

## Adquisición y Dinámica de Preferencias en Programas Múltiples de Reforzamiento

Oscar Zamora Arévalo<sup>1</sup> & Arturo Bouzas Riaño  
Universidad Nacional Autónoma de México

### Resumen

Ocho palomas fueron entrenadas en programas múltiples de reforzamiento Razón Variable-Razón Variable (mult RV-RV) con cambios rápidos e imprevistos en las distribuciones de refuerzo en ambos componentes del programa múltiple. El objetivo principal fue evaluar cómo se ajustan las tasas de respuestas a cambios abruptos y no señalados en las condiciones de reforzamiento en distintos períodos y en particular determinar si la dinámica del ajuste del comportamiento es dirigida por la razón o por la diferencia en las probabilidades en las tasas de reforzamiento obtenido en dos componentes de un programa múltiple. Los principales hallazgos fueron que cuando las diferencias entre dos programas de reforzamiento (pobre y rico) son constantes, el desarrollo de la preferencia por una de las respuestas del programa múltiple fue más rápido cuando la razón de las probabilidades de reforzamiento fue mayor (5 a 1), lo cual es congruente con los resultados obtenidos en programas de reforzamiento concurrentes. Sin embargo, cuando la razón se mantuvo constante, la tasa de adquisición no fue más rápida cuando la diferencia entre la probabilidad de reforzamiento fue mayor, hallazgo distinto al reportado en experimentos con programas concurrentes. Los resultados resaltan la importancia de la discriminabilidad en probabilidades de reforzamiento entre la fase de entrenamiento y de transición.

*Palabras clave:* Elección, Dinámica del comportamiento, Adquisición, Programas múltiples, Pichones.

### Acquisition and Preference Dynamic in Multiple Schedules of Reinforcement

### Abstract

Eight pigeons were trained on multiple variable ratio-variable ratio schedules of reinforcement (mult VR-VR) with rapid and unexpected changes in reinforcement distribution within both components of the multiple schedule. The main objective of the study was to assess the adjustment of response rates to abrupt and unsignaled changes in the conditions of reinforcement in different periods, particularly whether the dynamics of behavioral change is guided by the ratio or by the difference between the probabilities of reinforcement obtained in the two components of the multiple schedule. The main findings were that when the differences between two schedules of

---

<sup>1</sup> Este escrito fue financiado por el subsidio del Proyecto CONACYT N° 104396 y el proyecto UNAM-PAPIIT-IN307310. Los autores agradecen a Vladimir Orduña la revisión previa de este escrito. Enviar correspondencia a Oscar Zamora Arevalo o Arturo Bouzas Riaño a: Ave. Universidad 3004, Col. Copilco Coyoacán, 1er Piso Edificio D, Posgrado, Facultad de Psicología, Universidad Nacional Autónoma de México, C. P. 04510. Dirección de correo electrónico: ozamoraa@gmail.com; abouzasr@gmail.com

reinforcement (lean and rich) are constant, the development of preference for one of the responses in the multiple schedule was faster when the ratio of the probability of reinforcement was higher (5 to 1), which is consistent with the results obtained with concurrent schedules of reinforcement. However, different from the results obtained using concurrent schedules, when the ratio of probabilities remained constant the rate of acquisition was not faster when the difference between the probabilities of reinforcement was higher. The results highlight the importance of the discriminability of reinforcement probability between training and transition phases.

**Keywords:** Choice, Behavioral dynamics, Acquisition, Multiple schedules, Pigeons.

De los numerosos estudios que sobre elección y preferencia se han realizado en casi cinco décadas (Davison & McCarthy, 1988; Herrnstein, 1997; Stephens, Brown, & Ydelberg, 2007; Stephens & Krebs, 1986; Williams, 1988, 1994), la vasta mayoría han utilizado un modelo experimental que consiste en el estudio de la distribución de dos respuestas, cada una reforzada de acuerdo a programas de reforzamiento independientes (programas concurrentes, programas concurrentes encadenados y procedimientos concurrentes de ensayo discreto) y se han enfocado sobre conducta de estado estable. En la mayor parte de esos estudios, el énfasis teórico ha sido sobre principios de equilibrio molar, estático y reversible de entre los cuales ha destacado la ley de igualación (Davison & McCarthy, 1988; Hernstein, 1997; Williams, 1988, 1994). Esta regla de elección es una relación empírica que bajo condiciones apropiadas establece que el comportamiento es distribuido en proporción al reforzamiento obtenido en las distintas alternativas disponibles. Igualación es un principio de equilibrio (más que una ley causal) ya que relaciona dos cantidades, tasas de respuesta y tasas de reforzamiento, que son mutuamente dependientes. El utilizar metodologías de estado estable ha estimulado el desarrollo de modelos cuantitativos que son capaces de describir la ejecución individual con un elevado nivel de precisión (e.g., Baum, 1974; Davison & McCarty, 1988; Grace & McLean, 2006; Mazur 2001), pero no han dado una respuesta clara a la cuestión del mecanismo que subyace la igualación. Por lo tanto, no es de extrañar que los cambios en las tasas relativas de respuesta ante cambios en las tasas de reforzamiento sea un tema de creciente interés en investigación, ya que entender la dinámica de la preferencia por responder a una de dos alternativas en transición puede dar una visión de los procesos responsables del comportamiento en estado estable.

En los últimos 20 años ha habido un creciente interés acerca de cómo se desarrollan tales preferencias en programas de reforzamiento concurrentes utilizando una amplia variedad de especies y procedimientos experimentales. En un conjunto de experimentos se examinó la elección bajo contingencias que cambian impredeciblemente a través de las sesiones (e.g., Grace, Bragason, & McLean, 2003; Kyonka & Grace, 2008; Maguire, Hughes, & Pitts, 2007; Schofield & Davison, 1997). En otra serie de estudios, Davison, Baum y sus colegas (Baum & Davison, 2004; Davison & Baum, 2000, 2002; 2003; 2006; 2007; Krägeloh & Davison, 2003; Landon & Davison, 2001; Landon, Davison & Elliffe, 2003) emplearon un nuevo método para investigar la dinámica de la elección. Utilizando un variante de un procedimiento descrito originalmente por Belke y Heyman

(1994), palomas son expuestas a programas concurrentes en el que dada una dimensión del refuerzo por ejemplo, tasa, cantidad, o demora, éstas cambian en varias ocasiones dentro de cada sesión.

En un conjunto de experimentos, antecedentes directos del actual, Mazur y sus colegas utilizaron un procedimiento con el cual, para cada sujeto, es posible obtener varias curvas de la adquisición de la preferencia por responder a una de dos alternativas, ante cambios en la tasa de reforzamiento que producen (Bailey & Mazur, 1990; Mazur, 1992; Mazur & Ratti, 1991). En sus estudios, expusieron a palomas a programa concurrentes. Cada condición consistió de un número de sesiones con iguales programas de reforzamiento para las dos respuestas, las cuales eran seguidas de sesiones en las que sorpresivamente una tecla empezaba a dar más reforzadores que la otra. Esta combinación de sesiones se repetía un número de veces. Con este procedimiento, examinaron el efecto de variar la razón y la diferencia entre la probabilidad de reforzamiento en las condiciones de transición de reforzamiento.

En el estudio reportado por Bailey y Mazur (1990) en un primer experimento sobre la tasa de adquisición de una preferencia, se observó a ocho palomas durante 10 condiciones de transición en un procedimiento de ensayos discretos de dos elecciones (en este tipo de procedimiento, al sujeto sólo se le permite hacer una sola respuesta por ensayo, la cual puede o no ser reforzada). Cada condición comenzó con sesiones con la misma probabilidad de reforzamiento en las dos teclas, la que cambiaba, por un número de sesiones, de tal manera que una tecla tenía una mayor probabilidad de reforzamiento. Los investigadores observaron 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 fue más alta. Además, observaron que los animales fueron sensibles a las razones entre las probabilidades de reforzamiento y no a las diferencias entre dichas probabilidades.

Los experimentos sobre el cambio en las preferencias por respuestas que proveen diferentes tasas de reforzamiento, han estudiado en forma exclusiva alguna variante de programas de reforzamiento concurrente. La simultaneidad de las opciones de respuesta y de reforzamiento tiene dos consecuencias. La primera es que la única medida de preferencia es la tasa relativa de respuesta, dado que la competencia entre respuestas por el tiempo disponible necesita que el incremento en la tasa de una respuesta vaya acompañado del decremento en la tasa de la otra opción de respuesta. La segunda consecuencia, es el interés teórico y experimental por estudiar prioritariamente el efecto de las tasas de reforzamiento obtenidas por las respuestas simultáneas sobre la formación de la preferencia entre respuestas.

Los programas múltiples brindan una oportunidad diferente para estudiar la elección entre respuestas. En estos programas las opciones de respuesta están temporalmente separadas en dos componentes y consecuentemente no compiten por un tiempo común (Herrnstein, 1970; Williams, 1983). En equilibrio, la tasa de respuesta durante uno de los componentes depende tanto de su tasa de reforzamiento como de la tasa de reforzamiento obtenida en el componente alterno (Williams, 1983). Los modelos de elección se extendieron con éxito para

dar cuenta de estos resultados. Sin embargo, los autores del presente estudio no conocen reportes publicados dedicados a estudiar los cambios en las tasas de respuesta en programas múltiples en períodos de transición similares a los estudiados con programas de reforzamiento concurrentes. Este trabajo es una primera aproximación para abordar este problema. El procedimiento que se empleó es equivalente al utilizado por Mazur y sus colegas para el caso de la elección simultánea (Bailey & Mazur, 1990; Mazur, 1992; 1995; 1996, 1997; Mazur & Ratti, 1991). En programas de reforzamiento múltiples RV-RV, con iguales requisitos de respuesta, se estudió el impacto sobre las tasas de respuesta de cambios abruptos en los valores de los programas RV. Al igual que en los estudios de Mazur y colaboradores, se varió tanto la razón como la diferencia entre las probabilidades de reforzamiento asociadas con las respuestas en los dos componentes, con la intención de determinar si su efecto es similar en situaciones de elección simultánea como sucesiva. En unas condiciones se mantuvo constante la diferencia en la probabilidad de reforzamiento en los dos componentes variándose su razón. En otras condiciones se mantuvo constante la razón de probabilidades variándose la diferencia entre ellas.

## **Método**

### *Sujetos*

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

### *Aparatos*

Se utilizaron ocho cámaras idénticas de condicionamiento operante para pichones diseñadas por el personal del Laboratorio de Comportamiento y Adaptación, con las siguientes dimensiones interiores, 37 cm de alto, 30 cm de ancho y 35 cm de fondo. Cada caja contaba con una luz general, localizada en el centro del techo de la caja y tres teclas de respuestas translúcidas con su correspondiente proyector de estímulos. Las teclas se encontraban en la pared frontal de la cámara y estaban separadas por 7 cm y cada una media 2.7 cm de diámetro. Las teclas izquierda y central no fueron utilizadas en este experimento. La tecla derecha podía ser iluminada por una luz roja o verde producidas por dos focos de 2 watts fijados detrás de la tecla. Se requirió de una fuerza de aproximadamente 0.15 N para operar cada tecla. Las cajas también contaban con un dispensador electromagnético de alimento (grano balanceado) que se localizaba en la pared con las teclas, a 5.5 cm del piso y a 10 cm de la tecla central. El reforzador consistió de dos segundos de acceso libre a una mezcla de granos en el dispensador. Durante ese tiempo se encendió una luz que iluminó el

dispensador (comedero) con una luz blanca proveniente de un foco de 2 watts fijado sobre éste y se apagó tanto la luz general como la de la tecla derecha. Las cámaras experimentales estuvieron ventiladas todo el tiempo que duró la sesión y cada una tenía un extractor que sacaba el aire del interior de la cámara, y que a su vez servía como generador de ruido blanco que atenuaba el sonido proveniente del exterior. El control de la secuencia de presentación de los estímulos dentro de las cámaras experimentales, así como el registro de las respuestas de cada animal, se realizó empleando una computadora por medio de una interfaz y lenguaje de programación (MEDSTATE®). Todos los eventos de una sesión se registraron con una exactitud de mili segundos. Los datos fueron almacenados en discos flexibles para su posterior tratamiento.

### *Procedimiento*

Los sujetos fueron alojados en cajas-habitación individuales teniendo los primeros 20 días acceso libre tanto al agua como al alimento. Durante este período, se registró diariamente su peso, posteriormente se tomó la media de este registro como punto de comparación posterior (peso ad libitum). Después de este período, se restringió paulatinamente la cantidad de alimento disponible hasta que se alcanzó el criterio de 80% (+5grm) de peso con respecto al valor ad libitum.

Después de 10 días en los que se mantuvo a los sujetos al 80% de su peso normal, se pasó a la fase de entrenamiento del picoteo mediante un programa de automoldeamiento del picoteo a la tecla (Brown & Jenkins, 1968), hasta conseguir de manera consistente la respuesta de picoteo en la tecla derecha, ante dos colores (verde y rojo). Posteriormente, para familiarizar a los sujetos con el programa de reforzamiento, se reforzó la respuesta a la misma tecla con programas RV, con valores que incrementaron hasta alcanzar un RV 100 en una semana de entrenamiento. Este requisito de respuesta fue el más alto de los utilizados en las condiciones experimentales.

Las palomas fueron expuestas a cinco condiciones experimentales. Cada condición se repitió cuatro veces y los sujetos pasaron por las condiciones en distinto orden. Cada condición, consistió de dos fases, la primera fue una fase de estado estable o entrenamiento y la segunda de transición o de prueba. Las fases de estado estable incluyeron tres días de entrenamiento, en los cuales las respuestas en los dos componentes fueron reforzadas con el mismo programa de reforzamiento (RVx), es decir, con la misma probabilidad de reforzamiento. Durante el cuarto y quinto día de cada condición experimental, se expuso a las palomas a dos sesiones diarias de transición o fase de prueba, en las que un par de programas de reforzamiento diferentes (Razón Variable RVx, Razón Variable RVy), operaron en cada componente. Los valores para las razones utilizadas fueron creados a partir de las progresiones sugeridas por Freshler y Hoffman (1962). Los componentes del programa múltiple duraron 60 s cada uno y las sesiones finalizaron a los 30 min. Las sesiones se condujeron durante seis días a la semana. En la primera parte de este experimento, durante las fases de prueba se usaron pares de programas de reforzamiento que generaron razones iguales

de probabilidades de reforzamiento (5 a 1), pero que generaron distintas diferencias en las probabilidades de reforzamiento (.16, .08 y .04). En la segunda parte, se usaron pares de programas de reforzamiento que producían una razón de probabilidad de 2 a 1 y 3 a 1 e iguales diferencias en las probabilidades de reforzamiento (.04), lo anterior para comparar el efecto de una razón de probabilidades diferentes, y el efecto de diferencias constantes en probabilidad. Las cinco condiciones experimentales con valores de los RVs y con las razones y las diferencias absolutas en entrenamiento y transición se indican en la Tabla 1.

Tabla 1

*Descripción de los programas de reforzamiento utilizados asociados a los componentes del programa múltiple para cada una de las fases y condiciones experimentales.*

Condición	Fase de Estado		Fase de Prueba o de Transición	Probabilidad Alta (RVx)	Probabilidad Baja (RVy)	Razón	Dif.	
	Estable	o de Entrenamiento						
A	0.12	RV 8 - RV 8	0.2	RV 5	0.04	RV 25	5a1	0.16
B	0.06	RV 17 - RV 17	0.1	RV 10	0.02	RV 50	5a1	0.08
C	0.03	RV 33 - RV 33	0.05	RV 20	0.01	RV 100	5a1	0.04
D	0.06	RV 17 - RV 17	0.08	RV 13	0.04	RV 25	2a1	0.04
E	0.04	RV 25 - RV 25	0.06	RV 17	0.02	RV 50	3a1	0.04

La luz general de la caja y la tecla derecha permanecieron encendidas a lo largo de la sesión, excepto durante los períodos de reforzamiento. La tecla derecha fue transiluminada de color rojo o verde. Cada color se asoció con cada uno de los componentes del programa múltiple RV-RV tanto en la fase de entrenamiento como en la fase de prueba. En las fases de transición para la mitad de los sujetos el color rojo fue asociado al componente rico y el color verde se asoció al componente pobre y viceversa para la otra mitad de los sujetos.

Las tasas de respuesta tanto absolutas como relativas se calcularon componente a componente para los tres días de entrenamiento y en cada uno de los dos días de transición.

## Resultados

El interés principal de este experimento fue observar el comportamiento de las respuestas en un programa múltiple en los períodos de transición de componentes con igual probabilidad de reforzamiento a componentes con probabilidades diferentes de reforzamiento. Para cada una de las condiciones experimentales a las que fueron sometidos los sujetos, se reportan tanto las tasas relativas como las tasas absolutas de respuesta computadas componente a componente. Los últimos dos componentes no se consideraron para el análisis gráfico y estadístico, ya que con valores tan pequeños de RV en los distintos programas de las diferentes condiciones, se presentó un efecto de saciedad que se manifestó en ambos componentes en todas las condiciones experimentales.

Para ambas gráficas de la Figura 1 se presentan los resultados para el segundo día de transición, en las abscisas se muestran intervalos sucesivos de 2 minutos, en los cuales los sujetos respondían a uno y a otro componente (rico y pobre) del programa múltiple. En tanto en las ordenadas se muestra la tasa relativa de respuestas, es decir la tasa de respuestas en el componente rico ( $P_r$ ) dividida por la tasa de respuestas en el componente rico más la tasa de respuestas en el componente pobre ( $P_r + P_p$ ). Esta medida fue obtenida para cada uno de los 14 intervalos sucesivos de dos minutos que constituyeron cada sesión. Los símbolos en cada una de las gráficas representan el promedio total de la tasa relativa de respuestas de los sujetos, en los cuatro ciclos para el segundo día de las fases de prueba o transición para cada una de las condiciones. En el panel superior se presenta la tasa relativa de respuestas en las tres condiciones cuando la razón entre las probabilidades de reforzamiento de los componentes de los programas múltiples se mantuvieron constantes en 5 a 1 (Condiciones RV5-RV25, RV10-RV50 y RV20-RV100), pero la diferencia entre las probabilidades de reforzamiento de ambos componentes fue variable (0.16, 0.08, 0.04). En el panel inferior se muestra la tasa relativa de respuestas en las tres condiciones cuando la diferencia entre las probabilidades de reforzamiento fue la misma (0.04) en las tres condiciones (RV13-RV25, RV17-RV50, RV20-RV100) pero la razón entre las probabilidades de reforzamiento fue variable (2 a 1, 3 a 1 y 5 a 1).

La Figura 1 muestra que la velocidad con la que cambiaron las tasas relativas de respuesta fue proporcional a la razón de la probabilidad de reforzamiento en los dos componentes del programa múltiple. Muestra también una diferencia muy pequeña y no ordenada entre la tasa relativa de respuesta y la diferencia en las probabilidades de reforzamiento. Esta impresión fue confirmada por un análisis de varianza múltiple de medidas repetida con dos factores: Condición e Intervalos con 3 y 14 niveles respectivamente, tanto en las condiciones donde la razón fue constante (Condiciones RV5-RV25, RV10-RV50 y RV20-RV100), como en las condiciones donde la diferencia en las probabilidades de reforzamiento fue la misma (Condiciones RV13-RV25, RV17-RV50 y RV20-RV100). En el primer análisis cuando las razones entre las probabilidades de reforzamiento de mantuvieron constantes (5 a 1), sólo el factor intervalo fue estadísticamente significativo  $F(13, 91) = 7.39, p < .05$ , ni el factor condición, ni la interacción entre

condición e intervalo fueron significativas. En el segundo análisis cuando las diferencias entre las probabilidades de reforzamiento se mantuvieron constantes (0.04), ambos factores y la interacción fueron estadísticamente significativos, condición  $F(2, 14) = 12.09, p < .05$ , intervalo  $F(13, 91) = 9.03, p < .05$ , e interacción condición e intervalo  $F(26, 182) = 1.58, p < .05$ .

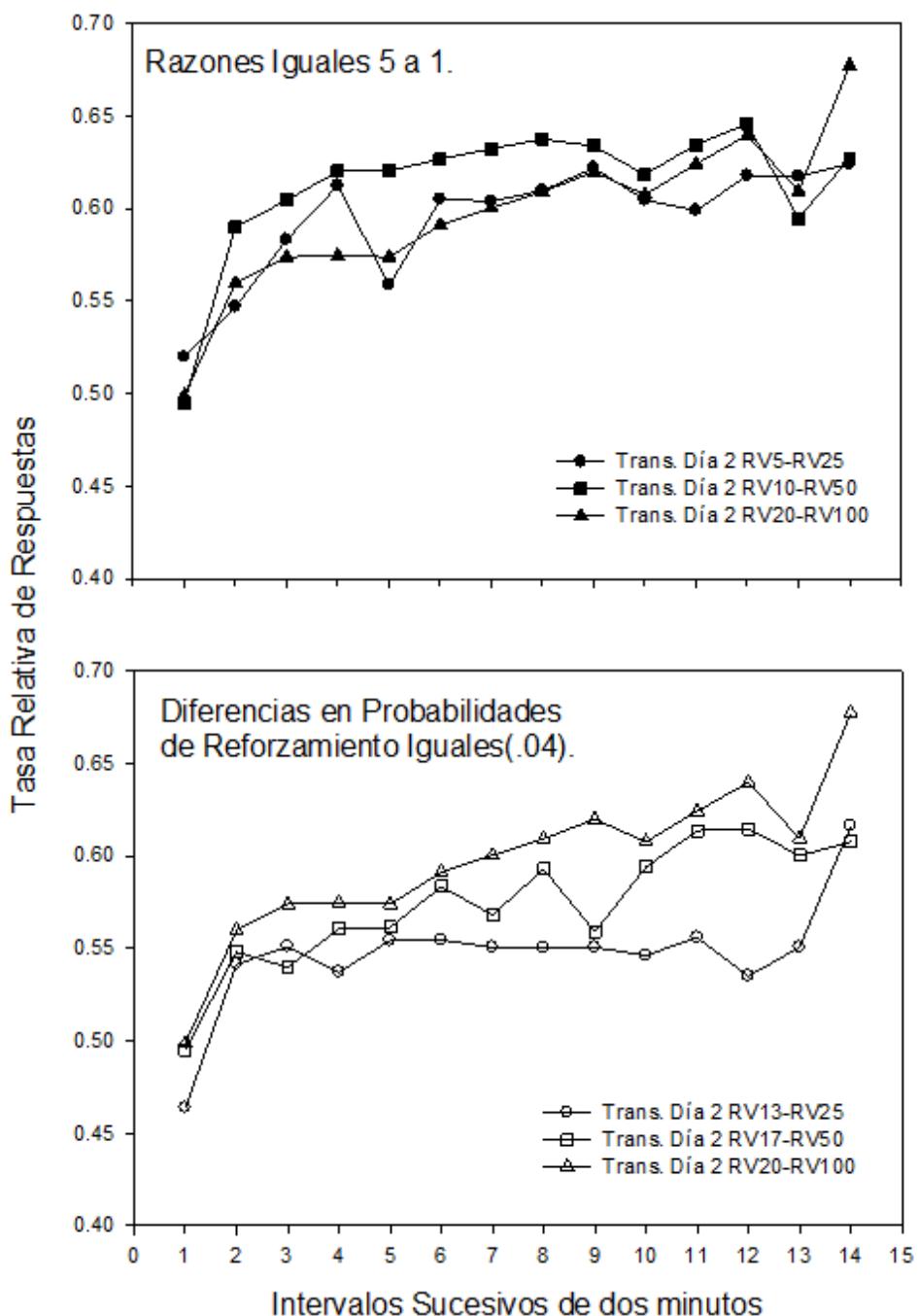


Figura 1. Promedio de todos los sujetos de las tasas relativas de respuesta como una función de intervalos sucesivos de 2 minutos.

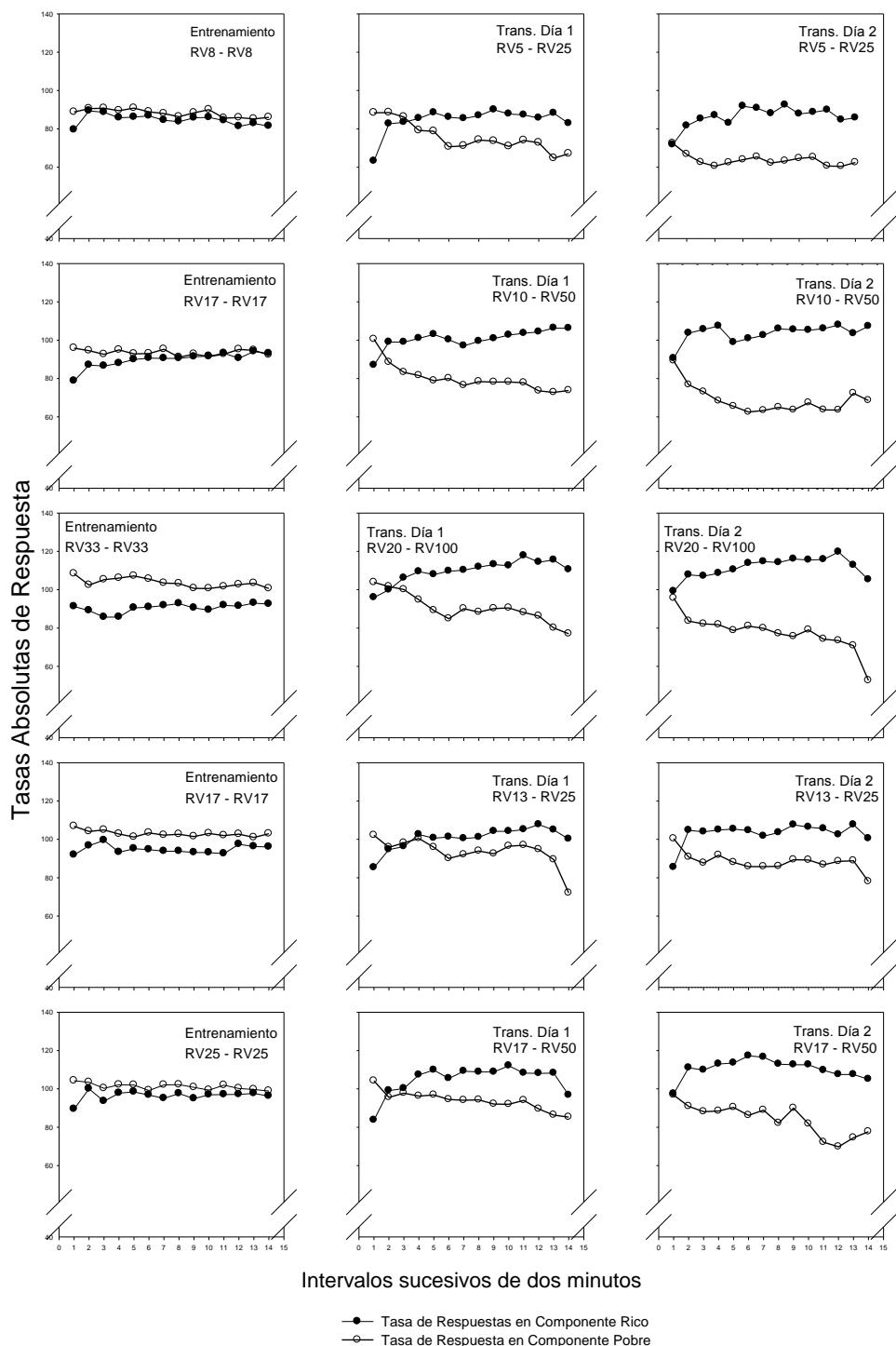
A diferencia de los programas concurrentes, en programas múltiples es pertinente estudiar los posibles cambios en las tasas absolutas de respuesta.

La Figura 2 muestra el promedio de las tasas absolutas de respuesta de todos los sujetos (número de respuesta/tiempo de exposición al componente para cada uno de los 14 componentes {ordenada}). Las tres columnas muestran respectivamente el promedio de los tres días de las fases de entrenamiento, el primer día de las fases de transición y el segundo día de las fases de transición. Los círculos llenos representan el componente asociado en las fases de transición al componente rico y los círculos vacíos representan el componente que en las fases de transición estuvo asociado al componente pobre. Se puede observar de manera muy general que durante las fases de entrenamiento, las tasas globales de respuesta fueron similares en ambos componentes de los programas múltiples. En cambio en todas las condiciones en los días de transición desde el primer día y el segundo intervalo, las tasas absolutas cambiaron, pero el cambio fue más pronunciado para el componente pobre. Este efecto fue aún mayor en el segundo día de transición, y en las tres condiciones con razón 5 a 1. Un efecto adicional, que puede verse en esta figura, es que en el primer intervalo del segundo día de transición, las tasas absolutas regresaron a los niveles observados en entrenamiento.

## Discusión

El propósito de este experimento fue presentar la evidencia empírica faltante sobre los factores que afectan la tasa de adquisición de una preferencia a responder diferencialmente cuando súbitamente dos alternativas temporalmente distantes proporcionan reforzadores con diferentes probabilidades.

La pregunta que se deseaba responder era si los resultados obtenidos con programas múltiples eran similares a los reportados en programas concurrentes manipulando variables similares. Los resultados permiten una serie de conclusiones. Primero, manteniendo constante la diferencia en la probabilidad de reforzamiento entre los componentes del programa múltiple, el cambio en las tasas de respuesta, absolutas y relativas, fue mayor y más rápido mientras mayor fue la razón entre dichas probabilidades de reforzamiento. Segundo, manteniendo constante la razón de probabilidades, la diferencia en la probabilidad de reforzamiento no tuvo un efecto, ni consistente ni significativo, sobre los niveles y la rapidez de los cambios en ambas medidas de respuesta. El primer resultado es consistente con los reportados por Mazur y colaboradores en programas concurrentes en procedimientos de ensayo discreto (Bailey & Mazur, 1990) y procedimientos de operante libre (Mazur, 1992; Mazur & Ratti, 1991). Sin embargo, el segundo no es consistente a lo reportado en los mismos experimentos.



*Figura 2.* Promedio total de las tasas globales de respuesta como una función de intervalos sucesivos de 2 minutos. Los círculos cerrados representan las tasas de respuesta asociadas a los componentes ricos y los círculos abiertos representan las tasas de respuesta asociadas a los componentes pobres tanto en las fases de entrenamiento como de transición.

Tercero, el cambio en la tasa relativa de respuesta obedece principalmente al decremento en la tasa absoluta de respuesta durante el componente pobre del programa múltiple de reforzamiento. Cuarto, el cambio en las tasas de respuesta fue mayor en el segundo día de transición que en el primero, sin embargo la experiencia con sólo dos componentes de un minuto fue suficiente para observar cambios en ambas medidas de respuesta. Quinto, la tasa de respuesta en el segundo día de transición en los minutos iniciales de la sesión regresó a los niveles obtenidos durante la fase de entrenamiento.

De estos resultados, uno en particular, el efecto de la diferencia en la probabilidad de reforzamiento en los dos componentes del programa múltiple resalta por su inconsistencia con los resultados obtenidos en programas de reforzamiento concurrentes. En particular, los resultados reportados por Mazur y colaboradores, implican que la rapidez en el cambio en las tasas absolutas de respuesta está relacionada monotónicamente a la tasa de reforzamiento global que conforma el contexto de una sesión experimental. Por ejemplo, comparando las condiciones RV5 – RV25 con RV20- RV100, ambas condiciones proporcionan reforzamiento con la misma razón de probabilidades (5 a 1), sin embargo, la primera condición permite una tasa de reforzamiento mucho mayor que la de la segunda. Consecuentemente, la primera condición proporciona un número mayor de oportunidades de aprendizaje por respuesta. Los datos aquí reportados son inconsistentes con esta interpretación del papel de la diferencia en probabilidad de reforzamiento en programas múltiples.

Una posible explicación de la diferencia en el efecto de la densidad de reforzamiento en programas concurrentes y múltiples, es las diferentes restricciones temporales asociadas con ellos. En el primer procedimiento hay una restricción lineal que implica que las respuestas compiten por el mismo tiempo disponible. En contraste, en los programas múltiples hay dos restricciones lineales, cada una de ellas asociada a uno de los componentes, y consecuentemente las respuestas reforzadas no compiten por un tiempo común. La implicación de esta diferencia en restricciones, es que en programas concurrentes, la competencia por el tiempo disponible magnifica el papel de la densidad de reforzamiento, de forma tal que la respuesta reforzada con mayor probabilidad, en este caso la respuesta reforzada con el RV 5, absorbe una mayor parte del tiempo disponible.

El efecto de la razón de probabilidades sobre la rapidez con la que cambian las tasas de respuesta es consistente con la posición que atribuye dicho efecto a la discriminabilidad entre la probabilidad de reforzamiento en el componente rico y el pobre. Al igual que en otras dimensiones sensoriales la discriminabilidad depende de las razones y no de las diferencias entre estímulos. La razón de probabilidades generada por el programa múltiple RV 20 RV 100 es mucho más discriminable que la razón que resulta del programa RV 13 RV 25, a pesar de que en ambas condiciones la diferencia entre probabilidades de reforzamiento es la misma (0.04).

La literatura en programas concurrentes asume que el proceso de comparación es entre la probabilidad de reforzamiento asociada con cada respuesta disponible simultáneamente. En el caso de los programas múltiples, las

diferentes extensiones de la ley del efecto relativa para dar cuenta de la ejecución que se obtiene en equilibrio (Williams, 1983), implican que por analogía, la comparación en transiciones es también entre las probabilidades de reforzamiento programadas para los dos componentes. Sin embargo, el patrón observado en las curvas de adquisición de las tasas absolutas de respuesta, sugieren reconsiderar las variables que operan en el aprendizaje en los programas múltiples. En el caso de los programas múltiples, no es fácil comparar las respuestas reforzadas en diferentes componentes, por lo que la adaptación a las nuevas condiciones en transición parece darse comparando la probabilidad de reforzamiento en cada componente separado, con las condiciones de reforzamiento en el mismo componente durante la fase de entrenamiento. Una forma de evaluar dicha hipótesis es tomando ventaja de que los programas múltiples permiten separar las tasas relativas de respuesta de las absolutas. El patrón de resultados observados en la Figura 2 es consistente con la suposición de que la discriminabilidad entre la probabilidad de reforzamiento en la fase de entrenamiento y la de transición es la variable importante. Así, el cambio más discriminativo en reforzamiento se dio en la condición RV20 – RV100, que en entrenamiento era RV33. Mientras que el cambio menos discriminativo en reforzamiento se obtuvo en la condición RV5 – RV25, que en entrenamiento era RV8. Por ejemplo, para el caso del programa múltiple RV 20 RV 100, con un programa en entrenamiento RV 33, las probabilidades de reforzamiento fueron respectivamente .05, .01 y .033. En este caso los animales compararían la razón de .05 a .033 igual a sólo 1.667 veces más probable el reforzamiento, en el componente rico y .01 a .033 igual a .333 (un tercio menos probable) en el componente pobre.

Consistente con la hipótesis que los animales comparan, para cada respuesta por separado, las probabilidades de reforzamiento en entrenamiento y transición, fue lo obtenido cuando la razón de probabilidades de reforzamiento fue igual a 5 a 1. En las tres condiciones hubo un decremento en la tasa de respuesta durante el componente pobre, conforme decremento su probabilidad de reforzamiento y simultáneamente, la tasa de respuesta en el componente rico también disminuyó conforme incremento su probabilidad de reforzamiento. Un efecto de techo es una explicación alternativa del menor cambio en la tasa absoluta de respuesta en el componente rico. Sin embargo, en otros experimentos con programas múltiples RV-RV la tasa de respuesta para requisitos de respuesta similares fue mucho más alta que la obtenida en el presente experimento (e.g., Zuriff, 1970).

Lo anterior no quiere decir que las condiciones de reforzamiento en el componente alterno durante transición no jueguen un papel. Es bien conocido que en programas múltiples, manteniendo otras variables constantes, la tasa de respuesta en un componente es una función inversa de la tasa de reforzamiento en el componente alterno, fenómeno conocido como contraste conductual (Williams, 1983). El diseño del estudio no tenía como uno de sus objetivos estudiar el curso del contraste conductual observado en programas múltiples. Sin embargo, en la Figura 2 puede verse que las condiciones que comparten un RV25 y un RV50, muestran clara evidencia de contraste conductual. La tasa de respuesta asociada al RV50 fue más baja cuando alternó con un RV10 que con un RV17. De igual forma, la tasa de respuesta asociada al RV25 fue más baja cuando alternó

con un RV5 que con RV13. En su conjunto, los resultados obtenidos indican la necesidad de obtener más datos que permitan separar las comparaciones sucesivas de las simultáneas y encontrar cuando una o la otra es más probable. Por ejemplo, es posible que reducir la duración de los componentes del programa múltiple, aumente la probabilidad de que los organismos lleven a cabo una comparación simultánea.

En resumen, los datos aquí reportados sugieren que para integrar los resultados de los experimentos en elección simultánea con los de elección intertemporal, los modelos cuantitativos de la adaptación a cambios en las condiciones de reforzamiento en programas múltiples, deben considerar el papel de las diferencias en restricciones temporales entre programas múltiples y concurrentes, las variables que determinan el peso de las comparaciones sucesivas y simultáneas en la probabilidad de reforzamiento y la recuperación espontánea frecuentemente observada (Dragoi & Staddon, 1999).

## Referencias

- Bailey, J. T., & Mazur, E. J. (1990). Choice behavior in transition development of preference for the higher probability of reinforcement. *Journal of the Experimental Analysis of Behavior*, 3, 409-422.
- Baum, W. M. (1974). On two types of deviation for the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior*, 22, 231-241.
- Baum, W. M., & Davison, M. (2004). Every reinforcer counts: Visit patterns in the dynamics of choice. *Journal of the Experimental Analysis of Behavior*, 81, 85-127.
- Brown, P. L., & Jenkins, H. M. (1968). Auto shaping of the pigeon's key peak. *Journal of the Experimental Analysis of Behavior*, 11, 151-160.
- Belke, T. W., & Heyman, G. M. (1994). Increasing and signaling background reinforcement: Effect on the foreground response-reinforcer relation. *Journal of the Experimental Analysis of Behavior*, 61, 65-81.
- Davison, M., & Baum W. (2000). Choice in a variable environment: Every reinforcer counts. *Journal of the Experimental Analysis of Behavior*, 74, 1-24.
- Davison, M., & Baum, W. M. (2002). Choice in a variable environment: Effects of blackout duration and extinction between components. *Journal of the Experimental Analysis of Behavior*, 77, 65-89.
- Davison, M., & Baum, W. M. (2003). Every reinforcer counts: Reinforcer magnitude and local preference. *Journal of the Experimental Analysis of Behavior*, 80, 95-129.
- Davison, M., & Baum, W. M. (2006). Do conditional reinforcers count? *Journal of the Experimental Analysis of Behavior*, 86, 269-283.
- Davison, M., & Baum, W. M. (2007). Local effects of delayed food. *Journal of the Experimental Analysis of Behavior*, 87, 241-260.
- Davison, M., & McCarthy, D. (1988). *The matching law: A research review*. Hillsdale, NJ:Erlbaum.

- Dragoi, V., & Staddon, J. E. R. (1999). The dynamics of operant conditioning. *Psychological Review, 106*, 20-61.
- Fleshler, M., & Hoffman, H. S. (1962) A progression for generating variable-interval schedules. *Journal of the Experimental Analysis of Behavior, 5*, 529-530.
- Grace, R. C., & McLean, A. P. (2006). Rapid acquisition in concurrent chains: Evidence for a decision model. *Journal of the Experimental Analysis of Behavior, 85*, 181–202.
- Grace, R. C., Bragason, O., & McLean, A. P. (2003). Rapid acquisition of preference in concurrent chains. *Journal of the Experimental Analysis of Behavior, 80*, 235–252.
- Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior, 13*, 243-266.
- Herrnstein, R. J. (1997). Relative and absolute strength of responses as a function of frequency of reinforcement. En H. Rachlin & D. I. Laibson (Eds.), *The matching law: Papers in psychology and economics* (pp. 15 -21). Cambridge, MA, E. U.: Harvard University Press.
- Kyonka, E. G. E., & Grace, R. C. (2008). Rapid acquisition of preference in concurrent chains when alternatives differ on multiple dimensions of reinforcement. *Journal of the Experimental Analysis of Behavior, 89*, 49–69.
- Krägeloh, C. U., & Davison, M. (2003). Concurrent schedule performance in transition: Changeover delays and signaled reinforcer ratios. *Journal of the Experimental Analysis of Behavior, 79*, 87–109.
- Landon, J., & Davison, M. (2001). Reinforcer-ratio variation and its effects on rate of adaptation. *Journal of the Experimental Analysis of Behavior, 75*, 207–234.
- Landon, J., Davison, M., & Elliffe, D. (2003). Choice in a variable environment: Effects of unequal reinforce distributions. *Journal of the Experimental Analysis of Behavior, 80*, 187–204.
- Maguire, D. R., Hughes, C. E., & Pitts, R. C. (2007). Rapid acquisition of preference in concurrent schedule: Effects of reinforcement amount. *Behavioural Processes, 75*, 213–219.
- 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, 24*, 93-103.
- Mazur, J. E. (1996). Past experience, recency, and spontaneous recovery in choice behavior. *Animal Learning & Behavior, 24*, 1-10.
- Mazur, J. E. (1997). Effects of rate of reinforcement and rate of change on choice behavior in transition. *Quarterly Journal of Experimental Psychology, 50B*, 111–128.
- Mazur, J. E. (2001). Transitional choice behavior in concurrent-chain schedules. *Behavioural Processes, 53*, 171-180.

- 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.
- Schofield, G., & Davison, M. (1997). Nonstable concurrent choice in pigeons. *Journal of the Experimental Analysis of Behavior*, 68, 219–232.
- Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton, NJ., E. U.: Princeton University Press.
- Stephens, D. W., Brown, J. S., & Ydelberg, R. C. (2007). *Foraging: Behavior and ecology*. Chicago, E.U.: The University Chicago Press.
- Williams, B. A. (1983). Another look at the contrast in multiple schedules. *Journal of the Experimental Analysis of Behavior*, 39, 345-384.
- Williams, B. A. (1994). Reinforcement and choice. En N. J. Mackintosh (Ed.), *Animal learning and cognition* (pp. 81–108). San Diego, CA., E.U.: Academic Press.
- 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, Vol. 2: Learning and cognition* (2a Ed., pp. 167-244.) New York, E.U.: Wiley.
- Zuriff, G. E. (1970). A comparison of variable-ratio and variable-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, 13, 369-374.