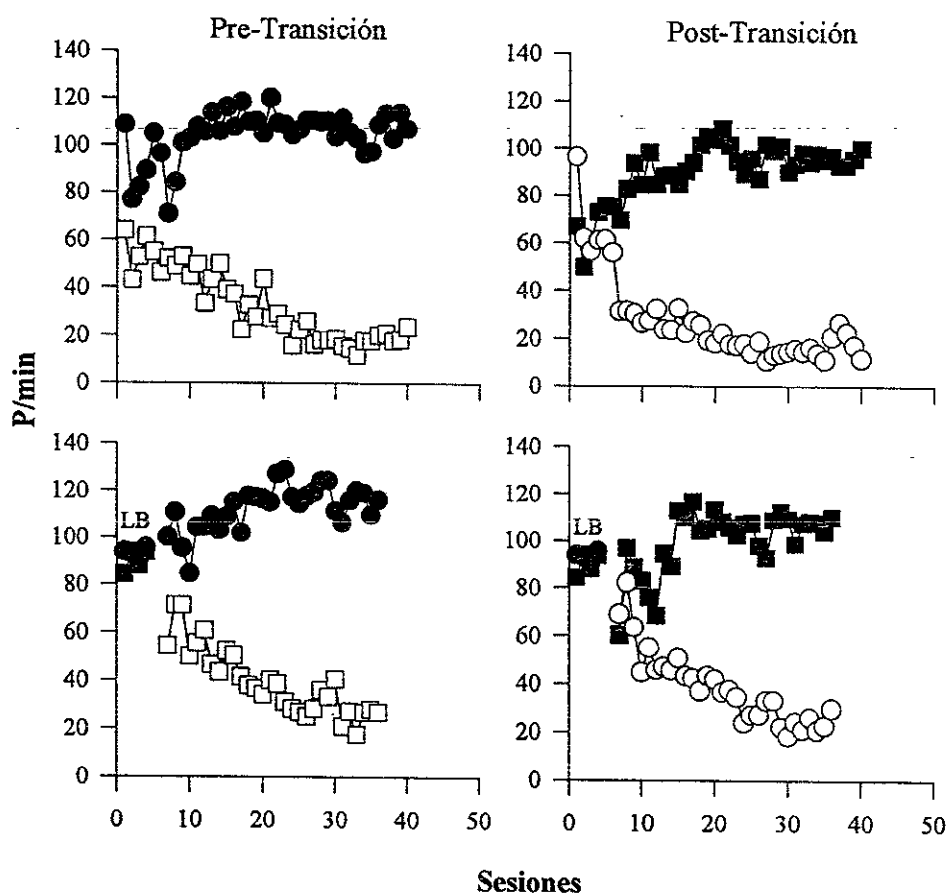


Fig. 2. Tasas absolutas promedio en cada componente del programa múltiple antes y después de la transición, durante la condición de cambio alterno intrasesión. Ciclo 1 (paneles superiores) y Ciclo 2 (paneles inferiores). Las respuestas al componente rico están indicadas por símbolos cerrados, y las respuestas al componente pobre por símbolos abiertos; cuando el componente fue rojo (círculos) ó verde (cuadrados).

La Figura 3 presenta las tasas absolutas de respuesta promedio durante los componentes del programa múltiple, al final del entrenamiento (panel superior), y en los dos primeros días de extinción (paneles inferiores), en el cambio entre sesión (panel izquierdo), y cambio intrasesión (panel derecho) del ciclo 1.

En el último día de entrenamiento en la condición de cambio entre sesión, se distingue el ajuste desde el primer componente de la sesión. Como se puede observar, en la condición de cambio intrasesión la elección del color de cada componente cambió hacia los últimos componentes anteriores a la transición, reflejándose en un ligero decremento en la tasa de respuesta en el componente rojo, y un incremento en la tasa de respuesta en el componente verde. Después de la transición, se aprecia un ajuste más lento en los primeros componentes.

En extinción, las respuestas al componente que en ese día particular correspondiera al componente rico, están representadas con símbolos abiertos con una cruz en medio.

En la condición entre sesión hay un patrón de alternación al componente que estaría asociado a reforzamiento, desde el último día de entrenamiento y durante los dos siguientes días en extinción. En la condición intrasesión, antes de la transición, hay una inversión en las distribuciones de respuestas para cada componente hacia el segundo tercio, esto es, el incremento inicial en la tasa de respuesta en el componente rico fue reemplazado por un decremento gradual y un incremento en el componente pobre en los últimos componentes. En el segundo día de extinción, este efecto se ve disminuido y solo se presenta en pocos componentes anteriores a la transición.

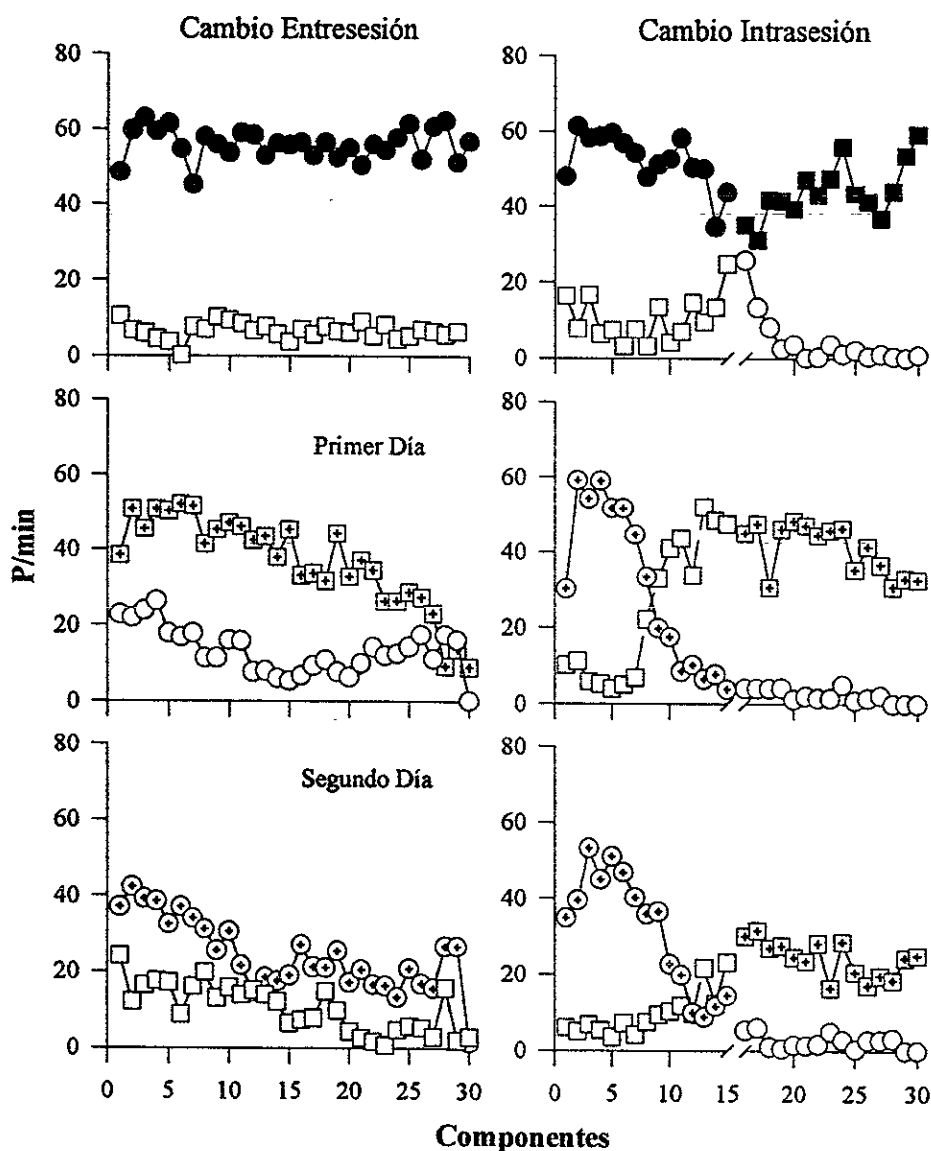


Fig. 3. Tasas de respuesta promedio al final del entrenamiento (panel superior), y en los dos primeros días de extinción (paneles inferiores), durante el cambio alterno entre sesión (panel izquierdo) y cambio alterno intrasesión (panel derecho) del Ciclo 1. Ver detalles en el texto.

Del mismo modo, la Figura 4 ilustra las tasas de respuesta en los componentes del programa múltiple al final del entrenamiento (panel superior), y en los dos siguientes días en extinción en el ciclo 2 (paneles inferiores), durante en el cambio entre sesión (panel izquierdo) y cambio intrasesión (panel derecho). Excepto por un número de respuestas ligeramente mayor a cada uno de los componentes del programa múltiple antes y después de la transición, el patrón de respuestas en ambos componentes en el ciclo 2, fue similar al observado en el ciclo 1. Estos resultados apoyan la idea de que los efectos de orden fueron prácticamente nulos.

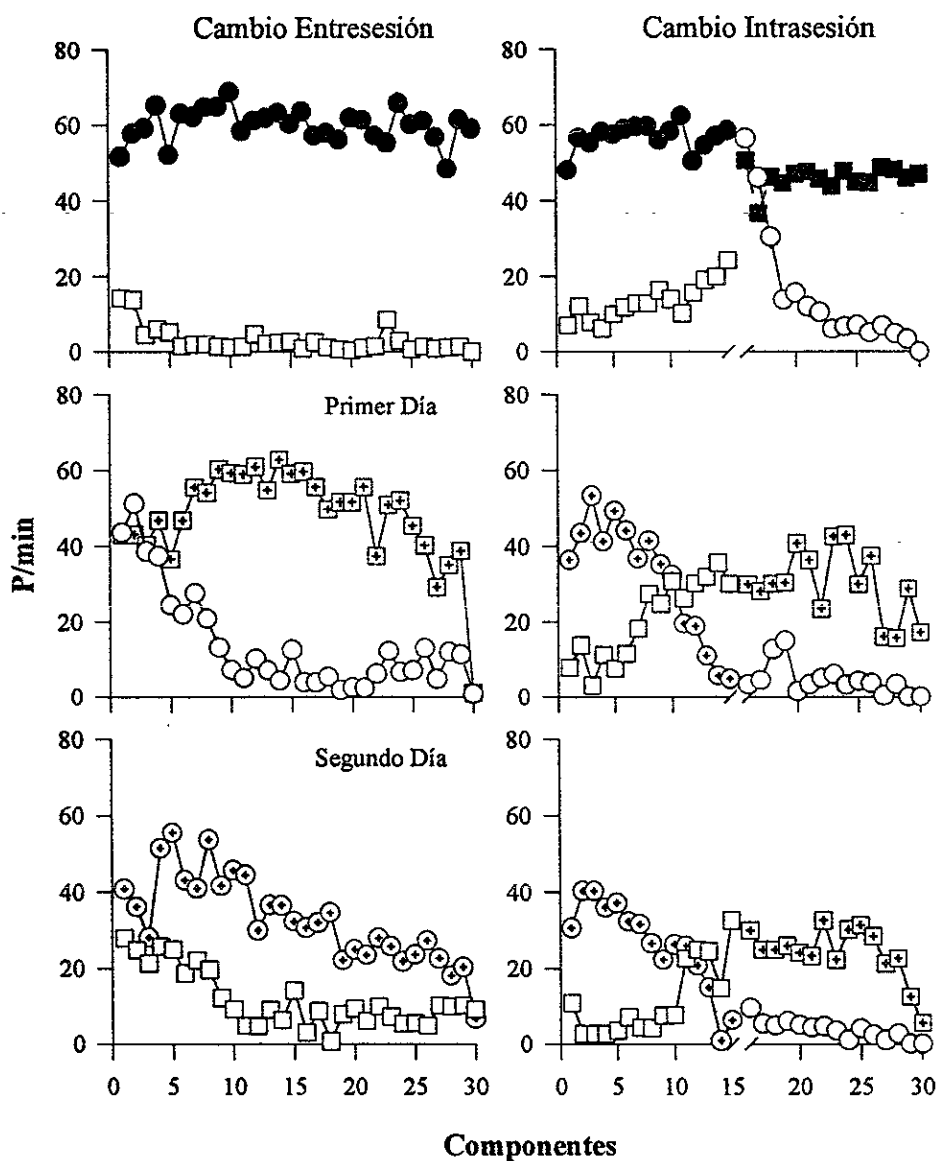


Fig. 4. Tasas de respuesta promedio al final del entrenamiento (panel superior), y en los dos primeros días de extinción (paneles inferiores), durante el cambio alterno entre sesión (panel izquierdo) y cambio alterno intrasesión (panel derecho) del Ciclo 2. Ver detalles en el texto.

Discusión

El presente experimento, evaluó las adaptaciones conductuales en situaciones de elección sucesiva durante periodos de transición, cuando los parámetros de reforzamiento variaron en forma alterna día a día y en una misma sesión experimental.

A pesar del hecho de que los animales antes de ser incluidos en el experimento ya tenían una buena historia de entrenamiento en programas múltiples, mostraron una mejoría progresiva en su ejecución en ambas condiciones, conforme incrementó el número de sesiones.

El patrón entre sesión de tasas absolutas observado, es muy parecido al que ha sido descrito en procedimientos de discriminación sucesiva con ensayos discretos (Woodward & Bitterman, 1976; Papini, 1997). La ejecución en ambos ciclos muestra que no hubo ningún efecto de presentación de las fases experimentales, y que los animales se ajustaron a las contingencias de reforzamiento actuales. En este aspecto, estos resultados coinciden con los hallazgos reportados en la literatura concurrente (e.g., Shettleworth & Plowright, 1992), ya que demuestran que los pichones fueron sensibles a cambios rápidos en su entorno, y utilizaron una regla adaptativa para cada situación experimental.

En la condición del cambio intrasesión, los patrones de anticipación al responder al componente previamente pobre observados durante extinción, sugieren que los animales estimaron la duración de la sesión, en donde las respuestas al componente rojo eran reforzadas en la primera mitad de la sesión y las respuestas al componente verde en la segunda mitad. A pesar de que el cambio en la distribución de reforzadores estaba señalado por 3 seg de tiempo

fuera, los animales no parecen haber incorporado esta señal externa como predictor del cambio. Sin embargo, utilizaron un estimado de la duración de la sesión experimental que les permitió adaptarse a los cambios frecuentes que experimentaron.

Una explicación sencilla para estos resultados es que en estas condiciones, el día y el tiempo de la sesión son predictores perfectos de los cambios, tal y como reportan Davis y Staddon (1990). Los animales parecen ser controlados por el día y la duración de la sesión, y en consecuencia, resulta fácil de asociar el cambio en el programa, con estas señales. Los pichones escogieron exactamente la alternativa rica cada día, de tal forma que la experiencia con refuerzo y no refuerzo que adquirieron durante el entrenamiento, sirvió como una señal en la memoria para la elección de la alternativa correcta al día siguiente.

La conclusión principal que podemos derivar de estos resultados es que los pichones pueden adaptarse, en la situación más sencilla de alternación diaria y alternación intrasesión, a cambios rápidos en los programas de refuerzo antes de llegar a equilibrio. Como los resultados obtenidos en experimentos previos con programas concurrentes (e.g., Davis & Staddon, 1990; Dreyfus, 1991; Mark & Gallistel, 1994), la adaptación fue más rápida para la condición de cambio intrasesión. Este ajuste rápido indica que la distribución de conducta cambia como una función de cambios locales en la tasa relativa de refuerzo. A partir de lo anterior, se puede concluir que los eventos que controlan la elección son agrupados en periodos más breves que los considerados anteriormente.

Experimento 2.

Cambio Aleatorio Diario y a la Segunda Mitad de la Sesión Experimental

El experimento 2 se llevó a cabo por los resultados obtenidos en el experimento 1. Una vez observado que los animales ajustan finamente su comportamiento a cambios diarios alternos y dentro de una misma sesión empleando programas múltiples, la siguiente pregunta que nos formulamos fue si ocurriría un ajuste y de qué tipo, ante cambios en las condiciones de reforzamiento que ocurrieran en forma aleatoria diaria y a la segunda mitad de la sesión.

Procedimiento

Para el experimento 2, los mismos sujetos y aparatos descritos en el experimento 1 fueron utilizados. Los grupos experimentales fueron constituidos por los sujetos que en el experimento 1 estuvieron en el grupo 1 y en el grupo 2, respectivamente.

En este experimento, el color asociado con reforzamiento fue asignado aleatoriamente tanto en los cambios entre sesión como en los cambios intrasesión. En esta última condición, en la primera mitad de la sesión, los reforzadores fueron programados para las respuestas al componente rojo. Al inicio de la transición hubo 3 segundos de tiempo fuera, que señalaban el cambio en la distribución de reforzadores, y los reforzadores fueron programados aleatoriamente a las respuestas a la tecla roja o verde, por el resto de la sesión. (Véase Apéndice A de la programación diaria en cada condición del experimento 2).

Una vez que la ejecución en ambas condiciones fue estable, se inició la condición de extinción (Múlt. EXT-EXT). Esta fase comenzó para la condición entre sesión, según la programación diaria previamente establecida. En la condición intrasesión, a la mitad de la sesión, siguió operando el periodo de tiempo fuera.

En la Tabla 2 se esquematiza la descripción de las condiciones experimentales.

TABLA 2. Descripción de las condiciones experimentales, en el experimento de cambio aleatorio diario y en la segunda mitad de la sesión experimental.

FASES EXPERIMENTALES		
Grupo 1	Entre sesiones Múltiple RV30-EXT Aleatorio	Extinción Múltiple EXT-EXT
Grupo 2	Intrasesiones Múltiple RV30-EXT Aleatorio	Extinción Múltiple EXT-EXT

Resultados

El objetivo principal en este experimento fue analizar a partir de los resultados del experimento 1, la posible adaptación a cambios aleatorios diarios y en la segunda mitad de la sesión experimental, en las distribuciones de reforzamiento en programas múltiples RV30-EXT.

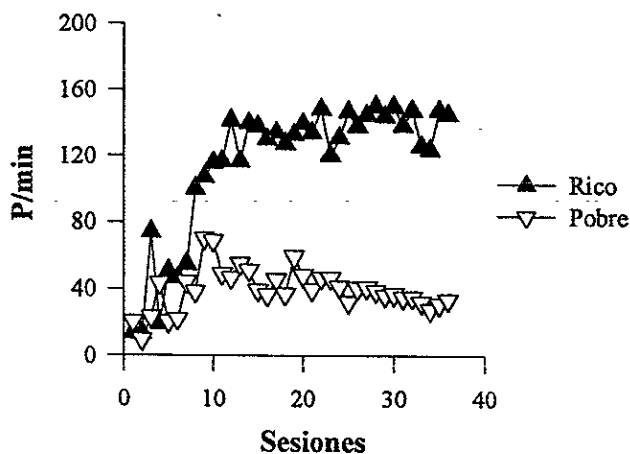


Fig. 5a. Ejecución asintótica promedio en cada componente del programa múltiple, durante la condición de cambio aleatorio entre sesión.

En la Figura 5a, están graficadas las tasas absolutas de respuesta en el componente rico y el componente pobre, promediado sobre los 6 sujetos para la condición de cambio entre sesión aleatorio.

La ejecución durante la adquisición empezó a un bajo nivel, y es a partir de la sesión 10, cuando se empieza a observar que las funciones se separan, mostrándose una mejoría conforme transcurren los días en la condición.

La Figura 5b, ilustra las tasas absolutas de respuesta para ambos componentes en la condición intrasesión. Se puede apreciar que antes de la transición, los animales responden gradualmente al componente rico y dejan de responder en el pobre, prácticamente desde el inicio de la condición. Sin embargo después de la transición, el ajuste ante los cambios fue considerablemente más lento, ya que los animales respondieron únicamente a un color y sólo después de 25 sesiones de entrenamiento, detectaron los cambios en la distribución de los reforzadores.

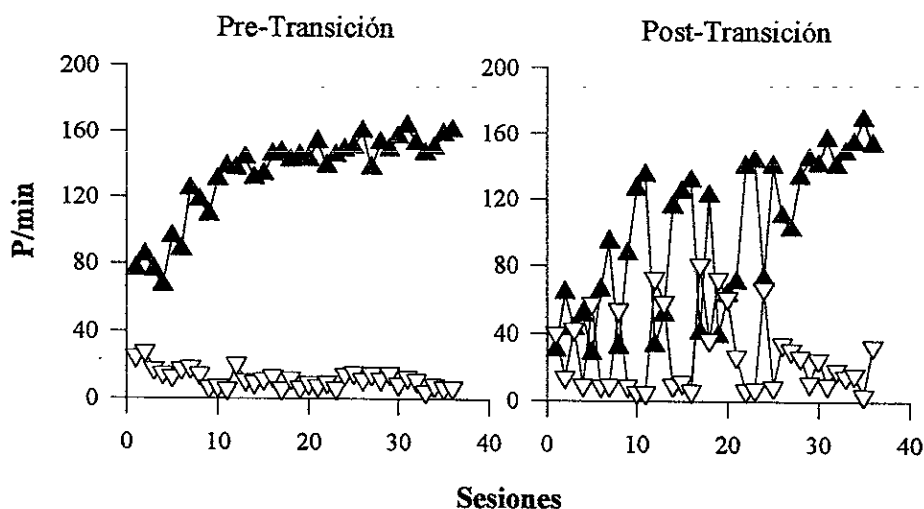


Fig. 5b. Ejecución asintótica promedio en cada componente del programa múltiple, durante la condición de cambio aleatorio en la segunda mitad de la sesión experimental. Los símbolos cerrados representan las respuestas al componente rico, y los símbolos abiertos las respuestas al componente pobre.

La Figura 6 presenta las tasas de respuesta durante los componentes del programa múltiple al final del entrenamiento (panel superior), y en los dos siguientes días de extinción (paneles inferiores) en la condición aleatoria cambio entre sesión (panel izquierdo), y cambio aleatorio en la segunda mitad de la sesión (panel derecho). Cada punto graficado representa el promedio de la tasa de respuesta al componente rojo (círculos) y al verde (cuadrados). En la mitad superior del panel, las respuestas al componente rico están representadas por símbolos cerrados, y las respuestas al componente pobre, por símbolos abiertos.

En el último día de la condición de cambio entre sesión, se observa un rápido ajuste desde el segundo componente. En la condición de cambio intrasesión, la obtención del reforzador por las respuestas en los primeros componentes antes de la transición, es la señal que los animales emplean para empezar a responder en el componente rico. Hacia el final de la primera mitad de la sesión no hubo algún patrón anticipatorio del cambio en la distribución de reforzadores, como el que se observó en el experimento 1. Después de la transición, los animales respondieron aproximadamente igual en ambos componentes, y fue la obtención del reforzador lo que hizo que respondieran al componente rico y decrementaran su tasa de respuesta en el componente pobre.

Es importante señalar que en la condición de extinción del cambio entre sesión, a partir del componente 10 se observó la inversión de la tasa de respuesta en el componente que al inicio de la sesión tenía la tasa de respuesta más alta, e incrementó la tasa de respuesta en el componente alterno. Esto indica que los pichones volvieron a muestrear en los componentes, cuando hubo transcurrido cierto periodo.

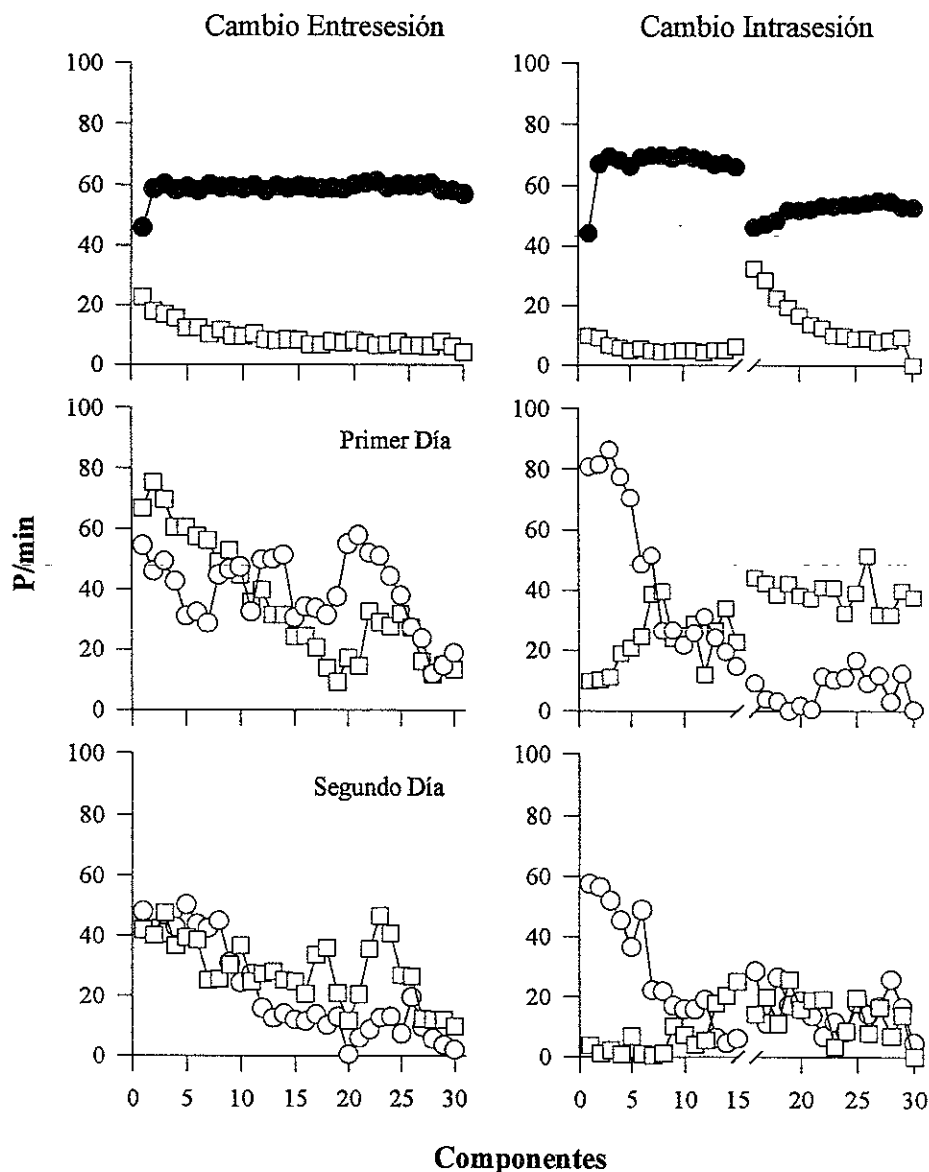


Fig. 6. Tasas de respuesta promedio al final del entrenamiento (panel superior), y en los dos primeros días de extinción (paneles inferiores), durante el cambio entre sesión (panel izquierdo) y cambio intrasesión (panel derecho) en la condición aleatoria diaria y en la segunda mitad de la sesión experimental. Ver detalles en el texto.

En la condición intrasesión, al final del primer tercio antes de la transición, decrementa la tasa de respuesta en el componente que había sido asociado con reforzamiento e incrementa la tasa en el componente alterno. Posteriormente, la tasa de respuesta en ambos componentes se mantiene aproximadamente igual antes del cambio. Después de la transición, se responde exclusivamente al componente verde a una tasa constante, con respuestas ocasionales al componente rojo. Esto parece indicar que los animales alternaron sus respuestas en los componentes, ya que en el último día de entrenamiento las respuestas que fueron reforzadas después de la transición fueron las respuestas al rojo, y en el primer día de extinción, la tasa de respuesta fue mayor para el componente verde. Esto sugiere un efecto entre sesiones, de lo que habían aprendido en las condiciones experimentales previas.

En el segundo día de extinción, se observa la inversión de tasas de respuesta tres componentes antes del cambio en la distribución de reforzadores; después de la transición, se aprecia una tasa de respuesta similar en ambos componentes, lo que hace suponer que los animales muestrean ante la incertidumbre de qué alternativa es la que está asociada con reforzamiento en este periodo.

Discusión

La conclusión principal del experimento 2 es que los animales no pueden seguir tan rápidamente, cambios aleatorios en programas de refuerzo múltiples antes de llegar a equilibrio. Hay un ajuste, pero ciertamente no tan evidente como en la situación de alternación.

Los datos observados en la condición intrasesión, hacen suponer que los animales muestrean ante la incertidumbre solo cuando el tiempo de la primera

mitad de la sesión ha transcurrido, ya que no es una estrategia adecuada muestrear continuamente si se tiene la experiencia de que el mismo color siempre ha sido reforzado en ese periodo. El hecho de que la tasa de respuesta sea similar en ambos componentes después de la transición y vaya decreciendo en el componente pobre, sugiere que dado que los cambios son impredecibles durante este periodo, el animal tiene que muestrear ininterrumpidamente hasta estimar las densidades de reforzamiento en ese día particular. En otras palabras, estos resultados parecen indicar que los animales utilizan el refuerzo como señal del cambio. Como en el experimento de Davis y Staddon (1990), los pichones requirieron de unos cuantos reforzadores para determinar la localización de la tecla rica cada día.

Los datos observados en extinción, en la condición de cambio entre sesión, indican que los animales invierten su tasa de respuesta en ambos componentes, hacia el componente 10 en ambos días. Una posible explicación de este patrón de alternación, es que los animales muestrean después de un cierto periodo en ambas alternativas, para obtener un estimado del reforzamiento en cada una de ellas. El hecho de que en ambos días hacia el componente 10 sea cuando invierten su tasa de respuesta, sugiere que los pichones parecen haber modificado la amplitud de su ventana de memoria para estimar el cambio, utilizando como periodo de muestreo el tiempo en las condiciones intrasesión a las que estuvieron expuestos previamente.

Estos resultados concuerdan con los resultados obtenidos en estudios de forrajeo óptimo (Shettleworth & Plowright, 1992; Lea & Dow, 1984), en los cuales los animales fueron capaces de cambiar la amplitud de su ventana de memoria para explorar los cambios recientes en el entorno y adaptarse exitosamente a éstos.

Experimento 3.

Transición Probabilística Intrasesión

El tercer experimento se condujo por los resultados obtenidos en los experimentos 1 y 2. A partir del hecho de que los animales pueden aprender la situación más sencilla de alternación diaria e intrasesión y a su vez, se adaptaron a cambios aleatorios entre sesiones y cambios señalados en la segunda mitad de la sesión en programas Múlt. RV30-EXT, evaluamos el otro extremo del continuo. En este experimento, la asignación del componente rico es completamente aleatoria y no existe una señal que prediga el cambio en la distribución de reforzadores.

Procedimiento

Para el experimento 3, los mismos sujetos (a excepción del sujeto O15 que murió antes de comenzar el experimento y fue reemplazado por el sujeto D5) y aparatos descritos en los experimentos 1 y 2 fueron utilizados.

A diferencia del experimento 2, en este experimento, los cambios aleatorios diarios ocurrieron desde el inicio de la sesión y a la mitad de la sesión se decidió aleatoriamente si cambiaba el color asociado con reforzamiento o si permanecía igual por el resto de la sesión (véase Apéndice A, Experimento 3); asimismo, los 3 seg de tiempo fuera que señalaban el cambio en la distribución de reforzadores a la mitad de la sesión, fueron eliminados.

Los sujetos se incluyeron en un grupo único, que fue expuesto a un programa Múlt. RV30-EXT aleatorio intrasesión.

Cuando las respuestas en esta condición alcanzaron niveles asintóticos, se inició la condición de extinción (Múlt. EXT-EXT). En esta condición, el tiempo total de la sesión fue de 90 minutos (en lugar de los 30 minutos empleados en los dos experimentos previos), ya que se pretendía analizar los posibles efectos tanto de regresión como del cambio entre sesiones que pudieran presentarse.

En la Tabla 3 se esquematiza la descripción de las condiciones experimentales.

TABLA 3. Descripción de las condiciones experimentales, en el experimento de transición probabilística intrasesión.

FASES EXPERIMENTALES		
Grupo único	Intrasesiones Múltiple RV30-EXT Aleatorio	Extinción Múltiple EXT-EXT

Resultados

El objetivo principal en este experimento fue analizar la adaptación a cambios intrasesión completamente aleatorios, en los parámetros de reforzamiento en programas múltiples RV30-EXT.

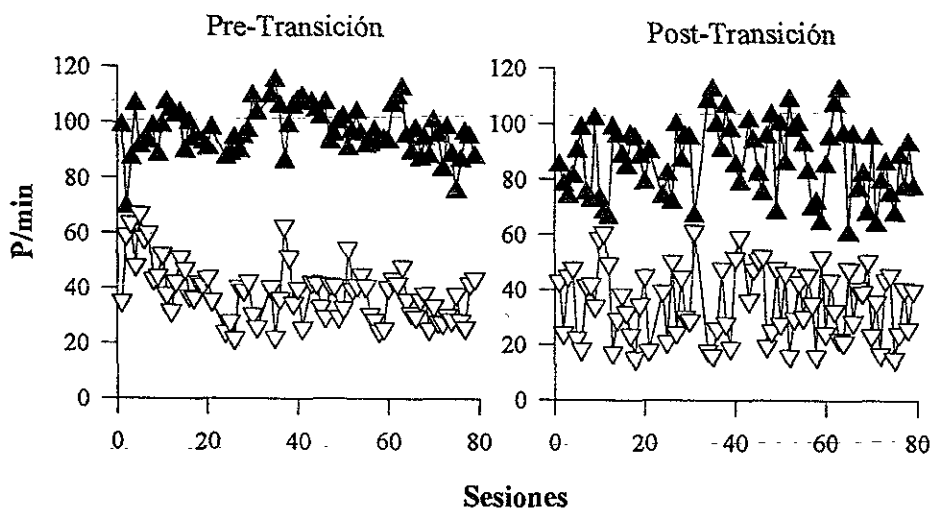


Fig.7. Tasas absolutas promedio para cada componente del programa múltiple, durante la condición transición probabilística intrasesión. Las respuestas al componente rico están representadas por los símbolos cerrados y las respuestas al componente pobre, por símbolos abiertos.

En la Figura 7, se presentan las tasas absolutas de respuesta en cada componente del programa múltiple, promediadas sobre los 12 sujetos en el cambio probabilístico intrasesión. El grupo de funciones promedio en la Figura 7 fue representativo de las funciones para los sujetos individuales.

Antes de la transición, las tasas de respuesta presentan fluctuaciones en ambos componentes, pero se observa un ligero ajuste, aunque no tan significativo como el observado en los experimentos anteriores. En contraste, las tasas de respuesta en ambos componentes presentan fluctuaciones más pronunciadas después de la transición.

La Figura 8 presenta las tasas de respuesta por componentes en los últimos tres días de la condición de transición probabilística intrasesión, cuando hubo cambios o no cambios en la distribución de reforzamiento.

En estas gráficas, se distingue que en los últimos días de la condición, únicamente en los componentes iniciales en la pre-transición, las tasas de respuesta son similares en ambos componentes del programa múltiple. Al segundo o tercer componente, los animales detectaron la tecla rica y siguieron respondiendo exclusivamente en ella hasta el final de la primera mitad de la sesión. Después de la transición, los animales muestrearon en ambos componentes hasta que la obtención de un reforzador señaló la tecla rica, y una vez sucedido esto, la tasa de respuesta incrementó en ese componente y decrementó en forma gradual en el componente alterno.

Cabe señalar que el penúltimo día de la condición, cuando las condiciones de reforzamiento permanecieron sin cambio, después de la transición, se observó un ligero incremento en los primeros componentes del programa pobre, pero la tasa de respuesta en el componente rico permaneció casi igual y se mantuvo constante con la obtención de los reforzadores.

De la Figura 8 se desprende el hecho de que los animales en la condición totalmente aleatoria, se van ajustando componente a componente durante la sesión completa.

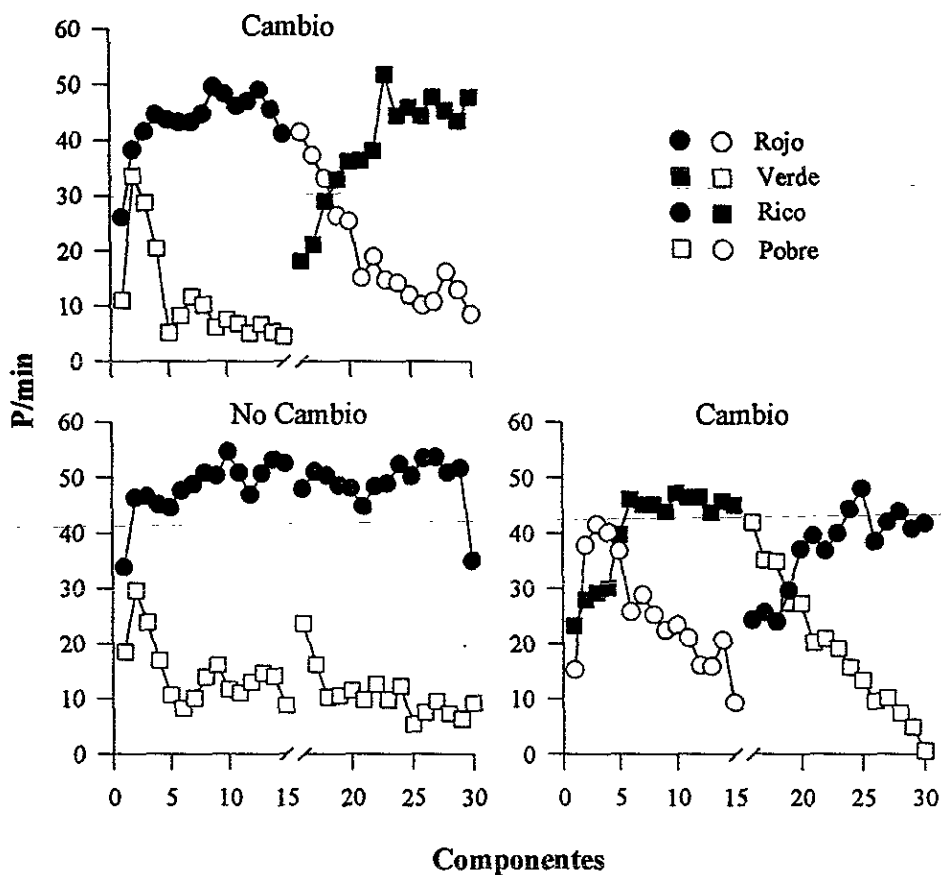


Fig. 8. Tasas de respuesta promedio al final del entrenamiento durante los tres últimos días de la condición transición probabilística intrasesión, cuando hubo cambios o no cambios en la distribución de reforzamiento de la tecla rica.

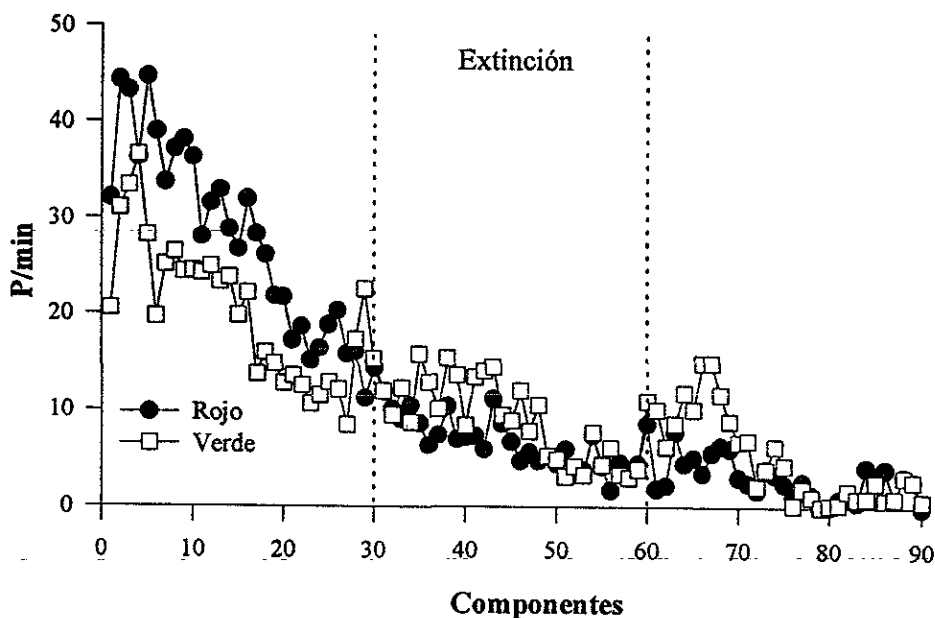


Fig. 9. Tasas absolutas de respuesta promedio para los componentes rojo y verde, durante el primer día de extinción en la condición transición probabilística intrasesión.

La Figura 9 ilustra la ejecución promedio de todos los sujetos componente por componente, en el primer día de extinción. La ejecución de los sujetos muestra una tasa de respuesta mayor al componente rojo durante los primeros 30 componentes. Una vez transcurrido este periodo, los animales invirtieron su distribución de respuestas, y comenzaron a responder a una tasa mayor en componente verde, continuando así hasta el componente 70.

Una vez más, es evidente el efecto del cambio entre sesión, en el cual los animales invierten sus preferencias aproximadamente cada 30 minutos durante toda la sesión, como si estuvieran estimando el transcurso de un día de experimentación o la duración de una sesión experimental.

Discusión

Los resultados del experimento 3, sugieren que la adaptación es mucho más lenta cuando los cambios en la distribución de reforzadores que experimentan los pichones, son completamente impredecibles. No obstante, se observa un ajuste componente a componente.

Al inicio del día, el animal no puede predecir cuál es el color rico y a la mitad de la sesión no puede predecir si va a continuar siéndolo. La única clave que aun tiene es una señal de que va a la mitad del día y lo que aprendió en condiciones anteriores. Lo que se observa en esta situación es el ajuste a sólo los cambios en el programa, ya que no es posible anticiparlos. Desde luego que con la experiencia que tiene el organismo, queda una clave fundamental: En el momento en que recibe reforzamiento, continúa respondiendo en esa opción; a la mitad de la sesión muestra, y en la alternativa que recibe reforzamiento, permanece respondiendo. Este patrón de comportamiento solo puede ser alcanzado, si el organismo es capaz de detectar cambios en el sitio de alimentación tan pronto como éstos ocurren. Si los cambios de bueno a malo son impredecibles y no están asociados a cualquier tipo de señal como tiempo, la única forma que tiene el organismo para explorar los cambios es muestrear ambas alternativas periódicamente, que fue precisamente lo que los animales hicieron bajo estas condiciones. Este resultado apoya los resultados obtenidos en estudios de forrajeo en ambientes variables (e.g., Dow & Lea, 1987; Kacelnik & Todd, 1992; Shettleworth, 1988; Shettleworth & Plowright, 1989, 1992), en donde el patrón de muestreo fué decisivo para la adaptación a los cambios que los animales experimentaron.

DISCUSIÓN GENERAL

El problema que enfrentan los organismos en su entorno es que de entre varias posibles respuestas, deben seleccionar aquella que es responsable de producir el reforzador (es lo mismo, pero más general, seleccionar el lugar donde se encuentra). Pero no solo ello, una vez seleccionada la respuesta, deben distinguir aquellos periodos en los cuales es exitosa, de aquellos en los que no lo es. Sabemos que en equilibrio, los animales pueden hacer esto y que la distribución de respuestas obedece a reglas muy sencillas como igualación. Para estudiar cómo los organismos se ajustan a estos cambios, se requiere de experimentos en periodos de transición antes de llegar a equilibrio. Existe una buena cantidad de datos para el tipo de adaptación simultánea, pero no es el caso para la adaptación sucesiva.

En el presente trabajo, se investigó la adaptación a cambios en las distribuciones de refuerzo con el mismo procedimiento empleado en elección simultánea, el bandido de dos brazos, para la situación de elección sucesiva. Nuestro principal resultado fue que los pichones se adaptan a cambios rápidos en la distribución de reforzadores, cuando ésta ocurre en forma alterna diaria y dentro de una misma sesión (Experimento 1, Fig. 1 y 2). Estos resultados coinciden con los resultados reportados por Davis y Staddon (1990), Dreyfus (1991), y Mark y Gallistel (1994), ya que demuestran que los organismos son sensibles a cambios rápidos en su entorno, y capaces de modificar su ventana de memoria, para ajustarse a cambios locales en la densidad de reforzamiento. Asimismo, los resultados revelan dos tipos de procesos. Un proceso de

memoria de largo plazo, en el cual se almacena la información de la tecla asociada con reforzamiento el día anterior, y un proceso de memoria de corto plazo, que le permite al animal ajustar su conducta a las condiciones de cada sesión diaria.

En el experimento 1, los animales parecen ser controlados por el día y la duración de la sesión, y en consecuencia, resulta fácil de asociar el cambio en el programa, con estas señales. Es importante mencionar, que la adaptación fue más rápida para la condición de cambio intrasesión que para el cambio entre sesión, resultados consistentes con experimentos de discriminación sucesiva con procedimientos de ensayos discretos que demuestran que cuando los sujetos son expuestos a cambios más frecuentes en la densidad de reforzamiento, la adaptación es más rápida (Papini, 1997); y con estudios de aprendizaje de cambios, cuyos resultados demuestran cambios de preferencias muy rápidos ante cambios frecuentes en la distribución de reforzamiento (e.g., cuando las contingencias cambiaron diariamente vs. cada dos o cada cuatro días, Davis & Staddon, 1990). Esta diferencia en la adquisición indica que el arreglo pudo haber retardado la adquisición en el cambio entre sesión debido a la apreciación del horizonte temporal que tuvieron los sujetos, ya que es probable que los animales varíen su comportamiento de acuerdo al horizonte temporal que estén esperando.

Por otra parte, el ajuste rápido observado indica que la distribución de conducta cambia como una función de cambios locales en la tasa relativa de refuerzo. El hecho de que se haya encontrado un ajuste en una escala temporal local, fortalece el argumento de que los eventos que controlan la elección son agrupados en periodos más breves que los considerados previamente por las teorías molares, y que la variable que controla el proceso de elección

concurrente (e.g., Dreyfus, 1991; Mark & Gallistel, 1994), pudiera ser la misma para situaciones de elección sucesiva, sugiriendo que el organismo aprende un contexto global en el que compara cada refuerzo recibido.

Cuando los cambios son asignados aleatoriamente, el ajuste no es tan rápido (Experimento 3, Fig. 7), aunque se observa un ajuste componente a componente (Experimento 3, Fig. 8). Estos resultados parecen indicar que los animales emplean como señal del cambio, el reforzamiento que reciben en determinada alternativa. La estrategia importante ante estas situaciones es muestrear continuamente para evaluar la densidad de reforzamiento en cada opción.

Asimismo, los datos obtenidos en el presente trabajo indican que ocurre una mejoría en la ejecución en este tipo de programas múltiples. En el experimento 1 hay una clara mejoría con subsecuentes cambios diarios. En el experimento 2, donde siempre se inicia con la misma condición como rica y aleatoriamente se cambia o no, también hay mejoría. En el tercer experimento (todo aleatorio al inicio y a la mitad de la sesión), hay una mejoría a través de cambios sucesivos pero no tan significativa como en los experimentos anteriores. El hecho de que se haya obtenido este resultado puede ser explicado de dos formas: La primera es que los sujetos pudieron haber empleado el resultado de la primer respuesta cada día como una señal de la identidad del estímulo reforzado en ese día. La segunda es que los sujetos computaron la proporción de respuestas correctas en este día, dada la proporción de respuestas ayer y las condiciones de reforzamiento hoy.

En particular, los experimentos 2 y 3 son una buena prueba del modelo de Davis et al. (1993), donde se observa que lo que controla a los sujetos son

los rendimientos por respuesta (i.e., cuántos reforzadores se reciben por número de respuesta), en otras palabras, la probabilidad de refuerzo.

Los efectos de regresión observados, sugieren que los animales integraron el tiempo en el cual los reforzadores estaban programados para las teclas roja y verde. Los animales necesitan un mecanismo por medio del cual integren eventos discretos en tasas de reforzamiento, para ello necesitan tener una memoria del componente rojo (i.e., de qué fue lo último que sucedió en rojo), y una memoria del componente verde (Capaldi, 1994). En los experimentos 1 y 2, los sujetos respondieron de acuerdo a este estimado, sin considerar como predictor del cambio los 3 seg de tiempo fuera. Esto fortalece la suposición de que los animales integran la información del pasado reciente con información de un pasado más distante, para modificar su comportamiento ante cambios en su entorno. Para ello, emplean diferentes tipos de reglas, tales como cambiar la amplitud de la ventana de memoria en forma óptima (según la experiencia individual del organismo) que permita la exploración de cambios frecuentes; y asignar un peso temporal para los estimados de la densidad de refuerzo global así como para los estimados de densidad de refuerzo local en las sesiones experimentales.

Igualmente, retener una preferencia aprendida para un sitio de alimentación podría parecer tener una función adaptativa en la naturaleza: En la ausencia de cualquier otra información, el mejor predictor de dónde se encuentra hoy la comida puede ser donde se encontraba ayer. En consecuencia, responder a las alternativas o sitios de alimentación que fueron últimamente los mejores, puede ser importante. Así, cuando la información es reciente, encontramos que los últimos resultados predicen el comportamiento con una completa precisión, como ha sido descrito en los experimentos reportados en la

literatura (e.g., Cuthill et al., 1990; Mark & Gallistel, 1994; Shettleworth & Plowright, 1992).

En consecuencia, la conducta en circunstancias variables es posiblemente más decisiva para la supervivencia de un animal y por ello, más probable de que permita revelar el mecanismo que influye en distribución del comportamiento (e.g., véase Bailey & Mazur, 1990; Dow & Lea, 1987; Kacelnik et al., 1987). Estos datos particulares, así como los datos reportados en estudios de forrajeo óptimo y programas operantes convencionales, revelan que no hay un intervalo específico en el cual los eventos en diversas situaciones sean integrados (i.e., una ventana de memoria de amplitud fija). Sin embargo, la necesidad de que la conducta permanezca sensible a los cambios medio ambientales (Killeen, 1981; Lea & Dow, 1984) sugiere que ciertamente debe haber límites superiores e inferiores en la duración de este intervalo. Dentro de estos límites, el intervalo depende de factores tales como la tasa total de reforzamiento (Lea & Dow, 1984).

Finalmente, los modelos descriptivos que se han discutido representan solo una posible aproximación al problema de cómo los animales responden a variaciones en la tasa de refuerzo o la probabilidad de refuerzo. La mayoría de las versiones derivadas del modelo de operador lineal, es incapaz de explicar ajustes a cambios muy simples en el ambiente. Una aproximación alternativa sería modelar el comportamiento del animal como un problema de adquisición de la información, preguntándonos cómo los animales deben de muestrear su medio ambiente en forma eficiente para adquirir información acerca de las densidades de reforzamiento actuales. Este tipo de modelamiento podría proporcionar interesantes implicaciones acerca de la capacidades cognitivas de los animales. Asimismo, los resultados encontrados, ofrecen elementos para

fundamentar algunos procedimientos experimentales dirigidos a entender cómo los animales integran la información de su medio ambiente y como esto se transforma en comportamiento.

Recientemente, Bouzas y col. han empezado a estudiar en el contexto de la elección sucesiva, el intervalo de integración de los reforzadores, el papel del valor proporcional y absoluto de las fuentes de reforzamiento, y las características de su ajuste dinámico. El siguiente paso es establecer la valoración de los modelos de ajuste dinámico a la situación de elección sucesiva y generar un modelo capaz de integrar los mecanismos que determinan qué cambios en los componentes en la estructura fina del comportamiento, producen cambios observados en las variables más moleculares.

REFERENCIAS

- Bailey, J. T., & Mazur, J. E. (1990). Choice behavior in transition: Development of preference for the higher probability of reinforcement. Journal of the Experimental Analysis of Behavior, 53, 409-422.
- Bateson, M., & Kacelnik, A. (1995). Preferences for fixed and variable food resources: Variability in amount and delay. Journal of the Experimental Analysis of Behavior, 63, 313-329.
- Beale, I. L. (1970). The effects of amount of training per reversal on successive reversals of a color discrimination. Journal of the Experimental Analysis of Behavior, 14, 345-352.
- Belke, T. W. (1992). Stimulus preference and the transitivity of preference. Animal Learning & Behavior, 20, 401-406.
- Brunner, D., Kacelnik, A., & Gibbon, J. (1996). Memory for inter-reinforcement interval variability and patch departure decisions in the starling, *Sturnus vulgaris*. Animal Behavior, 51, 1025-1045.
- Bush, R. R., & Mosteller, F. (1951). A mathematical model for simple learning. Psychological Review, 58, 313-323.
- Bush, R. R., & Mosteller, F. (1955). Stochastic models for learning. New York: Wiley.
- Capaldi, E. J. (1994). The sequential view: From rapidly fading stimulus traces to the organization of memory and the abstract concept of number. Psychonomic Bulletin & Review, 1, 156-181.
- Catania, A. C. (1961). Behavioral contrast in a multiple and concurrent schedule of reinforcement. Journal of the Experimental Analysis of Behavior, 4, 335-342.

- Catania, A. C. (1963). Concurrent performances: Reinforcement interaction and response independence. Journal of the Experimental Analysis of Behavior, 6, 253-263.
- Catania, A. C. (1969). Concurrent Performances: Inhibition of one response by reinforcement of another. Journal of the Experimental Analysis of Behavior, 12, 731-744.
- Charman, L., & Davison, M. (1982). On the effects of component durations and component reinforcement rates in multiple schedules. Journal of the Experimental Analysis of Behavior, 37, 417-439.
- Colwill, R. M., & Delamater, B. A. (1995). An associative analysis of instrumental biconditional discrimination learning. Animal Learning & Behavior, 23, 218-233.
- Couvillon, P. A., Brandon, S. E., Woodard, W. T., & Bitterman, M. E. (1980). Performance of pigeons in patterned sequences of rewarded and nonrewarded trials. Journal of Experimental Psychology: Animal Behavior Processes, 6, 137-154.
- Couvillon, P. A., & Bitterman, M. E. (1985). Analysis of choice in honeybees. Animal Learning & Behavior, 13, 246-252.
- Couvillon, P. A., & Bitterman, M. E. (1986). Performance of honeybees in reversal and ambiguous-cue problems: Test of a choice model. Animal Learning & Behavior, 14, 225-231.
- Cuthill, I. C., Kacelnik, A., Krebs, J. R., Haccou, P., & Iwasa, Y. (1990). Starlings exploiting patches: The effect of recent experience on foraging decisions. Animal Behaviour, 10, 625-640.
- Daly, H. B., & Daly, J. T. (1982). A mathematical model of reward and aversive nonreward: Its application in over 30 appetitive learning situations. Journal of Experimental Psychology: General, 111, 441-480.
- Davis, D. G. S., & Staddon, J. E. R. (1990). Memory for reward in probabilistic choice: Markovian and non-Markovian properties. Behaviour, 114, 37-64.

- Davis, D. G. S., Staddon, J. E. R., Machado, A., & Palmer, R. G. (1993). The process of recurrent choice. Psychological Review, 100, 320-341.
- Davison, M. C., & Hunter, I. W. (1979). Concurrent schedules: Undermatching and control by previous experimental conditions. Journal of the Experimental Analysis of Behavior, 32, 233-244.
- Davison, M., & McCarthy, D. (1988). The matching law: A research review. Hillsdale, NJ: Erlbaum.
- Devenport, L., Hill, T., Wilson, M., & Ogden, E. (1997). Tracking and averaging in variable environments: A transition rule. Journal of Experimental Psychology: Animal Behavior Processes, 23, 450-460.
- de Villiers, P. A. (1977). Choice in concurrent schedules and a quantitative formulation of the law of effect. En W. K. Honig & J. E. R. Staddon (Eds.), Handbook of operant behavior. New York: Prentice-Hall, pp. 233-287.
- Dow, S. M., & Lea, S. E. G. (1987). Foraging in a changing environment: Simulations in the operant laboratory. En M. L. Commons, A. Kacelnik, & S. J. Shettleworth (Eds.), Quantitative Analyses of Behavior: Vol. 6. Foraging (pp. 90-113). Hillsdale, NJ: Erlbaum.
- Dragoi, V. (1997). A dynamic theory of acquisition and extinction in operant learning. Neural Networks, 9, 1-29.
- Dragoi, V., & Staddon, J. E. R. (En prensa). The dynamics of operant conditioning. Psychological Review.
- 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.
- Estes, W. K. (1950). Toward a statistical theory of learning. Psychological Review, 57, 94-104.
- Gallistel, C. R. (1990). The organization of learning. Cambridge, MA: Bradford Press.

- Gallistel, C. R. (1994). Foraging for brain stimulation: Toward a neurobiology of computation. Cognition, 50, 151-170.
- Gibbon, J., Church, R. M., Fairhurst, S., & 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, 208-215.
- Herrnstein, R. J. (1961). Relative and absolute strength of responses as a function of frequency of reinforcement. Journal of the Experimental Analysis of Behavior, 4, 267-272.
- Herrnstein, R. J. (1970). On the law of effect. Journal of the Experimental Analysis of Behavior, 13, 243-266.
- Herrnstein, R. J. (1982). Melioration as behavioral dynamism. En M. L. Commons, R. J. Herrnstein, & H. Rachlin (Eds.), Quantitative Analyses of Behavior: Vol. 2. Matching and maximizing accounts (pp. 433-458). Cambridge, MA: Ballinger.
- Herrnstein, R. J., & Brady, J. (1958). Interaction among components of a multiple schedule. Journal of the Experimental Analysis of Behavior, 1, 293-300.
- Herrnstein, R. J., & Loveland, D. H. (1975). Maximizing and matching on concurrent ratio schedules. Journal of the Experimental Analysis of Behavior, 24, 107-116.
- Herrnstein, R. J., & Vaughan, W., Jr. (1980). Melioration and behavioral allocation. En J. E. R. Staddon (Ed.), Limits to action: The allocation of individual behavior (pp. 143-176). San Diego CA: Academic Press.
- Higa, J. J. (1996a). Dynamics of time discrimination: II The effects of multiple impulses. Journal of the Experimental Analysis of Behavior, 66, 117-134.
- Higa, J. J. (1996b). Rapid timing of a single transition in interfood interval duration by rats. Animal Learning & Behavior, 25, 177-184.

- Higa, J. J., Wynne, C. D. L., & Staddon, J. E. R. (1991). Dynamics of time discrimination. Journal of Experimental Psychology: Animal Behavior Processes, 17, 281-291.
- Horner, J. M., Staddon, J. E. R., & Lozano, K. (1997). Integration of reinforcements effects over time. Animal Learning & Behavior, 25, 84-98.
- Hunter, I., & Davison, M. C. (1985). Determination of a behavioral transfer function: White-noise analysis of session-to-session response-ratio dynamics on concurrent VI VI schedules. Journal of the Experimental Analysis of Behavior, 43, 43-59.
- Kacelnik, A., Krebs, J. R., & Ens, B. (1987). Foraging in a changing environment: An experiment with Starlings (*Sturnus vulgaris*). En M. L. Commons, A. Kacelnik, & S. J. Shettleworth (Eds.), Quantitative Analyses of Behavior: Vol. 6 Foraging (pp. 63-87). Hillsdale, NJ: Erlbaum.
- Kacelnik, A., & Todd, I. A. (1992). Psychological mechanisms and the marginal value theorem: Effect of variability in travel time on patch exploitation. Animal Behaviour, 43, 313-322.
- Killeen, P. R. (1968). On the measurement of reinforcement frequency in the study of preference. Journal of the Experimental Analysis of Behavior, 11, 263-269.
- Killeen, P. R. (1981). Averaging theory. En C. M. Bradshaw, E. Szabadi, & C. F. Lowe (Eds.), Recent developments in the quantification of steady-state operant behavior (pp. 21-34). Amsterdam: Elsevier/ North Holland Biomedical Press.
- Killeen, P. R. (1982). Incentive theory: II. Models for choice. Journal of the Experimental Analysis of Behavior, 38, 217-232.
- Killeen, P. R. (1994). Mathematical principles of reinforcement. Behavioral and Brain Sciences, 17, 105-172.

- Killeen, P. R. (En prensa). The first principle of reinforcement. En J. E. R. Staddon & C. D. L. Wynne (Eds.), Models of action. Hillsdale, NJ: Erlbaum.
- Lander, D. G., & Irwin, R. J. (1968). Multiple schedules: Effects of the distribution of reinforcements between components on the distribution of responses between components. Journal of the Experimental Analysis of Behavior, 11, 517-524.
- Lave, C., & March, J. G. (1975). An introduction to models in the social sciences. New York: Harper & Row.
- Lea, S. E. G., & Dow, S. M. (1984). The integration of reinforcements over time. Annals of the New York Academy of Sciences, 423, 269-277.
- Mackintosh, N. J. (1974). The psychology of animal learning. NY: Academic Press.
- March, J. G. (1996). Learning to be risk averse. Psychological Review, 103, 309-319.
- Mark, T. A., & Gallistel, C. R. (1994). Kinetics of matching. Journal of Experimental Psychology: Animal Behavior Processes, 20, 79-95.
- 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. (1995). Development of preference and spontaneous recovery in choice behavior with concurrent variable-interval schedules. Animal Learning & Behavior, 23, 93-103.
- Mazur, J. E. (1996). Past experience, recency and spontaneous recovery in choice behavior. Animal Learning & Behavior, 24, 1-10.
- Mazur, J. E., & Hastie, R. (1978). Learning as accumulation: A reexamination of the learning curve. Psychological Bulletin, 85, 1256-1274.

- Mazur, J. E., & Ratti, T. A. (1991). Choice behavior in transition: Development of preference in a free-operant procedure. Animal Learning & Behavior, 19, 241-248.
- McLean, A. P., & White, K. G. (1981). Undermatching and contrast within components of multiple schedules. Journal of the Experimental Analysis of Behavior, 35, 283-291.
- McSweeney, F. K., Farmer, V. A., Dougan, J. D., & Whipple, J. E. (1986). The generalized matching law as a description of multiple-schedule responding. Journal of the Experimental Analysis of Behavior, 45, 83-101.
- Menlove, R. L. (1975). Local patterns of responding maintained by concurrent and multiple schedules. Journal of the Experimental Analysis of Behavior, 23, 309-337.
- Myerson, J., & Hale, S. (1988). Choice in transition: A comparison of melioration and the kinetic model. Journal of the Experimental Analysis of Behavior, 49, 291-302.
- Myerson, J., & Miezin, F. M. (1980). The kinetics of choice: An operant system analysis. Psychological Review, 87, 160-174.
- Nevin, J. A. (1974). Response strength in multiple schedules. Journal of the Experimental Analysis of Behavior, 21, 389-408.
- Nevin, J. A., & Shettleworth, S. J. (1966). An analysis of contrast effects in multiple schedules. Journal of the Experimental Analysis of Behavior, 9, 305-315.
- Palya, W. L. (1992). Dynamics in the fine structure of schedule-controlled behavior. Journal of the Experimental Analysis of Behavior, 57, 267-287.
- Papini, M. R. (1997). Role of reinforcement in spaced-trial operant learning in pigeons (*Columbia livia*). Journal of Comparative Psychology, 111, 275-285.

- Pliskoff, S. S., Shull, R. L., & Gollub, L. R. (1968). The relation between response rates and reinforcer rates in a multiple schedule. Journal of the Experimental Analysis of Behavior, 11, 271-284.
- Rachlin, H. (1973). Contrast and matching. Psychological Review, 80, 217-234.
- Rachlin, H., & Burkhard, B. (1978). The temporal triangle: Response substitution in instrumental conditioning. Psychological Review, 85, 22-47.
- Rachlin, H., Battalio, R., Kagel, J., & Green, L. (1981). Maximization theory in behavioral psychology. The Behavioral and Brain Sciences, 4, 371-417.
- Real P. G., & Dreyfus, L. R. (1985). Levels of aggregation: Relative time allocation in concurrent-schedule performance. Journal of the Experimental Analysis of Behavior, 43, 97-113.
- Reynolds, G. S. (1961a). An analysis of interactions in a multiple schedule. Journal of the Experimental Analysis of Behavior, 4, 107-117.
- Reynolds, G. S. (1961b). Behavioral contrast. Journal of the Experimental Analysis of Behavior, 4, 57-71.
- Reynolds, G. S. (1961c). Relativity of response rate and reinforcement frequency in a multiple schedule. Journal of the Experimental Analysis of Behavior, 4, 179-184.
- Reynolds, G. S. (1963a). On some determinants of choice in pigeons. Journal of the Experimental Analysis of Behavior, 6, 53-59.
- Reynolds, G. S. (1963b). Some limitations on behavioral contrast and induction during successive discrimination. Journal of the Experimental Analysis of Behavior, 6, 131-139.
- Robbins, S. J. (1990). Mechanisms underlying spontaneous recovery in autoshaping. Journal of Experimental Psychology: Animal Behavior Processes, 16, 235-249.

- Rosenthal, R., & Rosnow, R. L. (1991). Essentials of behavioral research: Methods and data analysis. McGraw-Hill.
- Schmajuk, N. A., Lam, Y. W., & Gray, J. A. (1996). Latent inhibition: A neural network approach. Journal of Experimental Psychology: Animal Behavior Processes, 22, 321-349.
- Shettleworth, S. J. (1988). Foraging as operant behavior and operant behavior as foraging: What we have learned? The Psychology of Learning and Motivation, 22, 1-49.
- Shettleworth, S. J., Krebs, J. R., Stephens, D. W., & Gibbon, J. (1988). Tracking a fluctuating environment: A study of sampling. Animal Behaviour, 36, 87-105.
- Shettleworth, S. J., & Plowright, C. M. S. (1989). Time horizons of pigeons on a two-armed bandit. Animal Behaviour, 37, 610-623.
- Shettleworth, S. J., & Plowright, C. M. S. (1992). How pigeons estimate rates of prey encounter. Journal of Experimental Psychology: Animal Behavior Processes, 18, 219-235.
- Shimp, C. P. (1966). Probabilistically reinforced choice behavior in pigeons. Journal of the Experimental Analysis of Behavior, 9, 443-455.
- Staddon, J. E. R. (1977). On Herrnstein's equation and related forms. Journal of the Experimental Analysis of Behavior, 28, 163-170.
- Staddon, J. E. R. (1982). Behavioral competition, contrast and matching. En M. L. Commons, R. J. Herrnstein, & H. Rachlin (Eds.), Quantitative Analyses of Behavior: Vol. 2. Matching and maximizing accounts (pp. 243-261) Cambridge, MA: Ballinger.
- 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. (1992). Rationality, melioration and law-of-effect models for choice. Psychological Science, 3, 136-141.

- Staddon, J. E. R. (1993). On rate-sensitive habituation. Adaptive Behavior, 1, 421-436.
- Staddon, J. E. R., & Frank, J. (1974). Mechanisms of discrimination reversals. Animal Behaviour, 22, 802-828.
- Staddon, J. E. R., & 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.
- Staddon, J. E. R., Davis, D. G. S., Machado, A., & Palmer, R. G. (1994). Cumulative effects model: A response to Williams (1994). Psychological Review, 101, 708-710.
- Sutherland, N. S., & Mackintosh, N. J. (1971). Mechanisms of animal discrimination learning. London: Academic Press.
- Theios, J. (1965). The mathematical structure of reversal learning in a shock-escape T-maze: Overtraining and successive reversals. Journal of Mathematical Psychology, 2, 26-52.
- Vaughan W., Jr. (1981). Melioration, matching and maximization. Journal of the Experimental Analysis of Behavior, 36, 141-149.
- Wilton, R. N., & Gay, R. A. (1969). Behavioral contrast in one component of a multiple schedule as a function of the reinforcement conditions operating in the following component. Journal of the Experimental Analysis of Behavior, 12, 239-246.
- Williams, B. A. (1976a). Short-term retention of response outcome as a determinant of serial reversal learning. Learning & Motivation, 7, 418-430.
- Williams, B. A. (1976b). Behavioral contrast as a function of the temporal location of reinforcement. Journal of the Experimental Analysis of Behavior, 26, 57-64.

- Williams, B. A. (1976c). Elicited responding to signals for reinforcement: The effects of overall versus local changes in reinforcement probability. Journal of the Experimental Analysis of Behavior, 26, 213-220.
- Williams, B. A. (1979). Contrast, component duration and the following schedule of reinforcement. Journal of Experimental Psychology: Animal Behavior Processes, 5, 379-396.
- Williams, B. A. (1981). The following schedule of reinforcement as a fundamental determinant of steady state contrast in multiple schedules. Journal of the Experimental Analysis of Behavior, 35, 293-310.
- Williams, B. A. (1982). Do interactions in multiple and concurrent schedules have a common basis? En M. L. Commons, R. J. Herrnstein, & H. Rachlin (Eds.), Quantitative Analyses of Behavior: Vol. 2. Matching and maximizing accounts (pp. 281-302). Cambridge, MA: Ballinger.
- Williams, B. A. (1983). Another look at contrast in multiple schedules. Journal of the Experimental Analysis of Behavior, 39, 345-384.
- Williams, B. A. (1988). Reinforcement, choice and response strength. En R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), Steven's handbook of experimental psychology (2^a ed., Vol. 2. pp. 167-244). New York: Wiley.
- Williams, B. A. (1991). Behavioral contrast and reinforcement value. Animal Learning & Behavior, 19, 337-344.
- Williams, B. A. (1992). Inverse relations between preference and contrast. Journal of the Experimental Analysis of Behavior, 58, 303-312.
- Williams, B. A. (1993). Molar vs. local reinforcement probability as determinants of stimulus value. Journal of the Experimental Analysis of Behavior, 59, 163-172.
- Williams, B. A. (1994). The role of probability of reinforcement in models of choice. Psychological Review, 101, 704-707.

- Williams, B. A. (1997). What is learned? Revisiting an old issue. Journal of the Experimental Analysis of Behavior, 67, 255-258.
- Williams, B. A., & Wixted, J. T. (1986). An equation for behavioral contrast. Journal of the Experimental Analysis of Behavior, 45, 47-62.
- Williams, B. A., & Royalty, P. (1989). A test of the melioration theory of matching. Journal of Experimental Psychology: Animal Behavior Processes, 15, 99-113.
- Williams, B. A., & Bell, M. C. (1996). Changeover behavior and preference in concurrent schedules. Journal of the Experimental Analysis of Behavior, 65, 513-526.
- Woodard, W. T., & Bitterman, M. E. (1976). Asymptotic reversal learning in pigeons: Mechanisms for reducing inhibition. Journal of Experimental Psychology: Animal Behavior Processes, 2, 57-66.
- Wynne, C. D. L., Staddon, J. E. R., & Delius, J. D. (1996). Dynamics of waiting in pigeons. Journal of the Experimental Analysis of Behavior, 65, 603-618.

APÉNDICE A

Experimento 1.**Múltiple RV30-EXT. Cambio Alterno Diario y Cambio Intrasesión****MÚLTIPLE RV30-EXT ALTERNO**

DÍA	CAMBIO ENTRE SESIONES Componente Rico	CAMBIO INTRASESIONES Componente Rico
1	Rojo	Rojo → Verde
2	Verde	Rojo → Verde
3	Rojo	Rojo → Verde
4	Verde	Rojo → Verde
5	Rojo	Rojo → Verde
6	Verde	Rojo → Verde
7	Rojo	Rojo → Verde
8	Verde	Rojo → Verde
9	Rojo	Rojo → Verde
10	Verde	Rojo → Verde
11	Rojo	Rojo → Verde
12	Verde	Rojo → Verde
13	Rojo	Rojo → Verde
14	Verde	Rojo → Verde
15	Rojo	Rojo → Verde
16	Verde	Rojo → Verde
17	Rojo	Rojo → Verde
18	Verde	Rojo → Verde
19	Rojo	Rojo → Verde
20	Verde	Rojo → Verde
21	Rojo	Rojo → Verde
22	Verde	Rojo → Verde
23	Rojo	Rojo → Verde
24	Verde	Rojo → Verde
25	Rojo * fin ciclo 2	Rojo → Verde
26	Verde	Rojo → Verde
27	Rojo	Rojo → Verde
28	Verde	Rojo → Verde
29	Rojo	Rojo → Verde
30	Verde	Rojo → Verde * fin ciclo 2

MÚLTIPLE RV30-EXT ALTERNO (Cont.)

DÍA	CAMBIO ENTRE SESIONES Componente Rico	CAMBIO INTRASESIONES Componente Rico
31	Rojo	Rojo → Verde
32	Verde	Rojo → Verde
33	Rojo	Rojo → Verde
34	Verde	Rojo → Verde
35	Rojo	Rojo → Verde
36	Verde	Rojo → Verde
37	Rojo	Rojo → Verde
38	Verde	Rojo → Verde
39	Rojo	Rojo → Verde
40	Verde	Rojo → Verde * fin ciclo 1
41	Rojo	Rojo → Verde
42	Verde	Rojo → Verde
43	Rojo	Rojo → Verde
44	Verde	Rojo → Verde
45	Rojo * fin ciclo 1	Rojo → Verde

MÚLTIPLE EXT-EXT ALTERNO CICLO 1

DÍA	CAMBIO ENTRE SESIONES	CAMBIO INTRASESIONES
1	Verde	Rojo → Verde
2	Rojo	Rojo → Verde
3	Verde	Rojo → Verde
4	Rojo	Rojo → Verde

MÚLTIPLE EXT-EXT ALTERNO CICLO 2

DÍA	CAMBIO ENTRE SESIONES	CAMBIO INTRASESIONES
1	Verde	Rojo → Verde
2	Rojo	Rojo → Verde
3	Verde	Rojo → Verde
4	Rojo	Rojo → Verde
5	Verde	Rojo → Verde
6	Rojo	Rojo → Verde
7	Verde	Rojo → Verde
8	Rojo	Rojo → Verde
9	Verde	Rojo → Verde
10	Rojo	Rojo → Verde
11	Verde	Rojo → Verde
12	Rojo	Rojo → Verde
13	Verde	Rojo → Verde
14	Rojo	Rojo → Verde

Experimento 2.**Múltiple RV30-EXT. Cambio Aleatorio Diario y Cambio Intrasesión**

MÚLTIPLE RV30-EXT ALEATORIO		
DÍA	CAMBIO ENTRE SESIONES Componente rico	CAMBIO INTRASESIONES Componente Rico
1	Verde	Rojo → Verde
2	Rojo	Rojo → Rojo
3	Rojo	Rojo → Verde
4	Verde	Rojo → Rojo
5	Verde	Rojo → Verde
6	Verde	Rojo → Rojo
7	Rojo	Rojo → Rojo
8	Rojo	Rojo → Verde
9	Verde	Rojo → Rojo
10	Rojo	Rojo → Rojo
11	Verde	Rojo → Rojo
12	Verde	Rojo → Verde
13	Rojo	Rojo → Verde
14	Verde	Rojo → Rojo
15	Verde	Rojo → Rojo
16	Verde	Rojo → Rojo
17	Rojo	Rojo → Verde
18	Rojo	Rojo → Rojo
19	Verde	Rojo → Verde
20	Verde	Rojo → Verde
21	Rojo	Rojo → Verde
22	Verde	Rojo → Rojo
23	Rojo	Rojo → Rojo
24	Rojo	Rojo → Verde
25	Rojo	Rojo → Rojo
26	Verde	Rojo → Verde
27	Rojo	Rojo → Verde
28	Rojo	Rojo → Verde
29	Verde	Rojo → Rojo
30	Verde	Rojo → Verde

MÚLTIPLE RV30-EXT ALEATORIO (Cont.)

DÍA	CAMBIO ENTRE SESIONES Componente rico	CAMBIO INTRASESIONES Componente Rico
31	Rojo	Rojo → Rojo
32	Verde	Rojo → Verde
33	Rojo	Rojo → Verde
34	Verde	Rojo → Rojo
35	Verde	Rojo → Rojo
36	Rojo	Rojo → Rojo

MÚLTIPLE EXT-EXT ALEATORIO

DÍA	CAMBIO ENTRE SESIONES	CAMBIO INTRASESIONES
1	Verde	Rojo → Verde
2	Rojo	Rojo → Rojo
3	Rojo	Rojo → Verde
4	Verde	Rojo → Rojo
5	Verde	Rojo → Verde
6	Verde	Rojo → Rojo
7	Rojo	Rojo → Rojo
8	Rojo	Rojo → Verde

Experimento 3. Múltiple RV30-EXT. Transición Probabilística Intrasesión

MÚLTIPLE RV30-EXT. TRANSICIÓN PROBABILÍSTICA INTRASESIÓN

DÍA	SECUENCIA DE COLOR RICO
-----	-------------------------

1	Verde → Rojo
2	Rojo → Rojo
3	Verde → Rojo
4	Verde → Rojo
5	Verde → Verde
6	Rojo → Rojo
7	Verde → Verde
8	Rojo → Verde
9	Rojo → Verde
10	Rojo → Rojo
11	Verde → Rojo
12	Verde → Rojo
13	Rojo → Verde
14	Rojo → Rojo
15	Verde → Verde
16	Rojo → Verde
17	Rojo → Verde
18	Rojo → Verde
19	Verde → Verde
20	Rojo → Rojo
21	Verde → Rojo
22	Rojo → Rojo
23	Verde → Rojo
24	Verde → Verde
25	Rojo → Verde
26	Verde → Verde
27	Verde → Rojo
28	Verde → Verde
29	Rojo → Verde
30	Rojo → Verde
31	Verde → Verde

MÚLTIPLE RV30-EXT. TRANSICIÓN PROBABILÍSTICA INTRASESIÓN (Cont.)

DÍA	SECUENCIA DE COLOR RICO
-----	-------------------------

32	Verde→Rojo
33	Verde→Rojo
34	Verde→Rojo
35	Verde→Verde
36	Rojo→Rojo
37	Verde→Verde
38	Rojo→Verde
39	Rojo→Rojo
40	Rojo→Rojo
41	Verde→Rojo
42	Rojo→Verde
43	Rojo→Rojo
44	Verde→Verde
45	Rojo→Verde
46	Rojo→Verde
47	Rojo→Verde
48	Verde→Verde
49	Rojo→Rojo
50	Verde→Rojo
51	Rojo→Rojo
52	Verde→Rojo
53	Verde→Verde
54	Rojo→Verde
55	Verde→Verde
56	Verde→Verde
57	Rojo→Verde
58	Rojo→Verde
59	Verde→Verde
60	Verde→Rojo
61	Rojo→Rojo
62	Verde→Rojo
63	Verde→Rojo
64	Verde→Verde
65	Rojo→Rojo
66	Rojo→Verde

MÚLTIPLE RV30-EXT. TRANSICIÓN PROBABILÍSTICA INTRASESIÓN (Cont.)

DÍA	SECUENCIA DE COLOR RICO
67	Rojo→Rojo
68	Verde→Rojo
69	Verde→Rojo
70	Rojo→Verde
71	Rojo→Rojo
72	Verde→Verde
73	Rojo→Verde
74	Rojo→Verde
75	Verde→Rojo
76	Verde→Verde
77	Rojo→Rojo
78	Rojo→Verde
79	Rojo→Rojo
80	Verde→Rojo
