



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

**FACULTAD DE PSICOLOGÍA
DIVISIÓN DE ESTUDIOS PROFESIONALES**

**REFORZAMIENTO CONDICIONADO EN
PROGRAMAS DE REFORZAMIENTO
ENCADENADOS**

T E S I S

QUE PARA OBTENER EL TÍTULO DE
LICENCIADA EN PSICOLOGÍA

P R E S E N T A

ALMA PATRICIA OLGUIN GARCÍA

DIRECTOR: DR. ROGELIO ESCOBAR HERNÁNDEZ

REVISORA: DRA. ALICIA ROCA COGORDAN

SINODALES: DR. GUSTAVO BACHÁ MÉNDEZ

DR. LUIS RODOLFO BERNAL GAMBOA

DRA. JUDITH MARINA MENEZ DÍAZ



**Facultad
de Psicología**

APOYO DEL PROYECTO PAPPIT No. IT300517, DGAPA UNAM

Ciudad Universitaria, CDMX, 2018



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

A mi madre y a mi hermano

Agradecimientos

La presente tesis fue realizada gracias al apoyo del proyecto PAPIIT no. IT300517, DGAPA UNAM.

Agradezco a la Universidad Nacional Autónoma de México por brindarme la oportunidad y las herramientas para tener una formación profesional.

Agradezco a mi Director de tesis el Dr. Rogelio Escobar Hernández por todas sus enseñanzas, consejos y correcciones a lo largo de mi formación académica. Especialmente porque no sólo me ha enseñado cómo hacer investigación, sino a ver y entender el mundo de una manera distinta.

Agradezco a mi revisora la Dra. Alicia Roca, por sus valiosos comentarios y por todas las lecciones dentro del aula y del laboratorio.

Les agradezco a los miembros del Comité de Sinodales por sus invaluable comentarios: al Dr. Gustavo Bachá por ser una gran influencia para orientarme hacía la investigación desde los primeros semestres de mi carrera; a la Dra. Marina Ménez por su paciencia y por contagiarme de su gran entusiasmo y curiosidad en todo momento; y al Dr. Rodolfo Bernal por su formidable disposición y apoyo en las revisiones de esta tesis.

Agradezco a mis compañeros del laboratorio: Brissa, Katya, Rodrigo, Kenia, Olga y Sandra, por sus comentarios en las versiones preliminares de este trabajo y por acompañarme en esta etapa de formación académica.

A mis amigos, Soledad y Diego, quienes me han acompañado y han enriquecido mi vida con su maravillosa compañía. Al “Súper Animal Team”: Emmanuel, Lupita y Rodrigo, por compartir no sólo el gusto por la investigación sino por brindarme una bonita amistad. A Rodrigo Alba, por todas sus enseñanzas, apoyo y las interminables risas. A

Brissa, quién siempre ha estado dispuesta a escucharme y a compartir una taza de café conmigo. A David, Elisa, Paola, Gustavo, Matteo, Carlos y Ernesto, por todas las experiencias, risas, pláticas y discusiones que hemos tenido. A mis amigos de la ESIME, especialmente a Waldir, quienes me han escuchado hablar de ratas, creyendo que en realidad mi tesis es sobre los “esimios”. A todas las personas que han estado y que están en mi vida, los llevo siempre conmigo.

Finalmente, agradezco infinitamente a mi familia quienes han sido una pieza fundamental en mi desarrollo. A mi abuelito, por todas sus enseñanzas y por ser un hombre que jamás dejará de sorprenderme. A mi abuelita, porque siempre he encontrado calidez en su hombro. A mi hermano, quien nunca me ha abandonado y siempre ha estado a mi lado apoyándome. Y especialmente a mi madre, quien me ha apoyado y amado incondicionalmente; gracias mamá por siempre alentarme y por indicarme el camino correcto en todo momento.

Tabla de contenidos

Resumen	8
Reforzamiento condicionado en programas de reforzamiento encadenados	9
Propósito.....	19
Método.....	22
Sujetos.....	22
Aparatos	22
Procedimiento	23
Fase de Entrenamiento.	23
Programa de reforzamiento encadenado.....	23
Programa de reforzamiento tándem.....	25
Fase experimental.....	25
Condición A: Programa tándem con respuestas de observación.	25
Condición B: Eliminación del estímulo en el eslabón inicial.....	27
Condición C: Eliminación del estímulo en el eslabón terminal.	27
Resultados.....	29
Discusión	50
Referencias	68

Tablas y figuras

Tabla 1. Número de sesiones a las que estuvieron expuestos cada sujeto por condición en la fase experimental.....	28
Tabla 2. Duración promedio de la sesión en cada una de las condiciones a las que estuvieron expuestos cada sujeto.....	29
<i>Figura 1.</i> Respuestas por minuto en la palanca de reforzamiento durante las últimas seis sesiones de cada condición. Para cada uno de los sujetos la escala del eje de la ordenada es distinta.	31
<i>Figura 2.</i> Respuestas por minuto en la palanca de observación durante las últimas seis sesiones de cada condición. Para cada uno de los sujetos la escala es distinta.	34
<i>Figura 3.</i> Promedio de respuestas en la palanca de reforzamiento en presencia de cada estímulo en cada uno de los eslabones del programa. Se muestran las últimas seis sesiones de cada condición.	36
<i>Figura 4.</i> Pausa post reforzamiento en segundos a la palanca de reforzamiento en las últimas seis sesiones de cada condición.	39
<i>Figura 5.</i> Pausa post reforzamiento en segundos a la palanca de observación en las últimas seis sesiones de cada condición. La discontinuidad en la serie de datos del sujeto RP1 para la condición sin E terminal, se debe a que en esa sesión no se emitió alguna respuesta de observación.	40
<i>Figura 6.</i> Proporción de respuestas de observación subsecuentes que ocurrieron con una latencia entre 0 a <4, 4 a <8, 8 a <12, 12 a <16, 16 a <20 y >20 s desde que se apagó el estímulo del eslabón terminal (abscisa) y el estímulo del eslabón inicial (ordenada) durante las seis últimas sesiones de las condiciones en las que estuvieron presentes ambos	

estímulos. La diagonal representa una línea de referencia que indica indiferencia hacía alguno de los estímulos. 44

Figura 7. Porcentaje relativo de respuestas en la palanca de reforzamiento en cada decil del intervalo entre reforzadores como promedio de las últimas seis sesiones de cada condición.

Para la condición donde se presentaban ambos estímulos se promediaron las tres condiciones a las que estuvieron expuestos los sujetos RP2 y RP3, y las últimas dos exposiciones a esta condición para el sujeto RP1. La línea punteada indica que antes de ella se muestran las respuestas de observación durante el eslabón inicial y posterior a ella las respuestas de observación durante el eslabón terminal. 48

Figura 8. Porcentaje relativo de respuestas de observación en cada decil del intervalo entre reforzadores como promedio de las últimas seis sesiones de cada condición. Para la

condición donde se presentaban ambos estímulos se promediaron las tres condiciones a las que estuvieron expuestos los sujetos RP2 y RP3, y las últimas dos exposiciones a esta condición para el sujeto RP1. La línea punteada indica que antes de ella se muestran las respuestas de observación durante el eslabón inicial y posterior a ella las respuestas de observación durante el eslabón terminal. 49

Resumen

Un reforzador condicionado es un estímulo que adquiere valor reforzante a partir de su asociación con un reforzador primario. Los programas de reforzamiento encadenados han sido útiles para estudiar el reforzamiento condicionado. Éstos consisten en dos o más programas de reforzamiento simples llamados eslabones señalados diferencialmente. En este tipo de programas con dos eslabones, cumplir el requisito del eslabón inicial resulta en el eslabón terminal y cumplir el requisito del eslabón terminal resulta en la entrega del reforzador primario. Para algunos autores, los estímulos adquieren una función dual, como reforzadores condicionados de las respuestas que ocurren en el eslabón precedente y como estímulo discriminativo. Sin embargo, analizar dicha función dual de los estímulos es difícil principalmente porque la misma respuesta que produce el cambio de eslabón, junto con el estímulo asociado, produce también el reforzador primario al final del eslabón terminal. En el presente trabajo, con el objetivo de estudiar la función de los estímulos en cada eslabón en un programa de reforzamiento encadenado, se separaron los efectos reforzantes y discriminativos de los estímulos utilizando dos operandos. Se expuso a tres ratas a un programa de reforzamiento encadenado con dos eslabones; las presiones a una de las palancas producían la transición entre eslabones y el reforzador primario, y las presiones a otra palanca producían un estímulo asociado diferencialmente a cada uno de los eslabones (respuestas de observación). En condiciones sucesivas se eliminó uno u otro estímulo asociado con el eslabón inicial y terminal. Se encontró que solamente el estímulo asociado con el eslabón terminal tiene propiedades reforzantes.

Palabras clave: Programas de reforzamiento encadenados, programas de reforzamiento tándem, reforzamiento condicionado, respuestas de observación

Reforzamiento condicionado en programas de reforzamiento encadenados

Un reforzador condicionado es un estímulo que adquiere valor reforzante dada su asociación con un reforzador primario y subsecuentemente puede servir como un reforzador efectivo por sí mismo (Kelleher & Gollub, 1962; Mazur, 2006).

Algunos de los primeros procedimientos usados para estudiar el fenómeno del reforzamiento condicionado involucraron asociar un estímulo con un reforzador primario y posteriormente presentar el estímulo de manera contingente a alguna respuesta en ausencia del reforzador primario. Con este tipo de procedimientos puede mantenerse una respuesta establecida previamente o incluso puede establecerse una nueva respuesta con el reforzador condicionado (Skinner, 1938). Sin embargo, realizar un análisis sobre los efectos del reforzamiento condicionado con estos procedimientos puede ser problemático porque la eliminación del reforzador primario resulta en la reducción de los efectos del reforzador condicionado, de tal forma que la respuesta se mantiene únicamente durante un período relativamente breve de tiempo. Esta dificultad no necesariamente cambia la utilidad del concepto de reforzamiento condicionado, sino que parece demostrar la competencia de otros principios de condicionamiento que se deben tomar en cuenta, ya que podrían contrarrestar el efecto del reforzador condicionado sobre la respuesta blanco (véase Williams, 1994). Así, es preferible utilizar procedimientos en los que se mantenga continuamente la asociación entre el reforzador primario y el reforzador condicionado (Gollub, 1977). Un ejemplo de este tipo de procedimientos son los programas de reforzamiento encadenados.

Formalmente, los programas de reforzamiento encadenados consisten en dos o más programas de reforzamiento simple, llamados eslabones, que se presentan de manera

secuencial y se señalan diferencialmente. Por ejemplo, en un programa de reforzamiento encadenado intervalo fijo (IF) 10 s IF 10 s, se encuentran vigentes dos programas de IF 10 s en secuencia; el primer eslabón se señala con un tono y la primera respuesta después de 10 s apagará el tono e iniciará el siguiente eslabón, es decir, el segundo programa de IF 10 s, el cual se señala con una luz. La primera respuesta después de 10 s apagará la luz y resultará en la entrega de una bolita de comida. Es importante señalar que la consecuencia inmediata de cumplir el requisito del eslabón inicial es la presentación del estímulo asociado con el eslabón terminal que, a su vez, está asociado con el reforzador primario. Estos programas son análogos a la formación de una cadena de respuestas que termina con la entrega del reforzador primario (Ferster & Skinner, 1957).

En una cadena de respuestas podría interpretarse que cada uno de los estímulos asociados a cada eslabón adquiere una función dual (véase Keller & Schoenfeld, 1950). Por ejemplo, considérese la ejecución de una rata cuyas presiones a una palanca son reforzadas con comida. Este patrón de respuesta relativamente sencillo puede ser descompuesto en una cadena de respuestas que involucra que la rata se mueva hacia la palanca, la oprima y consuma el alimento para después regresar a la palanca para reiniciar la cadena. La secuencia completa se mantiene por la entrega de la comida, pero la aproximación a la palanca podría estar mantenida directamente por un reforzador condicionado: el sonido del interruptor de la palanca (e.g. Skinner, 1934) y éste sonido podría, simultáneamente, tener una función como estímulo discriminativo al señalar que el siguiente elemento de la cadena, como aproximarse al comedero, puede producir el siguiente reforzador. Así, cada uno de los estímulos en la cadena puede tener una función doble, como estímulo discriminativo que señala la ocasión para que se produzca la siguiente conducta, y una

función como reforzador de la conducta precedente, esta hipótesis sobre la formación de cadenas conductuales se conoce como encadenamiento (Catania, 1975; Keller & Schoenfeld, 1950).

Para estudiar la formación de una cadena en condiciones controladas, los programas de reforzamiento encadenado resultan especialmente importantes ya que involucran un cambio de estimulación, una contingencia entre la respuesta y la sucesión de eslabones, así como una contingencia entre la respuesta y el cambio de estimulación. Esto resulta útil para el estudio del reforzamiento condicionado, siguiendo el razonamiento de que la conducta emitida en cada uno de los eslabones del programa está mantenida por las propiedades reforzantes del estímulo asociado con el eslabón subsecuente en la cadena (Royalty, Williams, & Fantino, 1987).

A pesar de la aparente simplicidad de los programas encadenados, demostrar que los estímulos tienen una función como reforzadores condicionados ha sido problemático debido a que cada estímulo dentro de la cadena podría funcionar como reforzador condicionado de las respuestas en el eslabón previo, pero al mismo tiempo, las respuestas en todos los eslabones son necesarias para producir el reforzador primario al final de la cadena. El efecto del reforzador primario, en consecuencia, podría ser el único responsable de los efectos reforzantes atribuidos al cambio de estimulación. Por lo tanto, una estrategia para tratar de separar dichos efectos ha consistido en comparar las respuestas mantenidas con programas de reforzamiento encadenados con programas en los que se omite el cambio de estimulación llamados programas de reforzamiento tándem. Los programas tándem, al igual que los programas encadenados, se componen de dos o más programas de reforzamiento simple que cambian de manera contingente a la conducta, sin embargo, no se

señalan de manera diferencial, sino que el mismo estímulo exteroceptivo permanece constante durante todos los eslabones del programa. De tal manera, las condiciones en cada uno de los eslabones de un programa de reforzamiento encadenado son equivalentes en un programa tándem, lo cual permite realizar un análisis independiente de cada uno de los eslabones y, en principio, aislar el valor reforzante de cada estímulo (Kelleher & Gollub, 1962). En esta comparación, si el cambio de estimulación entre eslabones tiene una función reforzante, la tasa de respuesta en los eslabones iniciales debería ser mayor en un programa de reforzamiento encadenado que en un programa tándem equivalente.

Gollub (1977) describió que usando programas de dos eslabones la tasa de respuesta en el primer eslabón fue mayor en un programa de reforzamiento encadenado que en un programa de reforzamiento tándem. Este hallazgo es congruente con la noción de que el estímulo en el eslabón terminal funciona como reforzador condicionado de las respuestas en el eslabón inicial. Sin embargo, en otros estudios similares la tasa de respuesta en los eslabones iniciales fue más baja en programas encadenados que en programas tándem (Kelleher & Fry, 1962; Wallace, Osborne, & Fantino, 1982), o no se han observado diferencias (Malagodi, DeWeese, & Johnston, 1973). Por ejemplo, Kelleher y Fry (1962) expusieron a dos palomas a un programa encadenado con tres eslabones iguales IF 60 s y de manera alternada a un programa tándem con los mismos eslabones. Encontraron que, en el programa tándem, las respuestas fueron cercanas a cero durante el inicio del primer eslabón y aumentaron consistentemente hasta la entrega del reforzador en el eslabón terminal. En contraste, en el programa de reforzamiento encadenado, el número de respuestas fue distinto en cada eslabón, pero en todos, las respuestas se distribuyeron hacia el final de estos. En el eslabón inicial, las respuestas fueron cercanas a cero, en el eslabón

intermedio fueron infrecuentes y durante el eslabón terminal aumentaron notablemente. Comparativamente, las respuestas fueron más frecuentes en el eslabón inicial e intermedio del programa tándem que en los mismos eslabones en el programa encadenado. Estos hallazgos sugieren que los estímulos del programa encadenado funcionan como estímulos discriminativos que modulan la tasa y la distribución temporal de las respuestas en su presencia. Sin embargo, no mostraron evidencia clara de que los estímulos funcionaran como reforzadores condicionados que aumentaran las respuestas en el eslabón precedente al que se presentaban.

La inconsistencia existente entre la hipótesis del encadenamiento y los patrones de respuesta que se encuentran al comparar programas de reforzamiento tándem con programas de reforzamiento encadenado ha puesto en duda la utilidad del concepto de reforzamiento condicionado. Staddon (1983) y Baum (1973) argumentaron que el uso del concepto de reforzamiento condicionado es innecesario para explicar los patrones de respuesta en los programas de reforzamiento encadenado. Ambos autores han propuesto que la distancia temporal entre la presentación del estímulo y la entrega del reforzador primario, y no un efecto de reforzamiento condicionado, es la variable responsable de los patrones de respuesta. Staddon y Cerutti (2003) llevaron a cabo un análisis teórico sobre las implicaciones del uso del concepto de reforzamiento condicionado en estos programas; los autores propusieron una explicación en términos del control temporal llamada espera lineal, la cual explica las pausas prolongadas en los primeros eslabones de un programa de reforzamiento encadenado. Suponiendo que la función principal de los estímulos en un programa encadenado es simplemente una función discriminativa del tiempo que transcurre hacia la entrega del reforzador primario, en un programa de reforzamiento encadenado de

tres eslabones donde cada eslabón tiene una duración fija de I , el tiempo hacia el reforzador es de $3I$. El cambio de estimulación señala que el tiempo hacia el reforzador es de $2I$ y el siguiente cambio señala una demora de I . El supuesto de la espera lineal, es que el espaciamiento (el tiempo antes de la primera respuesta) en cada eslabón está determinado por el tiempo hacia el reforzador, por lo tanto, la pausa disminuirá conforme avancen los eslabones. Congruentemente con la explicación de Staddon y Cerutti (2003), Baum (1973) explicó que las respuestas emitidas durante cada uno de los eslabones de un programa de reforzamiento encadenado se encuentran bajo el control de la transición entre eslabones y que los estímulos asociados con cada uno de los eslabones sólo tienen una función discriminativa debido a que señalan la disponibilidad demorada del reforzador primario (véase también Davison & Baum, 2006).

A pesar de las críticas que se han realizado en torno a la utilidad del concepto de reforzamiento condicionado, hay estudios que han reportado hallazgos consistentes con este fenómeno y que no pueden ser explicados solamente por la función discriminativa de los estímulos, o la distancia temporal hacia la entrega del reforzador primario. Royalty, Williams y Fantino (1987) realizaron uno de los experimentos que apoya la noción de que los estímulos en los programas de reforzamiento encadenados funcionan como reforzadores condicionados. Royalty, et al. investigaron la función del reforzamiento condicionado al comparar la transición inmediata y demorada entre eslabones en un programa de reforzamiento encadenado, sin alterar la proximidad temporal de cada eslabón con la entrega del reforzador primario. Expusieron a palomas a un programa de reforzamiento encadenado con tres eslabones intervalo variable (IV) 33 s IV 33 s IV 33. Cuando las respuestas se estabilizaron, en diferentes condiciones los autores introdujeron una demora

no señalada de 3 s entre el eslabón inicial y el eslabón intermedio, entre el eslabón intermedio y el terminal y en el eslabón terminal. Para mantener constante la duración nominal de 33 s de cada eslabón, el eslabón al que le precedía la demora se cambió por un programa IV 30 s. De manera global, se encontró que las respuestas del eslabón precedido por la demora disminuyeron con respecto a la línea base y las respuestas en los otros eslabones no tuvieron cambios. Esto sugiere que el estímulo asociado con cada eslabón tiene propiedades de reforzador condicionado. Sin embargo, Staddon y Cerutti (2003) argumentaron que en este trabajo la duración obtenida de cada eslabón, pudo haber aumentado debido a la disminución en la tasa de reforzamiento y pudo haber determinado a su vez la tasa de respuesta. En este sentido los hallazgos de Royalty et al., son congruentes con la hipótesis de la espera lineal. De acuerdo con Williams (1994) otras críticas están relacionadas con el hecho que si el estímulo tuviera una función sensorial o solamente sirviera como un marcador de la respuesta, la demora reduciría también el efecto del estímulo, pero no debido a su valor como reforzador condicionado.

En otro estudio, Bejarano y Hackenberg (2007), expusieron a palomas a un programa de reforzamiento encadenado de dos eslabones. Durante el eslabón inicial se encontraba vigente un programa en el cual la probabilidad de avanzar hacia el siguiente eslabón dependía de la duración de los tiempos entre respuestas; en el eslabón terminal estaba vigente un programa intervalo al azar (IA) 30 s. Los autores hipotetizaron que, si el cambio de estimulación durante la transición de los eslabones era reforzante, entonces los sujetos responderían con tiempos entre respuestas más largos cuando la probabilidad de transición entre eslabones fuera mayor, y en tanto la demora a la entrega del reforzador primario se mantuviera constante. Los autores reportaron que los tiempos entre respuestas

eran más largos cuando la probabilidad de transición al eslabón terminal variaba directamente con el tiempo entre respuestas inmediato anterior. En un segundo experimento, Bejarano y Hackenberg (2007) extendieron el programa de reforzamiento de dos eslabones a cuatro y en cada eslabón se mantuvo vigente el mismo programa de reforzamiento probabilístico. En este segundo experimento se reportó que los tiempos entre respuestas del tercer eslabón fueron mayores que los del cuarto, pero menores que los del segundo. Los resultados reportados por Bejarano y Hackenberg (2007) pueden ser explicados debido al efecto reforzante del cambio de estimulación y no pueden ser atribuidos a la distancia temporal a la entrega del reforzador primario que señalaban los estímulos, debido a que se mantuvo constante la duración promedio de cada uno de los eslabones. No obstante, los resultados del Experimento 2 deben ser objeto de análisis debido a que no son completamente congruentes con la hipótesis del encadenamiento; de ser así, se hubiera encontrado que los tiempos entre respuestas del cuarto eslabón eran más largos que los del tercer eslabón y estos a su vez, más largos que los del eslabón inicial. Esto sugiere que los estímulos asociados con cada uno de los eslabones pueden adquirir propiedades de reforzador condicionado, aunque también podrían tener otros efectos que deben ser analizados.

Llevar a cabo el análisis de la función de los estímulos asociados con cada uno de los eslabones de un programa de reforzamiento encadenado resulta complicado debido a que la misma respuesta que produce el cambio de estimulación, también produce la transición entre eslabones y el reforzador primario al final de la cadena. Una manera sencilla de aislar estas consecuencias sería disociando las respuestas que resultan en el reforzador primario de las respuestas cuya única consecuencia es la presentación de los

estímulos. Un procedimiento útil para disociar dichos efectos, es el procedimiento de respuestas de observación (Wyckoff, 1969). Una respuesta de observación es una respuesta distinta a la que produce el reforzador primario y que no afecta la disponibilidad ni la frecuencia de reforzamiento. Las respuestas de observación producen estímulos previamente correlacionados con algún período de reforzamiento o extinción (Wyckoff, 1952, 1969).

Wyckoff (1952) describió inicialmente el procedimiento de respuestas de observación mientras estaba investigando cómo se llevaba a cabo una discriminación y reconoció la importancia de que el organismo “atendiera” al estímulo, es decir, estaba interesado en la conducta de orientación hacía el estímulo relevante. Para ello, entrenó a un grupo de palomas en un programa de reforzamiento con dos períodos, uno de reforzamiento, IF 30 s, y otro de extinción 30 s, alternados al azar. Cada uno de los períodos se encontraba señalado diferencialmente, IF con la iluminación de una tecla roja (E+) y extinción con la iluminación de una tecla verde (E-). Posteriormente, Wyckoff dividió a las palomas en dos grupos, eliminó los estímulos asociados con cada período e introdujo un pedal en la parte inferior del panel frontal de la caja. Para uno de los grupos las presiones al pedal producían los estímulos previamente asociados con cada uno de los períodos y para el otro grupo las presiones en el pedal producían los estímulos de manera aleatoria. Wyckoff encontró que cuando los estímulos estaban correlacionados positivamente con el programa de reforzamiento vigente, se mantuvo una frecuencia de respuestas de observación estable. Mientras que, cuando los estímulos se encontraban correlacionados negativamente los sujetos no emitían respuestas de observación (i.e., no presionaban el pedal).

El procedimiento de respuestas de observación se emplea cuando se requiere analizar la adquisición y el mantenimiento de las respuestas que producen estímulos discriminativos asociados al programa de reforzamiento vigente sin alterarlo, de ahí su denominación como “respuestas de observación”. Las respuestas de observación no están asociadas con el incremento en la frecuencia del reforzador primario, es por ello que han servido para justificar el concepto de reforzamiento condicionado y también han servido como una medida para cuantificar el efecto reforzante de un estímulo (Dinsmoor, 1983).

El uso del procedimiento de observación podría permitir estudiar el valor reforzante de los estímulos asociados a los eslabones de los programas de reforzamiento encadenados (Hendry & Dillow, 1966). De manera general, con este procedimiento vigente en un programa de reforzamiento tándem, las respuestas de observación podrían transformar el programa tándem en un programa encadenado por un período corto de tiempo.

En un estudio realizado por Hendry y Dillow (1966) se utilizó un procedimiento de respuestas de observación añadido a un programa de reforzamiento tándem. Hendry y Dillow expusieron a palomas a un programa de reforzamiento encadenado IF 60 s IF 60 s IF 60 s y posteriormente se cambió a un programa tándem equivalente con una respuesta de observación que encendía los estímulos y transformaba el programa momentáneamente en un encadenado. Los autores reportaron que los estímulos en los tres eslabones mantuvieron las respuestas de observación, sin embargo, se observaron diferentes patrones de respuesta durante cada uno de los eslabones. La tasa de respuestas de observación fue mayor en el eslabón intermedio comparado con el eslabón terminal, y a su vez mayor que durante el eslabón inicial. Con estos datos, Hendry y Dillow argumentaron que los estímulos del eslabón terminal e intermedio podrían haber adquirido propiedades de reforzador

condicionado, pero que el estímulo del eslabón inicial debería de estudiarse con mayor cuidado porque parecía mantener una tasa de respuesta cercanas a cero en la tecla de reforzamiento, lo cual no es congruente con la interpretación de que este estímulo funcione como un reforzador condicionado. El hecho de que el estímulo del eslabón terminal haya mantenido una menor tasa de respuesta en comparación con el eslabón intermedio podría deberse a que en el eslabón terminal las respuestas de observación compitieron con las respuestas por el reforzador primario con un patrón típico de respuestas en un programa de IF (e.g., Slezak & Anderson, 2014), lo cual dificulta llevar a cabo un análisis de la función de los estímulos como reforzadores condicionados.

Además de calcular la tasa de respuesta de observación, otra forma de determinar la función que adquiere cada estímulo consiste en eliminar el estímulo del procedimiento y observar tanto la tasa de respuesta de observación como los patrones de respuesta en presencia y ausencia de los estímulos (e.g., Kendall, 1972; Lieberman, 1972). Por ejemplo, si eliminar el estímulo produce una disminución en las respuestas de observación, podría sugerirse que el estímulo tenía una función reforzante.

Propósito

Los programas de reforzamiento encadenado han sido útiles para el estudio del reforzamiento condicionado debido a que el cambio de estimulación que acompaña la transición entre eslabones, puede funcionar como reforzador condicionado de las respuestas en el eslabón precedente. No obstante, determinar la función de los estímulos como reforzadores condicionados ha sido problemática debido a que la misma respuesta que produce el cambio de eslabones (y con ello el cambio de estimulación), también produce el

reforzador primario al final de la cadena. Una manera de analizar la función de los estímulos, es comparar los programas de reforzamiento encadenados con los programas de reforzamiento tándem, a los cuales son equivalentes, pero se omite el cambio de estimulación entre eslabones. En esta comparación, se esperaría que, si los estímulos adquieren una función de reforzador condicionado, la tasa de respuesta en los eslabones iniciales de un programa de reforzamiento encadenado debería de ser mayor que en un programa equivalente tándem. No obstante, los estudios que han llevado a cabo estas comparaciones no han sido congruentes con esta hipótesis (e.g., Kelleher & Fry, 1962; Malagodi, et al., 1973; Wallace, et al., 1982).

Derivado de estas inconsistencias, algunos autores han puesto en duda la utilidad del concepto de reforzamiento condicionado y han explicado la función de los estímulos que acompañan a los eslabones de un programa de reforzamiento encadenado en términos de una función discriminativa que modula la tasa de respuesta en cada uno de los eslabones señalando la distancia temporal hacia la entrega del reforzador primario como la variable controladora de éste efecto (veáse Baum, 1973; Davison & Baum, 2006; Staddon, 1983; Staddon & Cerutti, 2003). Sin embargo, existen otros estudios que han demostrado que la explicación de ciertos patrones de respuesta en programas de reforzamiento encadenado puede sustentarse mediante una explicación de reforzamiento condicionado y no en términos de control discriminativo (e.g. Bejarano & Hackenberg, 2007; Royalty, et al., 1987).

La evaluación directa de las variables involucradas durante el establecimiento de un reforzador condicionado en un programa de reforzamiento encadenado resulta problemática, porque la misma respuesta que produce el cambio entre eslabones y con ello

el reforzador condicionado, también produce el reforzador primario al final de la cadena. Para disociar estas respuestas, se puede utilizar el procedimiento de respuestas de observación. Una respuesta de observación produce específicamente el reforzador condicionado sin alterar la frecuencia de reforzamiento del reforzador primario. Los datos disponibles con un procedimiento en el que se disociaron las respuestas que producen los estímulos de un programa encadenado de las repuestas que producen el cambio de eslabón y el reforzador primario, sugieren que sólo algunos de los estímulos podrían tener una función reforzante y que este valor reforzante disminuye conforme estos están más alejados del reforzador primario (e.g., Hendry & Dillow, 1966). Los datos de Hendry y Dillow (1966) son congruentes con los datos de Royalty et al. (1987), quienes mostraron que los estímulos de un programa de reforzamiento encadenado pueden funcionar como reforzadores condicionados debido a que, al añadir una demora entre la respuesta que produce el cambio de estímulo, disminuye la tasa de respuesta en el eslabón que precede la demora.

El objetivo del presente trabajo fue determinar si los estímulos del eslabón inicial y terminal de un programa de reforzamiento encadenado pueden funcionar como reforzadores condicionados que establecen y mantienen respuestas de observación. Además de analizar la función de cada uno de los estímulos al eliminar la presentación del estímulo del eslabón inicial y del eslabón terminal en condiciones sucesivas (e.g., Kendall, 1972). Se esperaría que si un estímulo tuviera una función reforzante, su eliminación produciría un decremento en las respuestas de observación que lo producen.

Método

Sujetos

Se utilizaron tres ratas Wistar macho de aproximadamente ocho meses de edad al comienzo del experimento y con experiencia en programas de reforzamiento simples. Las ratas se encontraban alojadas en cajas habitación individuales en la colonia del Laboratorio de Condicionamiento Operante de la Facultad de Psicología de la UNAM.

Los tres sujetos tenían libre acceso a comida y se mantuvieron en un ciclo de privación de agua en el cual las ratas tenían acceso a agua durante una hora después de cada sesión experimental y posteriormente se restringía su consumo hasta que la sesión experimental del siguiente día tuviera inicio.

Aparatos

Se utilizó una cámara de condicionamiento operante construida por medio de impresión en 3D y corte láser como se describe en Escobar (2016). La cámara medía 30 cm de ancho, 23 cm de alto y 20.5 cm de profundidad. En panel frontal se encontraba un recipiente para agua; dos palancas retractiles con un interruptor (Honeywell, 311SM4-t) y un servomotor con engranes de metal, que servía como mecanismo para expulsar y retraer las palancas fabricadas de lámina de acero inoxidable calibre 16, que se colocaron a cada lado del recipiente para agua. Era necesaria una fuerza de 0.15 N para activar el interruptor de cada palanca. Sobre cada una de las palancas se colocó un LED que emitía una luz blanca con una cubierta traslúcida. Detrás del panel frontal había un sonalert que emitía un tono de 80 dB a 2900 Hz; una bomba peristáltica que se activaba para dispensar 0.1 ml de agua en el recipiente del panel frontal y un motor de vibración que se activaba de manera

simultánea y durante el tiempo que la bomba permanecía activa. En el panel posterior se encontraba un LED que proporcionaba la iluminación general de la cámara experimental.

La cámara experimental se colocó en un cubículo de madera sonoamortiguador y equipado con un ventilador Steren® (Modelo VN4-012P) para facilitar la circulación de aire dentro de la cámara, así como dos bocinas conectadas a una computadora que servían para generar ruido blanco de 70 dB, almacenado en un archivo mp3, para enmascarar sonidos externos. Para controlar y registrar los eventos experimentales se usó la interfaz Arduino-Visual Basic descrita por Escobar y Pérez-Herrera (2015; Escobar, 2016).

Procedimiento

Las sesiones experimentales se llevaron a cabo diariamente, los siete días de la semana a la misma hora. Debido a la experiencia de los sujetos en programas de reforzamiento de IV, no fue necesario entrenar las presiones a la palanca y se les expuso de manera inmediata a las condiciones que a continuación se describen.

Fase de Entrenamiento. En esta fase se entrenó a los sujetos a presionar la palanca conforme a un programa de reforzamiento encadenado y posteriormente a un programa de reforzamiento tándem equivalente. A continuación, se describe cada uno de estos programas.

Programa de reforzamiento encadenado. En el panel frontal se encontraba disponible sólo la palanca izquierda (de ahora en adelante, palanca de reforzamiento), mientras que la palanca derecha (de ahora en adelante, palanca de observación) se encontraba retraída; la luz general iluminaba la cámara experimental desde el inicio hasta el término de la sesión.

Se expuso a los sujetos a un programa encadenado de dos eslabones con el mismo programa de reforzamiento. El eslabón inicial se señaló con la iluminación de la luz izquierda sobre la palanca de reforzamiento y el eslabón terminal se señaló por un estímulo auditivo constante generado por el sonalert. El reforzador consistió en la entrega de 0.1 ml de agua.

Todos los sujetos se expusieron a sesiones de entrenamiento en donde estuvo vigente un programa de reforzamiento encadenado IV 5 s IV 5 s, una vez que los sujetos obtuvieron 40 reforzadores en una sesión el valor de los IV se duplicó hasta llegar a un programa encadenado IV 20 s IV 20 s. Los intervalos en cada uno de los eslabones fueron programados de acuerdo a la progresión de Fleshler y Hoffman (1962) con 10 iteraciones. Cada sesión finalizó cuando se entregaron 40 reforzadores o cuando hubiera transcurrido el doble del tiempo promedio programado del programa encadenado; por ejemplo, en el caso del programa encadenado IV 20 s IV 20 s, el promedio del tiempo programado para obtener los 40 reforzadores es de 1600 s, por lo que el tiempo límite de la sesión eran 3200 s o 53.3 min.

Una vez que los sujetos estuvieron expuestos al programa encadenado IV 20 s IV 20 s, éste estuvo vigente por al menos 12 sesiones y posteriormente se determinó si las respuestas para ambos eslabones alcanzaron el criterio de estabilidad descrito por Schoenfeld, Cumming y Hearst (1956). Este criterio se cumplió cuando al considerar las últimas seis sesiones de la condición vigente (por ejemplo, sesión 6 a 12), la diferencia entre el promedio de las primeras tres (sesión 6 a 9) y el promedio de las últimas tres (sesión 10 a 12) no era mayor al 5% del promedio total de las seis; en caso contrario, la condición se prolongó un día más y se reevaluó cada día hasta que el criterio se satisficiera

o hasta que se cumplieran 24 sesiones. Para determinar la estabilidad de la tasa de respuesta se utilizó el programa *Stability check* (Costa & Cançado, 2012).

Programa de reforzamiento tándem. En esta condición se mantuvieron constantes las condiciones de la cámara experimental con respecto a la condición anterior. Se expuso a los sujetos a un programa tándem de dos eslabones con el mismo programa de reforzamiento vigente en la condición anterior, IV 20 s IV 20 s, en el cual no se señaló diferencialmente el eslabón inicial y el eslabón terminal. Cada sesión finalizó cuando se entregaron 40 reforzadores o después de 53.3 min. La estabilidad se determinó bajo los mismos criterios que en la fase anterior.

Fase experimental. Se llevó a cabo un diseño experimental reversible ABACA, el cual se describe a continuación. Durante cada una de las condiciones la luz general iluminó la cámara experimental durante toda la sesión y en el panel frontal de la cámara experimental estaban expuestas ambas palancas. Cada sesión terminó cuando fueron entregados 40 reforzadores o después de 53.3 min transcurridos, al igual que en la fase de entrenamiento; cada condición estuvo vigente por al menos 12 sesiones y posteriormente se determinó la estabilidad del número de respuestas en la palanca de reforzamiento y en la palanca de observación, tanto en el eslabón inicial como en el eslabón terminal, de la misma manera que en la fase de entrenamiento. La condición cambió cuando se cumplió el criterio de estabilidad o después de 24 sesiones. En la Tabla 1, se muestra el número de sesiones por condición a las que estuvieron expuestos los tres sujetos.

Condición A: Programa tándem con respuestas de observación. En esta condición, se mantuvo vigente el programa de reforzamiento tándem IV 20 s IV 20 s de la fase de entrenamiento. Las presiones a la palanca de reforzamiento producían el cambio entre los

eslabones, así como 0.1 ml de agua al cumplir el criterio del programa de reforzamiento en el eslabón terminal del programa tándem. Las presiones a la palanca de observación encendían durante 5 s los estímulos asociados con cada eslabón en el programa encadenado de la fase de entrenamiento, así, cada presión en la palanca de observación transformó el programa tándem en un programa encadenado por 5 s. Mientras los estímulos se encontraban encendidos, las presiones en la palanca de observación no tenían consecuencias programadas y si ocurría el cambio de eslabón o la entrega de reforzador durante este período, los estímulos se apagaban hasta que otra respuesta de observación los encendiera de nuevo.

Las respuestas de observación no fueron entrenadas de manera independiente, sino que solo se expuso la palanca de observación y al presionarla los sujetos tuvieron contacto con la presentación del estímulo correspondiente a cada eslabón durante 5 s. De tal manera que, al concluir esta condición, la conducta de presionar la palanca de observación se estableció solo para dos sujetos, RP1 y RP3. Para el sujeto RP2, el número de respuestas de observación por sesión eran cercanas a cero, de tal manera que este sujeto fue expuesto nuevamente a la condición A, pero las consecuencias de ambas palancas fueron invertidas, así, la palanca izquierda era entonces la palanca de observación y la palanca derecha la palanca de reforzamiento. Esta condición se mantuvo vigente durante 10 sesiones y posteriormente se restableció la configuración original de las palancas y se volvió a exponer a RP2 a la condición A con las mismas condiciones previamente descritas. Una vez que se observó que RP2 respondía confiablemente en la palanca de observación, la condición se mantuvo vigente por al menos 12 sesiones, como se describió previamente en la sección *Fase experimental*.

Condición B: Eliminación del estímulo en el eslabón inicial. En esta condición se mantuvo vigente el programa de reforzamiento tándem IV 20 s IV 20 s de las condiciones anteriores. Las presiones en la palanca de reforzamiento producían el cambio entre los eslabones, así como el reforzador primario al cumplir el criterio del programa de reforzamiento en el eslabón terminal del programa tándem.

Las presiones en la palanca de observación durante el eslabón inicial no tenían consecuencias programadas y no se producía el cambio de estimulación como en la Condición A. Sin embargo, las respuestas en la palanca de observación durante el eslabón terminal transformaban el programa de reforzamiento tándem en uno encadenado durante 5 s, de acuerdo al estímulo asociado con este eslabón en el programa de reforzamiento encadenado previamente asociado (tono).

Condición C: Eliminación del estímulo en el eslabón terminal. En esta condición, estuvo vigente el programa de reforzamiento tándem IV 20 s IV 20 s de las condiciones anteriores. Las presiones en la palanca de reforzamiento producían el cambio entre los eslabones, además de la entrega del reforzador primario al cumplir el criterio del programa de reforzamiento en el eslabón terminal del programa tándem.

Las presiones en la palanca de observación durante el eslabón inicial transformaban el programa de reforzamiento tándem por uno encadenado durante 5 s, de acuerdo al estímulo asociado con este eslabón en el programa de reforzamiento encadenado en la fase de Entrenamiento (luz). Sin embargo, las respuestas en la palanca de observación durante el eslabón terminal no tenían consecuencias programadas.

Tabla 1.

Número de sesiones a las que estuvieron expuestos cada sujeto por condición en la fase experimental

Sujeto	Condiciones				
	Ambos Estímulos	Sin estímulo del eslabón inicial	Ambos estímulos	Sin estímulo del eslabón terminal	Ambos estímulos
RP1	24	24	24	24	24
RP2	24	24	17	24	23
RP3	24	24	24	24	24

Nota: El número de sesiones por condición que se presentan para el sujeto RP2, son después de haber invertido las consecuencias de las palancas (ver texto).

Resultados

En todas las sesiones de cada una de las condiciones, los tres sujetos obtuvieron 40 reforzadores, el máximo de reforzadores programados. Debido al tipo de programa empleado en cada uno de los eslabones del programa de reforzamiento encadenado, se analizó la duración promedio de las últimas seis sesiones de cada condición por sujeto, y se calculó su respectiva desviación estándar, estos datos se muestran en la Tabla 2. Durante las condiciones, en ninguno de los sujetos se observaron diferencias sistemáticas en cuanto a la duración promedio de la sesión entre condiciones.

Tabla 2.

Duración promedio de la sesión en cada una de las condiciones a las que estuvieron expuestos cada sujeto

Sujeto	Condiciones									
	Ambos Estímulos		Sin estímulo del eslabón inicial		Ambos Estímulos		Sin estímulo del eslabón terminal		Ambos Estímulos	
	Duración	DE	Duración	DE	Duración	DE	Duración	DE	Duración	DE
RP1	37.45	3.78	32.09	0.78	31.33	0.25	33.53	1.45	33.44	0.80
RP2	33.72	1.43	34.29	1.35	35.61	1.29	40.00	4.82	39.71	4.88
RP3	32.21	1.24	30.64	6.54	36.45	2.94	34.97	1.67	39.86	2.48

Nota. La duración promedio de la sesión, así como la desviación estándar, se encuentran expresado en minutos.

DE = Desviación estándar

La Figura 1 muestra las respuestas por minuto en la palanca de reforzamiento para cada uno de los sujetos como la media de las últimas seis sesiones de cada condición. Los círculos blancos representan la tasa de respuesta mantenida con el reforzador primario durante el eslabón inicial mientras que, los círculos negros representan la tasa de respuesta mantenida por el reforzador primario durante el eslabón terminal. Esta variable se obtuvo al dividir el número de respuestas en la palanca de reforzamiento durante cada uno de los eslabones, entre el tiempo en minutos, que se encontró vigente el respectivo eslabón durante la sesión.

Como muestra la Figura 1, la tasa de respuesta en la palanca de reforzamiento durante el eslabón inicial fue menor que la tasa de respuesta durante el eslabón terminal, durante todas las condiciones en los sujetos RP2 y RP3. En el caso del sujeto RP1 esta diferencia se hizo evidente a partir de la segunda condición y permaneció con la misma tendencia el resto del experimento. No se observaron efectos sistemáticos al eliminar los estímulos en ninguna de las ratas.

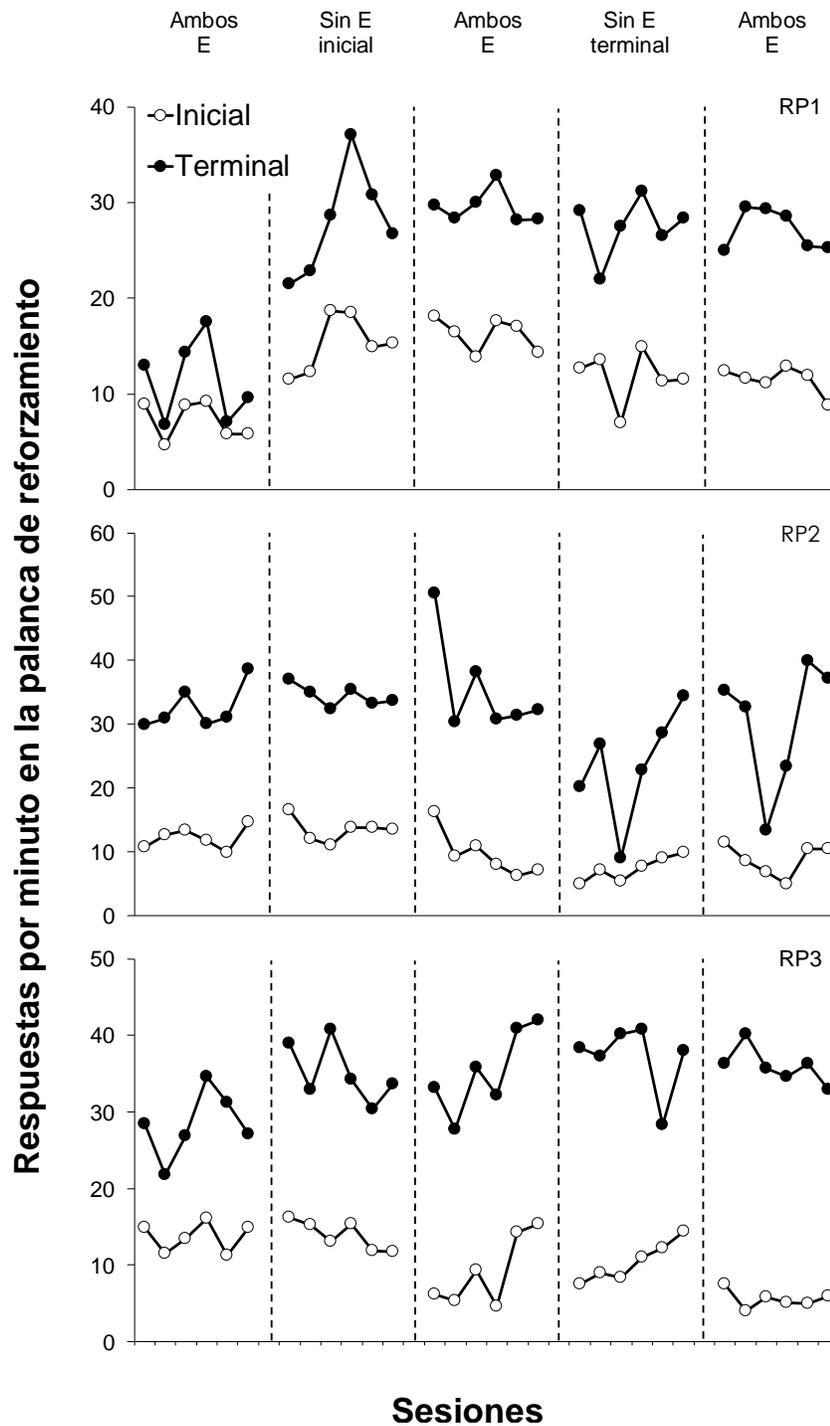


Figura 1. Respuestas por minuto en la palanca de reforzamiento durante las últimas seis sesiones de cada condición. Para cada uno de los sujetos la escala del eje de la ordenada es distinta.

La Figura 2 muestra las respuestas por minuto en la palanca de observación para cada uno de los sujetos durante las últimas seis sesiones de cada condición. La tasa de respuesta de observación se obtuvo al dividir el número de respuestas en la palanca de observación durante cada eslabón, entre la duración corregida en minutos, en que estuvo vigente el eslabón. La corrección de la duración de cada eslabón se llevó a cabo debido a que las respuestas de observación que se producían durante cada uno de los eslabones encendían los estímulos asociados con cada eslabón durante 5 s, a menos que hubiera un cambio de eslabón o la entrega de reforzador primario, y durante este período las respuestas en la palanca de observación no tuvieron consecuencias programadas. Así, se sustrajo el tiempo que permanecieron activados los estímulos a la duración total de los eslabones y con este dato se obtuvo la tasa de respuesta en la palanca de observación. En el caso de las condiciones en las que se eliminó la presentación de alguno de los estímulos y por motivos de comparación, la corrección se llevó a cabo con el tiempo que hubiera estado activado el estímulo, aunque este no haya sido presentado (véase Lieberman, 1972).

Como se puede observar en la Figura 2, en las condiciones donde estuvieron presentes ambos estímulos, la tasa de respuesta de observación durante el eslabón terminal fue mayor a la tasa de respuesta de observación durante el eslabón inicial, para al menos dos sujetos, RP2 y RP3; mientras que, para RP1, las respuestas por minuto en la palanca de observación fueron similares en ambos eslabones. En la condición en la que fue eliminado el estímulo del eslabón inicial, para los sujetos RP1 y RP3, la tasa de respuesta de observación aumentó en los dos eslabones respecto a la condición previa y posterior, donde estaban presentes ambos estímulos; en el caso del sujeto RP2, las respuestas de observación en esta condición parecen no tener diferencia con respecto a la condición donde se

encontraban ambos estímulos. Por otro lado, en la condición en la que se eliminó el estímulo del eslabón terminal, en los tres sujetos se observó una disminución de la tasa de respuesta de observación tanto en el eslabón inicial como en el eslabón terminal, con respecto a las condiciones previa y sucesiva, la tasa de respuesta de observación fue cercana a cero en ambos eslabones.

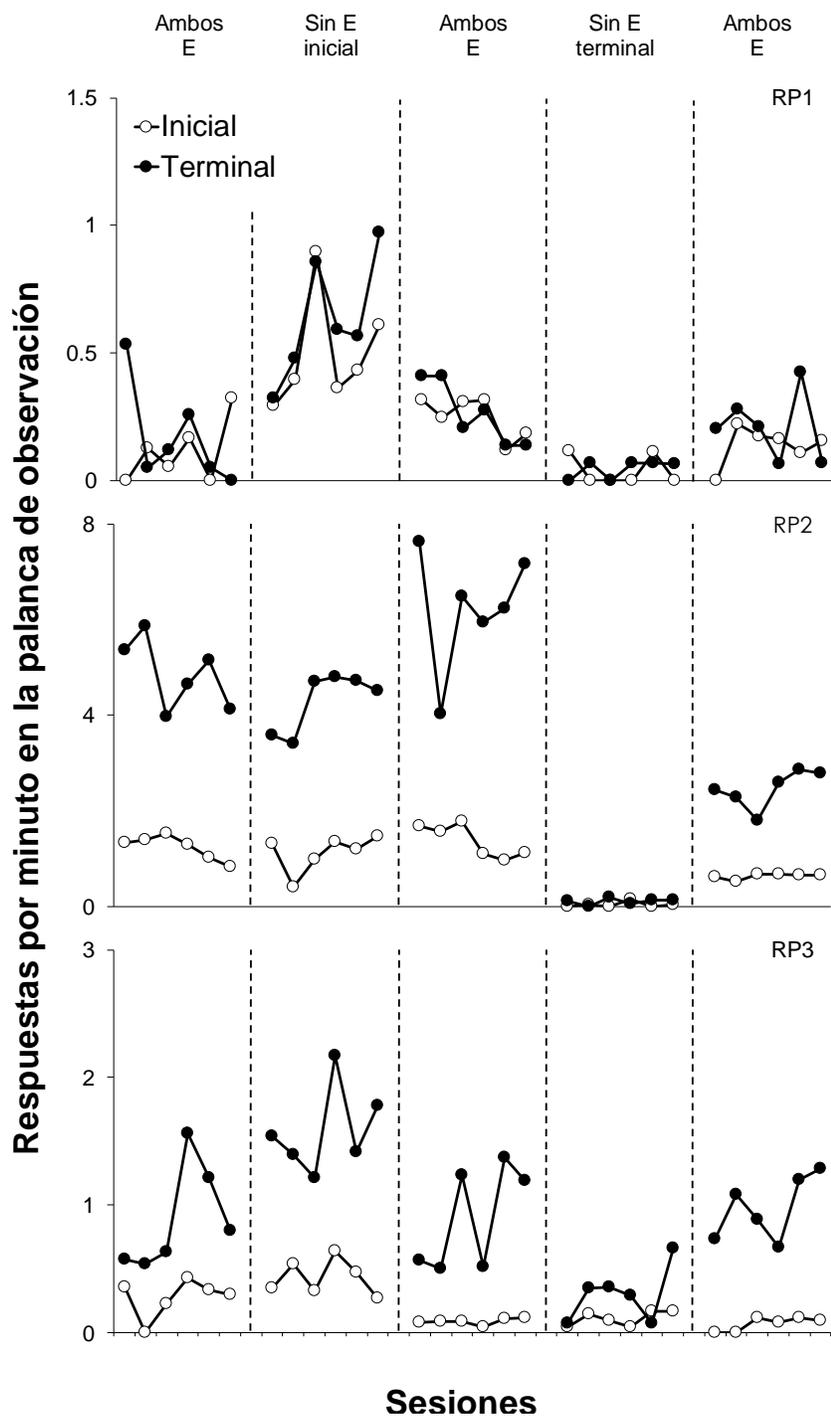


Figura 2. Respuestas por minuto en la palanca de observación durante las últimas seis sesiones de cada condición. Para cada uno de los sujetos la escala es distinta.

En la Figura 3 se muestra la media del número de respuestas en la palanca de reforzamiento durante los estímulos en las últimas seis sesiones de cada condición. Esta variable se obtuvo al dividir el número de respuestas en la palanca de reforzamiento mientras el estímulo permanecía encendido y posteriormente se dividió entre el número de veces que dicho estímulo se presentó durante la sesión. Los círculos blancos representan las respuestas en la palanca de reforzamiento en cada estímulo del eslabón inicial y los círculos negros representan las respuestas en la palanca de reforzamiento en cada estímulo del eslabón terminal que se presentaron durante la sesión. Así, en las condiciones en las que se eliminó la presentación de algún estímulo, solo se presenta la serie de datos en el eslabón en el que estuvo vigente la presentación del estímulo correspondiente.

Globalmente puede observarse en la Figura 3 que, para los tres sujetos, cuando el estímulo del eslabón terminal se encontraba encendido el número de respuestas en la palanca de reforzamiento fue mayor respecto a cuando se encontraba encendido el estímulo del eslabón inicial. Además, en las condiciones en las cuales se eliminó alguno de los estímulos, no varió el número de respuestas relativo a las condiciones donde se presentaban ambos estímulos.

Específicamente, los sujetos RP2 y RP3, a lo largo de las condiciones, emitieron entre 2 y 4 respuestas en promedio en presencia del estímulo del eslabón terminal; mientras que, en presencia del estímulo del eslabón inicial, el promedio de respuestas fue comparativamente más bajo.

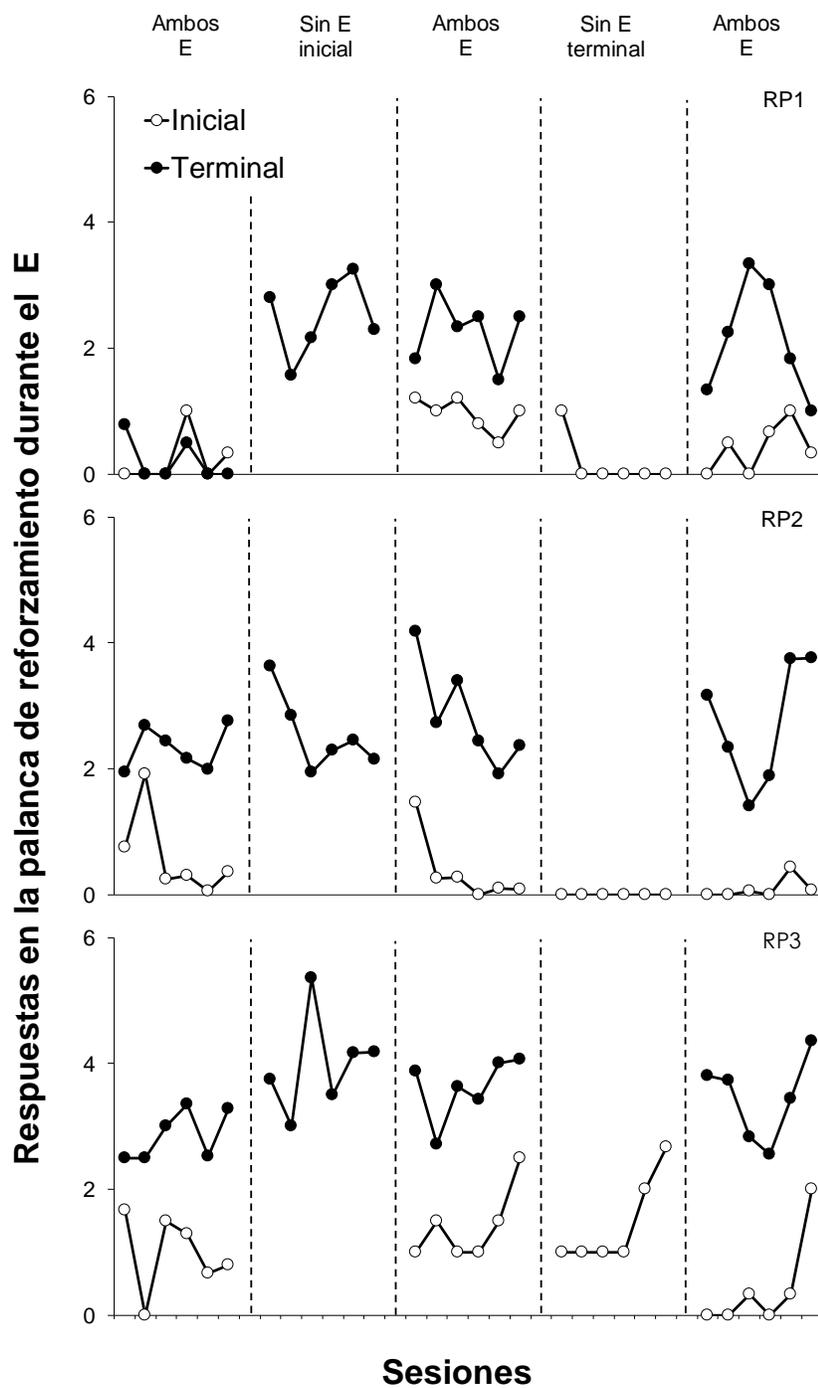


Figura 3. Promedio de respuestas en la palanca de reforzamiento en presencia de cada estímulo en cada uno de los eslabones del programa. Se muestran las últimas seis sesiones de cada condición.

En las Figuras 4 y 5 muestran el promedio de la duración de la pausa post reforzamiento a la palanca de reforzamiento y a la palanca de observación, respectivamente, para los tres sujetos en las últimas seis sesiones de cada condición. Para la duración de la pausa post reforzamiento en la palanca de reforzamiento, se calculó el tiempo que transcurría entre que se entregaba el reforzador y se emitía una respuesta en la palanca de reforzamiento dividido entre 40, el número de intervalos entre reforzadores por sesión. En el caso de la pausa post reforzamiento en la palanca de observación, la duración de la pausa se dividió entre el número de intervalos entre reforzadores en los que se emitieron respuestas de observación; de tal manera que, aquellos intervalos entre reforzadores en los que no se produjo ninguna respuesta de observación se excluyeron de este análisis.

En la Figura 4 se muestra que para ninguno de los sujetos hubo un efecto sistemático entre condiciones en la duración de la pausa post reforzamiento en la palanca de reforzamiento. Para los sujetos RP2 y RP3, se observa una tendencia creciente en la pausa post reforzamiento en las últimas tres condiciones, sin embargo, no parece ser un efecto de las manipulaciones realizadas.

En la Figura 5 se puede observar que en las condiciones en las que estuvieron presentes ambos estímulos, la pausa post reforzamiento a la palanca de observación se mantuvo relativamente estable para los tres sujetos y fue muy parecida a la condición en la que se eliminó el estímulo del eslabón inicial. Sin embargo, en la condición en la que se eliminó el estímulo del eslabón terminal, con respecto a las condiciones previa y posterior, la pausa post reforzamiento aumentó para el sujeto RP2 y aumentó en variabilidad para el sujeto RP1. La discontinuidad en la serie de datos para el sujeto RP1 en la condición sin el

estímulo del eslabón terminal se debe a que en esa sesión el sujeto no realizó alguna respuesta de observación, por lo tanto, no hubo pausa post reforzamiento.

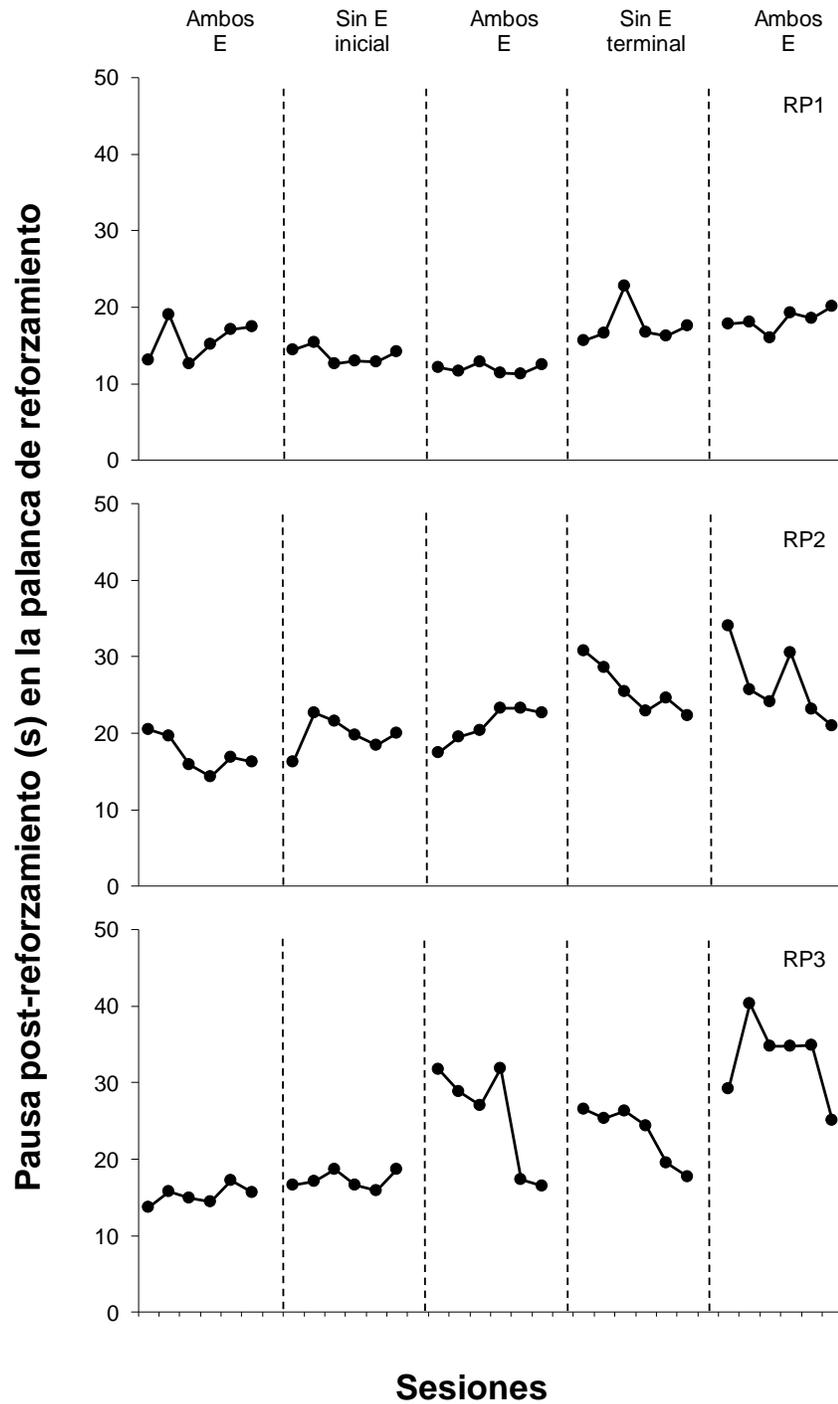


Figura 4. Pausa post reforzamiento en segundos a la palanca de reforzamiento en las últimas seis sesiones de cada condición.

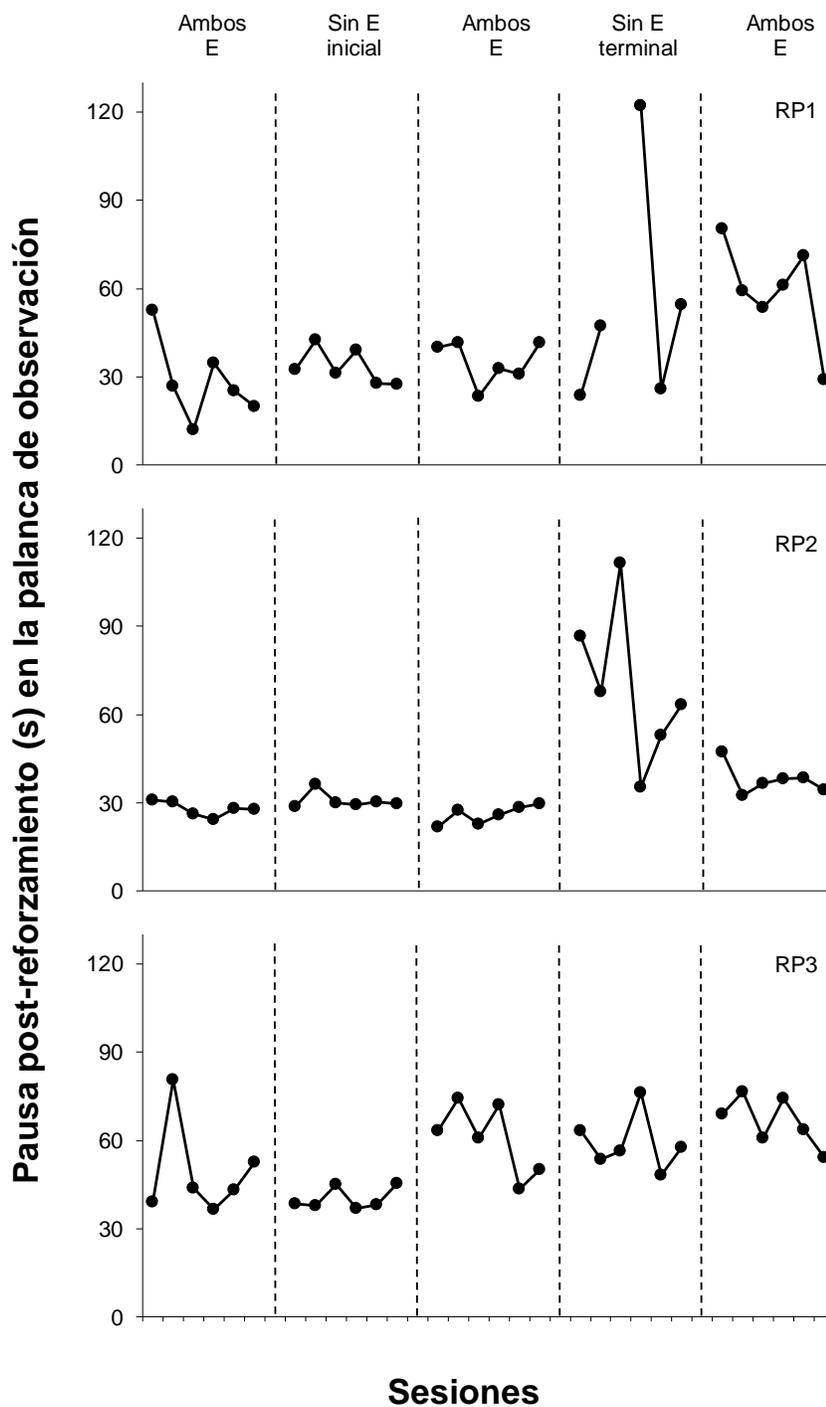


Figura 5. Pausa post reforzamiento en segundos a la palanca de observación en las últimas seis sesiones de cada condición. La discontinuidad en la serie de datos del sujeto RP1 para la condición sin E terminal, se debe a que en esa sesión no se emitió alguna respuesta de observación.

Para llevar a cabo un análisis más detallado de la función de los estímulos producidos con las respuestas de observación se realizó un análisis del tiempo que transcurría entre el encendido de un estímulo y una respuesta de observación subsecuente. Gaynor y Shull (2002) argumentaron que la función de los estímulos como reforzantes o aversivos podría estar relacionada con el tiempo que transcurre entre que se apagaba un estímulo y la ocurrencia de una respuesta de observación posterior o latencia que los volviera a encender. Describieron que un estímulo cuya función sea reforzante, debería controlar la ocurrencia de respuestas de observación inmediatamente después de que se apaga el estímulo. En contraste, un estímulo aversivo debería controlar que no ocurran respuestas de observación inmediatamente después del apagado de un estímulo. Para ello, se analizó por separado las respuestas de observación emitidas después del estímulo del eslabón inicial y después del estímulo del eslabón terminal solamente en las condiciones en las que estuvieron presentes ambos estímulos. Las latencias obtenidas se clasificaron en 6 sub intervalos de 4 segundos cada uno: de 0 a < 4, 4 a < 8, 8 a < 12, 12 a < 16, 16 a < 20 y >20 s, y posteriormente se obtuvo la proporción relativa de las latencias en el eslabón inicial con respecto al eslabón terminal y viceversa.

En cada uno de los eslabones, la proporción relativa de cada latencia se obtuvo al calcular la frecuencia de respuestas emitidas con cada una de las latencias de los sub intervalos, dividido entre el número total de estímulos que se presentaron en ambos eslabones por sesión. Así, si se emitía una respuesta de observación en el eslabón inicial, se calculaba la latencia entre que se apagaba el estímulo del eslabón inicial y la siguiente respuesta de observación y se clasificaba en cada uno de los subintervalos para el eslabón inicial; se llevó a cabo el mismo procedimiento para las latencias de las respuestas de

observación en el eslabón terminal y posteriormente la frecuencia de cada latencia se dividió entre el número total de estímulos presentados por sesión.

En la Figura 6 se muestran las proporciones relativas para cada uno de los sub intervalos del eslabón terminal con respecto al eslabón inicial. La línea diagonal, que se muestra en cada uno de los paneles sirve como referencia: si los puntos se ubican a lo largo de esta línea, la latencia de las respuestas de observación se habrán emitido de forma indiferenciada y con la misma proporción en ambos eslabones; si los puntos se encuentran por debajo de la referencia, las latencias entre las respuestas de observación habrán ocurrido en mayor proporción en el eslabón terminal que en el eslabón inicial, mientras que, si se ubican por encima la línea diagonal, las latencias habrán tenido una mayor proporción en el eslabón inicial con respecto al terminal.

En el caso del sujeto RP1, la cantidad de puntos dispersos en el primer panel de la Figura 6 es menor, relativa a los otros dos sujetos, esto porque, como se puede ver el Figura 2, las respuestas de observación de este sujeto fueron menores relativo a los otros dos sujetos y para poder calcular una latencia necesariamente debieron de haber ocurrido dos o más respuestas de observación durante un mismo intervalo entre reforzadores. No obstante, la proporción de latencias que se muestran para este sujeto, no parecen tener una tendencia, sino que en algunas ocasiones hubo latencias menores para las respuestas que produjeron en estímulo del eslabón inicial y en ocasiones para las respuestas que produjeron el estímulo del eslabón terminal.

En el caso de los sujetos RP2 y RP3, la mayoría de las latencias tienen una proporción mayor durante el eslabón terminal con respecto al inicial, puesto que la mayoría de los puntos se ubican por debajo del eje de referencia. En el sujeto RP2 la mayor

proporción de latencias son las de 0 a <4 s, es decir, que las respuestas de observación que encendieron los estímulos del eslabón terminal fueron seguidas de otras respuestas de observación con una latencia menor a 4 s. En el caso de RP3, la mayor proporción de latencias también se concentra en el eslabón terminal, pero hubo tanto latencias de 0 a <4 s como latencias mayores a 20 s.

El mismo análisis que se llevó a cabo en la Figura 6 para las condiciones donde estuvieron presentes ambos estímulos, se realizó en la condición donde solo estuvo presente el estímulo del eslabón terminal, sin embargo, los datos se distribuyeron similarmente a cuando estuvieron presentes ambos estímulos. En la condición en la que solo se presentó el estímulo del eslabón inicial, las respuestas de observación en ambos eslabones fueron cercanas a cero y no hubo ninguna latencia que analizar durante esta condición. Por esta razón estos datos no se incluyeron en el presente trabajo.

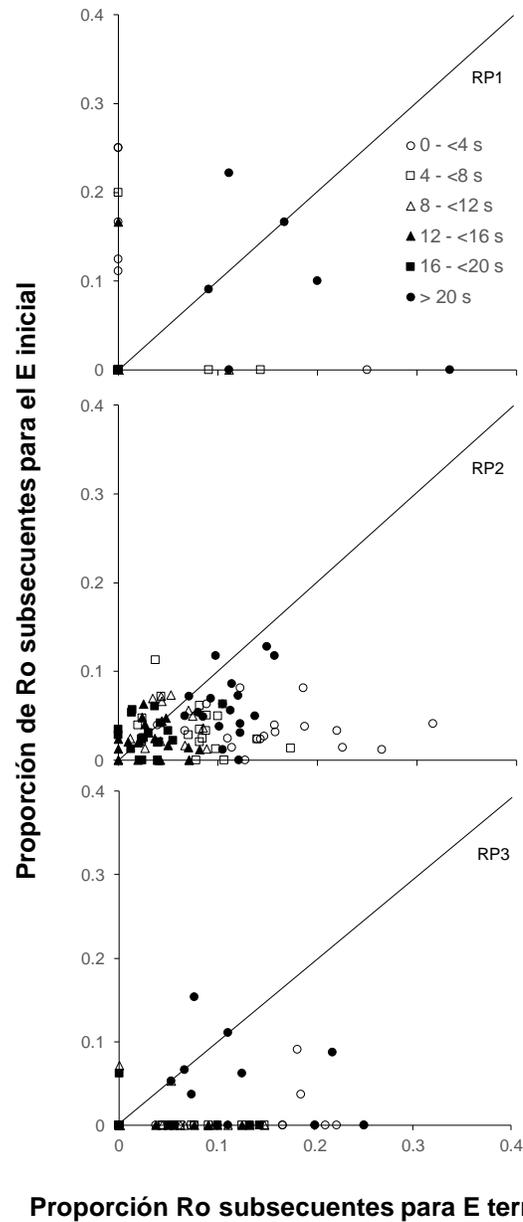


Figura 6. Proporción de respuestas de observación subsuces que ocurrieron con una latencia entre 0 a <4, 4 a <8, 8 a <12, 12 a <16, 16 a <20 y >20 s desde que se apagó el estímulo del eslabón terminal (abscisa) y el estímulo del eslabón inicial (ordenada) durante las seis últimas sesiones de las condiciones en las que estuvieron presentes ambos estímulos. La diagonal representa una línea de referencia que indica indiferencia hacia alguno de los estímulos.

Otra variable que se analizó fue la distribución temporal de respuestas en la palanca de reforzamiento y en la palanca de observación tanto en el eslabón inicial como en el eslabón terminal. Debido a que los programas de reforzamiento vigentes en cada uno de los eslabones eran un programa IV, se obtuvo un promedio de la proporción relativa de respuestas a ambas palancas en los deciles del intervalo entre reforzadores. Para esto, se analizaron las respuestas en la palanca de reforzamiento y en la palanca de observación ocurridas en el eslabón inicial y la duración de este eslabón diferencialmente de las ocurridas en el eslabón terminal y la duración de éste. Se obtuvo el tiempo que estuvo vigente cada eslabón y se dividió en 5 partes iguales, de tal manera que cada subintervalo de tiempo correspondía al 20% de ese eslabón y al 10% del intervalo entre reforzadores; posteriormente, se calculó la proporción relativa de respuestas en cada palanca ocurridas en los subintervalos dividiendo el número de respuestas entre el total de respuestas correspondientes a cada palanca durante el eslabón inicial y el eslabón terminal del intervalo entre reforzadores analizado. Finalmente, se obtuvo el promedio de la proporción relativa de respuestas en la palanca de reforzamiento de cada decil del intervalo entre reforzadores, dividiendo la proporción relativa de respuestas entre el número de intervalos entre reforzadores totales por sesión; en el caso de las respuestas de observación, la proporción relativa de respuestas se dividió entre los intervalos entre reforzadores en los que ocurrieron repuestas de observación.

En las Figuras 7 y 8 se muestra la distribución temporal de respuestas en la palanca de reforzamiento y en la palanca de observación respectivamente, para las últimas seis sesiones de cada condición y para los tres sujetos. En el panel izquierdo de ambas figuras se muestra la condición en la que estuvieron disponibles ambos estímulos, en el panel central

la condición donde sólo se presentó el estímulo del eslabón terminal y en el panel derecho, la condición en la cual sólo se presentaba el estímulo del eslabón inicial. En las condiciones en la que se presentaron los dos estímulos, se analizaron las tres condiciones que estuvieron vigentes de manera conjunta para los sujetos RP2 y RP3, mientras que para el sujeto RP1 se excluyeron del análisis los datos correspondientes a la primera exposición de la condición donde fueron presentados ambos estímulos debido a la variabilidad observada en esta condición respecto a las replicaciones posteriores.

En la Figura 7 se muestra que la distribución temporal de respuestas en la palanca de reforzamiento no tuvo cambios sistemáticos entre las condiciones en las que se presentaron ambos estímulos o cuando se omitió uno de ellos para ninguno de los sujetos.

En cuanto a la distribución temporal de respuestas de observación, en la Figura 8 se muestra que en la condición en la que se presentaron los dos estímulos, para los tres sujetos las respuestas de observación ocurrieron en mayor proporción hacia el final de cada uno de los eslabones y en mayor medida hacia el final del eslabón terminal. De manera general, en el primer 20% del intervalo entre reforzadores no ocurrieron respuestas de observación u ocurrieron con muy poca frecuencia.

Durante la condición en la que se eliminó el estímulo del eslabón inicial, para los tres sujetos, la proporción relativa de respuestas de observación aumentó durante el eslabón inicial, con respecto a la condición en la que se presentaban ambos estímulos, además que parecen mostrar una proporción creciente en cada uno de los eslabones, teniendo una mayor proporción relativa de respuestas hacia el final del eslabón terminal.

Durante la condición en la que se eliminó el estímulo del eslabón terminal, la distribución de la proporción relativa de respuestas de observación parece mantener el

mismo patrón que en las otras condiciones para los tres sujetos, una proporción menor de respuestas de observación al principio de cada uno de los eslabones que aumentó hacia el final de éste; sin embargo, la proporción parece ser mayor durante el eslabón terminal a pesar de que no se presentaba ningún estímulo durante este eslabón.

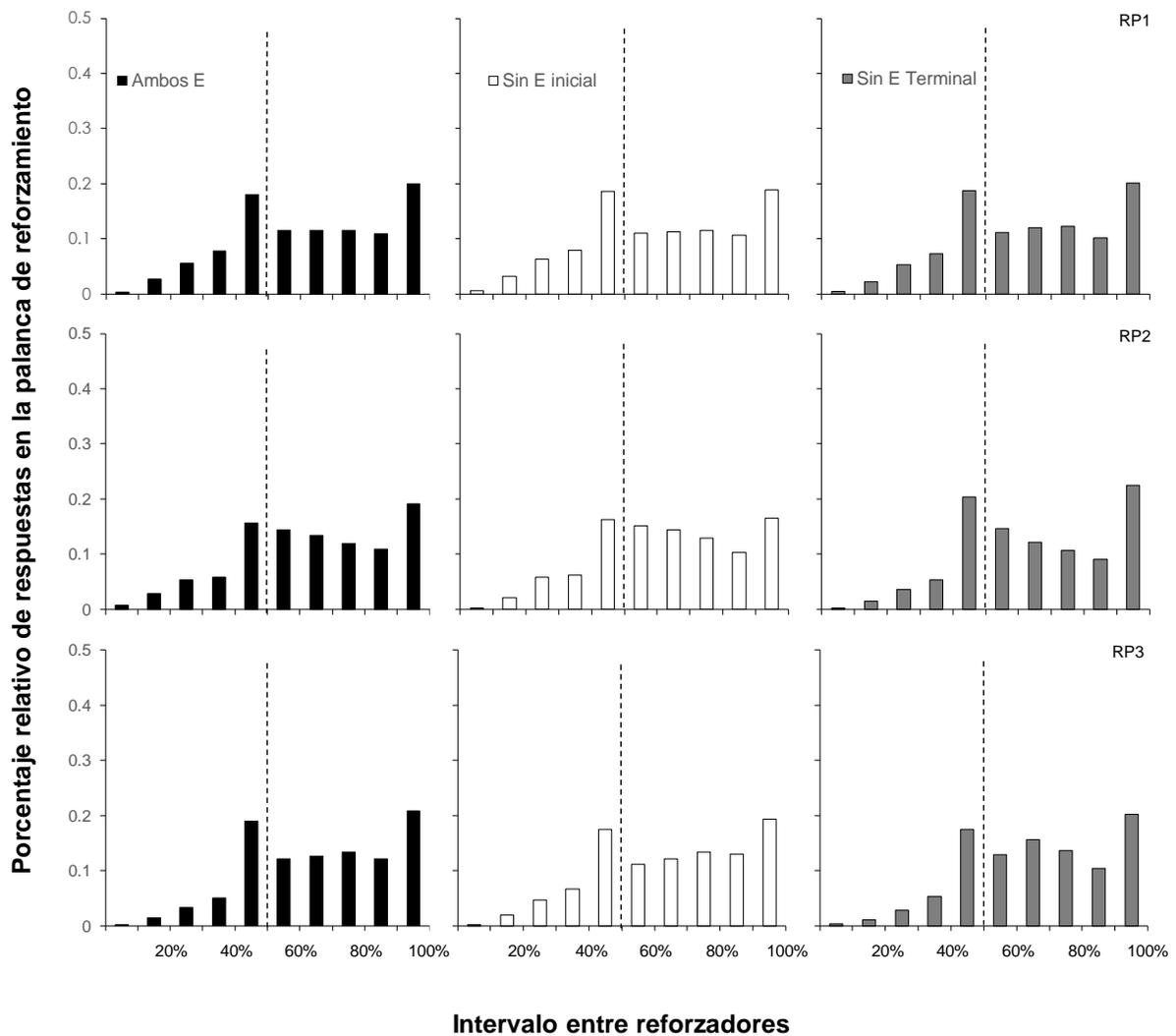


Figura 7. Porcentaje relativo de respuestas en la palanca de reforzamiento en cada decil del intervalo entre reforzadores como promedio de las últimas seis sesiones de cada condición. Para la condición donde se presentaban ambos estímulos se promediaron las tres condiciones a las que estuvieron expuestos los sujetos RP2 y RP3, y las últimas dos exposiciones a esta condición para el sujeto RP1. La línea punteada indica que antes de ella se muestran las respuestas de observación durante el eslabón inicial y posterior a ella las respuestas de observación durante el eslabón terminal.

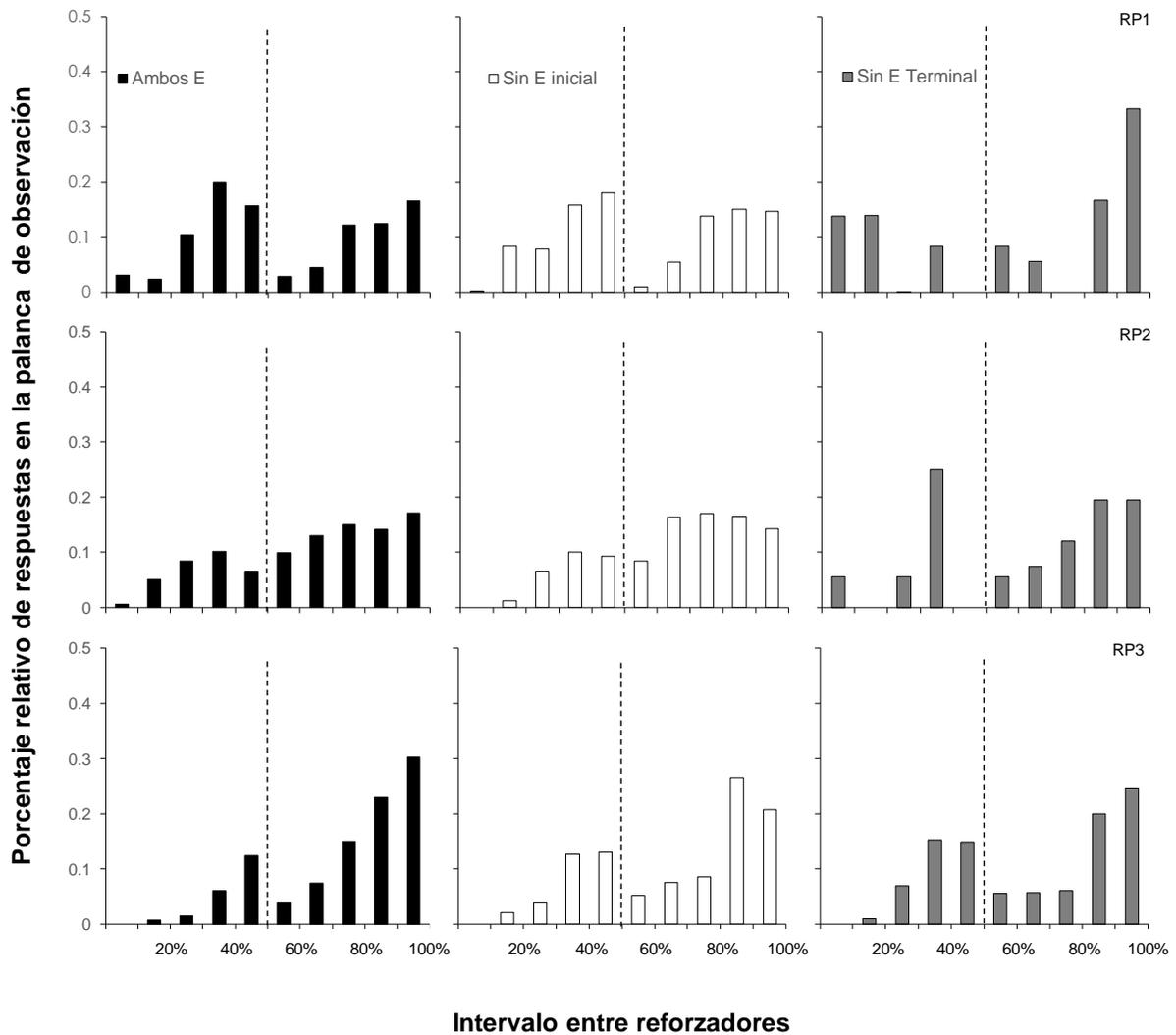


Figura 8. Porcentaje relativo de respuestas de observación en cada decil del intervalo entre reforzadores como promedio de las últimas seis sesiones de cada condición. Para la condición donde se presentaban ambos estímulos se promediaron las tres condiciones a las que estuvieron expuestos los sujetos RP2 y RP3, y las últimas dos exposiciones a esta condición para el sujeto RP1. La línea punteada indica que antes de ella se muestran las respuestas de observación durante el eslabón inicial y posterior a ella las respuestas de observación durante el eslabón terminal.

Discusión

En el presente estudio se investigó la función que adquieren los estímulos que señalan diferencialmente los eslabones en un programa de reforzamiento encadenado. Dicha función se analizó combinando un programa de reforzamiento tándem de dos eslabones IV 20 s IV20 s con un procedimiento de respuestas de observación en ratas como sujetos. Utilizar un programa de reforzamiento IV tuvo como objetivo reducir los efectos de una posible discriminación temporal que podría haber ocurrido con programas de reforzamiento de IF usados previamente (cf. Hendry & Dillow, 1966). Asimismo, las respuestas en la palanca de reforzamiento se disociaron de las respuestas que producían los estímulos asociados con cada eslabón (respuestas de observación); de tal manera que cada respuesta de observación transformó el programa de reforzamiento de tándem a encadenado durante 5 s. Con la finalidad de analizar la función de cada uno de los estímulos, en condiciones sucesivas se eliminó la presentación de uno u otro estímulo. Si el estímulo tuvo una función como reforzador condicionado, se esperaría que eliminar su presentación disminuyera las respuestas de observación.

Para los tres sujetos se establecieron respuestas de observación que produjeron únicamente los estímulos asociados con cada eslabón del programa. Este hallazgo sugiere una función reforzante de los estímulos, que permite establecer y mantener las respuestas de observación.

La interpretación que sugiere que los estímulos de un programa encadenado funcionan como un reforzador condicionado de las respuestas que los producen, ha generado críticas. Uno de los aspectos criticados está relacionado con los patrones de respuesta que se observan en un programa de reforzamiento encadenado en comparación

con un programa de reforzamiento tándem equivalente. Staddon (1983) y Staddon y Cerutti (2003), así como Baum (1973) y Davison y Baum (2006), han sugerido que estas diferencias pueden deberse únicamente a la función discriminativa de los estímulos, que indican la proximidad temporal entre el estímulo y la entrega del reforzador primario. Así, las pausas entre el inicio de un eslabón y el aumento en la tasa de respuesta durante este eslabón irán disminuyendo conforme el eslabón se encuentre más próximo al reforzador primario, a lo cual Staddon y Cerutti (2003) denominaron espera lineal. De acuerdo con estos autores, en un programa de reforzamiento encadenado los estímulos asociados a cada eslabón indican la distancia temporal hacia la entrega del reforzador primario mientras que en un programa tándem no hay tales estímulos discriminativos y por ello las tasas de respuesta son mayores.

Debido a que el procedimiento usado en el presente trabajo separó las respuestas que producen el reforzador primario y el cambio en el eslabón del programa de las respuestas cuya única consecuencia programada fue la presentación de los estímulos, la función discriminativa del estímulo pudo determinarse al analizar las respuestas en la palanca de reforzamiento. Con este propósito, en la Figura 3 se muestra el número de respuestas en la palanca de reforzamiento durante el tiempo que permaneció encendido cada estímulo. En presencia del estímulo del eslabón inicial, las respuestas en la palanca de reforzamiento fueron cercanas a cero, sugiriendo que el estímulo del eslabón inicial señalaba la entrega demorada del reforzador primario, teniendo una función similar a un estímulo delta. Mientras que, en presencia del estímulo del eslabón terminal ocurrieron desde 2 hasta 6 respuestas por estímulo (24 a 48 respuestas por minuto), sugiriendo que este estímulo señalaba la entrega del reforzador primario y funcionó como un estímulo

discriminativo. Si bien esta diferencia pudo deberse a que durante el eslabón inicial la tasa de respuesta a la palanca de reforzamiento fue menor respecto a la tasa durante el eslabón terminal como se muestra en la Figura 1, el hecho de que la tasa de respuesta durante el estímulo del eslabón inicial fuera cercana a cero y la tasa durante el eslabón terminal fue prácticamente el doble de la observada globalmente, sugiere que los estímulos durante el eslabón inicial y terminal tuvieron funciones como estímulo delta y como estímulo discriminativo, respectivamente. Estudios previos que buscaron evidencia para el modelo de espera lineal (véase Staddon & Cerutti, 2003) reportaron que la tasa de respuesta durante el eslabón inicial de un programa encadenado fue menor que la tasa durante el eslabón terminal debido a que el estímulo señala una demora relativamente larga hacia el reforzador. En el presente estudio, el estímulo podía encenderse y apagarse durante el eslabón inicial y el número de respuestas durante el estímulo fue cercano a cero. A pesar de la diferencia en las tasas de respuestas, el hallazgo del presente estudio es congruente con la noción general del modelo de espera lineal debido a que, probablemente, las diferencias se deben a que, en un programa encadenado convencional, el estímulo del eslabón inicial se mantiene encendido durante el eslabón completo y los animales deben responder para que el programa avance al siguiente eslabón. En todo caso, los efectos de los estímulos observados en la palanca de reforzamiento fueron más extremos que los esperados por un modelo de espera lineal.

La distribución temporal de respuestas en la palanca de reforzamiento en cada uno de los eslabones (Figura 7) evidenció que, en los primeros subintervalos del eslabón inicial, el porcentaje relativo de respuestas fue menor que en los primeros subintervalos del eslabón terminal. Este hallazgo es congruente en términos generales con el modelo de espera lineal,

pero, al mismo tiempo parece sugerir que los estímulos no tuvieron un control discriminativo que alterara la distribución de respuestas en la palanca de reforzamiento ya que no hubo diferencias entre condiciones. Podrían existir dos posibles explicaciones para este efecto, una es que es posible que espera lineal dependa de la entrega del reforzador primario y de la distancia temporal respectiva de cada eslabón (véase Staddon, 1983), pero no de los estímulos asociados a cada eslabón. La otra explicación es que es necesario que el estímulo permanezca encendido durante todo el eslabón para observar los efectos esperados por el modelo de espera lineal en la palanca de reforzamiento. Estudios futuros podrían analizar las diferencias entre estas hipótesis.

Aunque los efectos del programa sobre las presiones en la palanca de reforzamiento son, en términos generales, congruentes con la noción de espera lineal, el hecho de que se establecieran y mantuvieran respuestas cuya única consecuencia programada fue la presentación de los estímulos es incongruente con la aproximación general del modelo de espera lineal y sugiere un efecto reforzante de los estímulos (e.g., Hendry & Dillow, 1966). Cabe aclarar que en los procedimientos que involucran programas de reforzamiento encadenado tradicionales, el valor reforzante del estímulo en el eslabón terminal se determina a partir de la tasa de respuesta en el eslabón inicial. El valor reforzante del estímulo del eslabón inicial no puede determinarse sin hacer alguna modificación al procedimiento. Para determinar el valor reforzante de dos estímulos, es necesario agregar un tercer eslabón, de tal forma que la tasa de respuesta en el eslabón inicial es un índice del valor reforzante del estímulo en el eslabón intermedio (e.g., Royalty, et al., 1987). Debido a que en el presente estudio se modificó el programa encadenado en uno tándem con un procedimiento de respuestas de observación que transformó temporalmente el programa

tándem en un encadenado, el valor reforzante del estímulo en el eslabón terminal puede inferirse a partir del número de repuestas de observación en el eslabón terminal y el valor reforzante del estímulo del eslabón inicial puede inferirse a partir del número de respuestas de observación en el mismo eslabón inicial. De esta forma, los datos de la palanca de observación en el presente trabajo son comparables con los análisis de las respuestas en programas encadenados tradicionales de tres componentes (cf. Hendry y Dillow, 1966; Kelleher & Fry, 1962).

Hendry y Dillow (1966) reportaron el establecimiento de respuestas de observación que producían los estímulos asociados a tres eslabones de un programa de reforzamiento encadenado con palomas como sujetos. Igual que en el presente estudio, las respuestas de observación transformaron el programa de reforzamiento tándem en uno encadenado. Hendry y Dillow reportaron que las respuestas de observación aumentaron del primero al segundo eslabón y disminuyeron durante el eslabón terminal. Aunque este último hallazgo es difícil de entender dada la literatura sobre reforzamiento condicionado, pudo haberse debido a la competencia entre las respuestas que producían el reforzador primario mantenidas con programas de IF y las respuestas de observación (Slezak & Anderson, 2014). En el presente estudio, para simplificar el procedimiento se utilizaron únicamente dos eslabones con programas de reforzamiento de IV y se encontró que las respuestas de observación aumentaron del eslabón inicial al eslabón terminal. Este hallazgo es congruente con la noción de que el estímulo en el eslabón terminal es un reforzador condicionado más potente que el estímulo asociado con el eslabón inicial (véase Kelleher & Gollub, 1962).

Aunque la tasa de respuesta de observación puede considerarse como un índice del valor reforzante de los estímulos que producen, los cambios en la tasa pueden deberse a la

competencia entre respuestas y no necesariamente al valor reforzante de los estímulos como lo mostró el estudio de Hendry y Dillow (1966). Por lo tanto, en las siguientes condiciones del estudio se eliminaron los estímulos para determinar su función sobre las respuestas de observación.

De acuerdo con la hipótesis del encadenamiento, un reforzador condicionado adquiere valor reforzante dada su asociación con el reforzador primario y debido a un proceso similar al condicionamiento de orden superior (véase Williams, 1994). En un programa de reforzamiento encadenado de dos eslabones, las respuestas en el eslabón inicial producen la transición al eslabón terminal y con ello el cambio de estimulación, finalmente cumplir con el requisito de respuestas en el eslabón terminal resulta en la entrega del reforzador primario. En este ejemplo, el estímulo en el eslabón terminal, al estar asociado con el reforzador primario, adquiere una función como reforzador condicionado. El estímulo del eslabón inicial al estar asociado con el estímulo del eslabón terminal adquiere propiedades reforzantes, aunque en menor medida que el estímulo del eslabón terminal (véase Williams, Ploog & Bell, 1995). En el presente estudio, al disociar las respuestas que producen únicamente los estímulos asociados con cada uno de los eslabones, de las respuestas que producen el cambio entre eslabones y el reforzador primario al final de la cadena, ambos estímulos pudieron funcionar como reforzadores condicionados de las respuestas de observación que los producían. En este sentido, el estímulo del eslabón terminal podría adquirir propiedades reforzantes dada su asociación con el reforzador primario, y el estímulo del eslabón inicial podría adquirir una función como reforzador condicionado debido a su asociación con el estímulo del eslabón terminal.

La tasa de respuestas de observación que se presenta en la Figura 2, cuando estaban presentes ambos estímulos, es un hallazgo congruente con la hipótesis del encadenamiento, ya que el estímulo del eslabón terminal mantiene una tasa de respuestas de observación mayor que el estímulo del eslabón inicial. Siguiendo con esta hipótesis, en la condición en la que se eliminó el estímulo del eslabón terminal, se rompería la cadena y las respuestas de observación disminuirían; en la Figura 2 también se muestra datos congruentes con esta hipótesis. Por otro lado, al eliminar el estímulo del eslabón inicial la tasa de respuestas de observación debería disminuir solamente en el eslabón inicial. Sin embargo, los datos no apoyan la hipótesis del encadenamiento. La tasa de respuestas de observación en el sujeto RP2 no cambió sistemáticamente e incluso en los sujetos RP1 y RP3 las tasas aumentaron respecto a las condiciones donde se presentó el estímulo.

La tasa de respuestas de observación durante la condición donde se presentaban ambos estímulos y la condición en la que se eliminó el estímulo del eslabón terminal, son congruentes con la hipótesis del encadenamiento, sin embargo, la tasa de respuestas de observación cuando se eliminó el estímulo del eslabón inicial sugiere que este estímulo no adquiere una función reforzante. Este hallazgo es congruente con el estudio reportado por Kendall (1972), quién realizó un experimento para analizar la función de cada estímulo respecto a la distancia temporal con la entrega del reforzador primario.

Kendall (1972) expuso a tres palomas a un IF 3 min con un procedimiento de reloj opcional similar al reportado por Hendry y Dillow (1966) pero en el que podría conceptualizarse que el cambio de un eslabón a otro ocurrió independiente de la respuesta (véase también Slezak & Anderson, 2014). Las respuestas en la tecla central producían el reforzador primario al final del IF 3 min, mientras las respuestas a una tecla lateral

(respuestas de observación) producían un cambio de estimulación en la tecla central. El IF se dividió en subintervalos de 1 min y cada uno estaba asociado con un estímulo diferente. En la primera condición las respuestas de observación producían un cambio de estimulación en los tres segmentos del intervalo. En la una siguiente condición las respuestas de observación sólo producían los estímulos asociados con el primer y segundo subintervalo y en la última condición solo se producía el estímulo asociado con el último subintervalo. Cuando se producían los tres estímulos o cuando solamente se producía el estímulo del último subintervalo, las respuestas en cada segmento del intervalo fueron muy parecidas y tuvieron un aumento del minuto uno al minuto dos y una disminución durante el último subintervalo; este hallazgo sugiere que los estímulos de los primeros subintervalos no tienen un efecto reforzante. Sin embargo, cuando no se presentaba el estímulo asociado al último subintervalo, se produjo una tasa de respuestas de observación cercana a cero durante todo el IF. Kendall argumentó que la eliminación del estímulo asociado con el subintervalo final produjo un descenso en la tasa de respuesta en los subintervalos precedentes porque se eliminó la fuente de reforzamiento para las respuestas de observación. A pesar que estos resultados sugieren que los estímulos en mayor lejanía temporal con el reforzador primario no resultan ser reforzadores condicionados, es difícil determinar si los resultados de Kendall podrían replicarse usando un programa de reforzamiento tándem con un procedimiento de observación añadido, debido a que solo se utilizó un programa de reforzamiento IF y la tasa de respuestas de observación pudo haber interactuado con el patrón típico de respuestas de los IF (Ferster & Skinner, 1957).

Hendry y Dillow (1966), así como el estudio reportado por Royalty, et al. (1987), sugieren que todos los estímulos asociados a un programa de reforzamiento encadenado

tienen propiedades reforzantes, aunque los autores apoyaron una hipótesis basada en la transmisión de información de los estímulos, sus datos fueron congruentes con la hipótesis del encadenamiento. En el presente estudio es difícil sostener dicha hipótesis ya que, al analizar la función de los estímulos sobre el efecto de las respuestas de observación en las diferentes condiciones se puede observar una tasa de respuestas de observación distinta en presencia de ambos estímulos y en ausencia de alguno de ellos. La diferencia encontrada en la tasa de respuestas en cada uno de los eslabones sugiere que no todos los estímulos son reforzantes y que las respuestas de observación en el eslabón inicial podrían estar mantenidas por el estímulo del eslabón terminal (véase también Kendall, 1972), pues como se observa en la Figura 8, las respuestas de observación durante el eslabón inicial ocurrieron en mayor proporción al final de este eslabón para los tres sujetos. En este sentido, el estímulo del eslabón terminal parece ser reforzante, pero no el estímulo del eslabón inicial. Cuando solo ocurría el estímulo del eslabón inicial contingente a las respuestas de observación, estas disminuyeron; mientras que, cuando se presentaban ambos estímulos, para al menos dos sujetos, parece que tuvo un efecto supresor de las respuestas de observación. Esto sugiere que el estímulo del eslabón inicial pudo adquirir una función aversiva.

Thomas (1966) y Leung (1994), sugirieron que el estímulo del eslabón inicial de un programa de reforzamiento encadenado es aversivo ya que señala la ausencia de la entrega del reforzador primario en su presencia, o la asociación entre este estímulo y la entrega del reforzador primario es muy débil. Thomas (1966) realizó un estudio en el que expuso a palomas a un programa de reforzamiento encadenado con dos eslabones. El primer eslabón era un IF 3 min, señalado por una luz verde o roja, y el segundo fue un componente de 30 s

con un programa de razón fija (RF) 50 señalado con el cambio del color de la luz. En los últimos 10 s del componente de RF, se encontraba disponible una tecla de evitación, si las palomas emitían una respuesta en esta tecla, se posponía el cambio al IF durante 30 s. En condiciones sucesivas, Thomas aumentó el valor del IF mientras permanecía constante el requisito de RF y también aumento el requisito del programa de RF mientras permanecía constante el IF 3 min. Thomas encontró que las palomas evitaban consistentemente regresar al IF respondiendo en la tecla de evitación durante el componente de RF, además de que la tasa de respuestas de evitación aumentaba en función del aumento en la duración de IF. Por otro lado, en las condiciones en las que se aumentó el requisito de RF, las respuestas de evitación disminuyeron en programas RF 40, 60 y 80, pero aumentaron en programas de RF100 y RF120. A pesar que los tres sujetos, consistentemente respondieron para no regresar al programa de IF, es difícil interpretar si las palomas respondían de esta manera para evitar el contacto con el estímulo en el IF o sólo para prolongar el tiempo del componente de RF, como cuando las RF eran muy grandes y las palomas presionaban la tecla de evitación para tener tiempo suficiente para cumplir con el requisito de respuestas.

Para tratar aclarar la función del estímulo del eslabón inicial, Leung (1994) realizó un estudio en el cual expuso a palomas a un programa encadenado de dos eslabones tiempo fijo (TF) TF, cada uno de los eslabones se encontraba señalado diferencialmente con una luz; concurrentemente con el programa de reforzamiento TF se encontraba disponible una tecla que apagaba los estímulos durante 1 s. Leung encontró que los sujetos presionaban consistentemente la tecla que apagaba los estímulos durante el eslabón inicial pero no la presionaban durante el eslabón terminal. De esta manera, parece ser que el estímulo del eslabón inicial adquiere propiedades aversivas respecto al estímulo del eslabón terminal.

Estos hallazgos, aunque inconsistentes con el resultado de Royalty et al. (1987) son congruentes con el reporte de que las tasas de respuesta en los primeros eslabones de un programa de reforzamiento tándem son más altas que en programas de reforzamiento encadenados (Kelleher & Gollub, 1962).

En la Figura 2 se puede observar, con al menos dos sujetos, que cuando se eliminó el estímulo del eslabón inicial, las respuestas de observación aumentaron en ambos eslabones. Si el estímulo del eslabón inicial adquiere una función aversiva, el aumento de las repuestas de observación sería congruente con ello, al no existir consecuencias aversivas.

Si el estímulo del eslabón inicial tuviese alguna función aversiva, la pausa post reforzamiento debería haber disminuido respecto a las condiciones en las que se presentaron ambos estímulos. En la Figura 5 se puede observar que, en la condición en la que se eliminó el estímulo del eslabón inicial, la pausa post reforzamiento sólo disminuyó para el sujeto RP3, mientras que para RP1 y RP2 no hubo cambios respecto a las condiciones donde se presentaron ambos estímulos. Este efecto, sin embargo, también pudo deberse a que la frecuencia de la presentación del estímulo del eslabón inicial fue relativamente baja respecto a la frecuencia de la presentación del estímulo del eslabón terminal en las condiciones en las que se presentaban ambos estímulos.

Siguiendo con la suposición de que el estímulo del eslabón inicial es aversivo, Gaynor y Shull (2002), argumentaron que los sujetos en procedimientos de respuestas de observación realizan una observación selectiva, es decir, que las respuestas de observación que producen los estímulos asociados con algún período de tiempo (extinción o de reforzamiento), se emiten de manera diferencial. Con base en esta hipótesis, las respuestas

de observación durante el eslabón terminal deberían establecer la ocasión para seguir “observando”, mientras que las respuestas de observación durante el eslabón inicial deberían de tener el efecto contrario. Así, si el estímulo del eslabón inicial tiene una función aversiva, en la Figura 6 se debería observar una mayor proporción de latencias cortas por debajo de la línea de referencia. Este efecto se observa sólo para los sujetos RP2 y RP3. En estos sujetos parece que las respuestas de observación seguidas de la aparición del estímulo del eslabón terminal tienden a ser seguidas de otras respuestas de observación con un período corto de tiempo entre ellas. En el caso del sujeto RP1, la proporción de latencias largas también se encuentra en las respuestas de observación que producen los estímulos del eslabón terminal. Los datos analizados bajo la hipótesis de la observación selectiva, si bien sugieren que para uno de los sujetos el estímulo del eslabón inicial parece no ser aversivo, las respuestas de observación que se produjeron en cada uno de los eslabones, encendían los estímulos durante 5 s y cada uno de los eslabones tenía una duración promedio de 20 s, además, como se observa en la Figura 2, la tasa de respuesta en la palanca de observación para el sujeto RP1 fue menor respecto a los otros dos sujetos. Por lo tanto, las respuestas de observación ocurrieron con tiempos largos entre respuestas y esto pudo provocar que no fuera claro el efecto de observación selectiva.

Los efectos de la eliminación del estímulo del eslabón terminal son congruentes con la hipótesis de que este estímulo adquirió propiedades reforzantes debido a que la tasa de respuestas de observación disminuyó cuando éste no se presentaba. Sin embargo, la función del estímulo del eslabón inicial pareció ser similar a un estímulo neutro para una rata (RP2) y adquirió una función como estímulo aversivo (RP1 y RP3) para las otras dos. Explicar por qué el estímulo del eslabón terminal adquiere una función reforzante y el estímulo del

eslabón inicial podría tener una función aversiva es problemático en términos de la hipótesis del encadenamiento y del modelo de espera lineal. Sin embargo, la función de los estímulos asociados con los eslabones de una cadena, y que son congruentes con la función reforzante de los estímulos, puede ser explicada por la hipótesis de la reducción de la demora, la cual plantea que la efectividad de un estímulo como reforzador condicionado depende de la reducción del tiempo hacia la entrega del reforzador primario a partir de que se enciende el estímulo (Fantino, 1977) relativo al intervalo entre reforzadores. Así, el valor reforzante de un estímulo se encuentra determinada por:

$$A = f\left(\frac{T-t_a}{T}\right)$$

donde A es el valor reforzante de un estímulo, t_a representa el intervalo de tiempo entre que se enciende el estímulo y la entrega del reforzador primario, mientras que T es el tiempo promedio entre reforzadores (Fantino, Preston & Dunn, 1993). Este modelo contempla que, mientras la ocurrencia de un estímulo este correlacionado con una mejora, en términos de la proximidad temporal hacia la entrega del reforzador primario, mayor será la efectividad de ese estímulo como reforzador condicionado. Este modelo fue desarrollado y aplicado principalmente en investigaciones de toma de decisiones, especialmente en procedimientos que utilizan programas de reforzamiento concurrentes encadenados.

Case y Fantino (1981) extendieron la hipótesis de la reducción de la demora de situaciones de toma de decisiones a procedimientos en los cuales la producción de un estímulo no afecta la distribución temporal del reforzador primario como lo son procedimientos de observación. Case y Fantino llevaron a cabo un estudio en donde expusieron a palomas a un programa de reforzamiento mixto con tres componentes con respuestas de observación que transformaban el programa en un múltiple. En un programa

de reforzamiento múltiple el reforzador se presenta en programas de reforzamiento alternados y hay estímulos exteroceptivos que indican el componente que se encuentra en curso; un programa mixto es equivalente a un múltiple, pero sin estímulos exteroceptivos. Con base en la hipótesis de la reducción de la demora, Case y Fantino argumentaron que el estímulo correlacionado con un programa de reforzamiento que proporcionara de manera más inmediata el reforzador primario estaría correlacionado con una reducción en el tiempo hacia el reforzador y debería de mantener respuestas de observación; mientras que un estímulo correlacionado con un incremento en el tiempo hacia el reforzador no debería de mantener respuestas de observación. En condiciones sucesivas los autores expusieron a palomas a un programa mixto IF 20 s, IF 40 s, IF 180 s en una condición e IF 20 s, IF 120 s, IF 180 s, en otra condición. Los autores reportaron que la tasa de respuesta de observación, en cada uno de los componentes, estuvo en función de la reducción de la demora correlacionada con el estímulo que producían.

En el presente estudio se utilizó un programa de reforzamiento encadenado de dos eslabones IV 20 s IV 20 s, de tal manera que el tiempo promedio entre reforzadores era de 40 s. De acuerdo con la ecuación para calcular el valor reforzante de un estímulo, T es igual a 40 s; para el estímulo del eslabón inicial el intervalo de tiempo entre que se encendía el estímulo y la entrega del reforzador primario, eran también 40 s, mientras que para el estímulo del eslabón terminal 20 s. El estímulo del eslabón inicial, entonces, no representaba una reducción en la demora, mientras que el estímulo del eslabón terminal la reducía en un 50 %. Los resultados que se presentan en la Figura 8 son congruentes con el modelo de la hipótesis de la reducción de la demora. En la distribución temporal de la proporción relativa de respuestas de observación, durante cada uno de los eslabones se

observa que en la condición en la que se presentaron ambos estímulos, para los tres sujetos, la proporción de respuestas aumentó desde el primer 20% de cada eslabón hacia el final de cada uno de ellos respectivamente, esta misma tendencia se mantuvo cuando solo se presentó el estímulo del eslabón terminal, sugiriendo que el estímulo del eslabón inicial no tenía una función en este patrón de respuesta. Sin embargo, cuando sólo se presentaba el estímulo del eslabón inicial las respuestas de observación fueron cercanas a cero y no mostraron una tendencia de distribución.

Explicar los hallazgos de este experimento en términos de la hipótesis de la reducción de la demora contempla que el estímulo del eslabón terminal es reforzante ya que su presentación reduce un 50 % el tiempo hacía la entrega del reforzador. Por lo tanto, este estímulo adquiere una función de reforzador condicionado capaz de establecer y mantener respuestas de observación. Una vez que este estímulo es eliminado, la tasa de respuesta disminuye a valores cercanos a cero debido a que, con su omisión, no se reduce la demora a la entrega del reforzador y por ello las respuestas de observación ya no se encuentran bajo el control del reforzador condicionado. En el caso del estímulo del eslabón inicial, su aparición no reduce la demora hacía la entrega del reforzador primario y no adquiere una función como reforzador condicionado.

Una versión simétrica de la hipótesis de la reducción de la demora, propuesta por Jwaideh y Mulvaney (1976), sostiene que los estímulos correlacionados con un incremento en el tiempo hacia el reforzador primario pueden adquirir una función como estímulos aversivos. Estos autores expusieron a palomas a un programa de reforzamiento mixto de dos componentes IV30 s IV 2 min, con tres teclas disponibles, de las cuales, en dos de ellas se encontraba vigente un procedimiento de respuestas de observación. Su principal interés

era averiguar si un estímulo correlacionado con un programa de reforzamiento con una menor frecuencia de reforzamiento relativa funcionaba como un estímulo aversivo que suprimiera las respuestas de observación. Las respuestas a las teclas de observación producían los estímulos correlacionados a cada uno de los programas de reforzamiento de cada componente, las respuestas a una tecla producían ambos estímulos, y en condiciones sucesivas, las presiones a la otra tecla solo producían uno de los estímulos. Jwaideh y Mulvaney reportaron que cuando las respuestas de observación producían sólo el estímulo correlacionado con el componente IV 2 min, los sujetos presionaban en mayor proporción la tecla que producía ambos estímulos; mientras que, cuando las respuestas de observación sólo producían el estímulo correlacionado con el componente IV 30 s, los sujetos respondían más en esta tecla que en la que producía ambos estímulos. Con base en sus resultados, Jwaideh y Mulvaney argumentaron que el estímulo asociado con un programa de reforzamiento relativamente “pobre” no mantiene respuestas de observación y adquiere la función de estímulo aversivo porque indica una mayor lejanía temporal con la entrega del reforzador primario y propusieron la hipótesis del aumento en la demora del castigo condicionado: mientras el estímulo indique una mayor lejanía temporal con la entrega del reforzador primario este estímulo será más aversivo.

Si bien los programas de reforzamiento múltiples y mixtos difieren significativamente de los programas de reforzamiento encadenados y tándem, es posible que las respuestas en los eslabones iniciales de un programa de reforzamiento encadenado o tándem sean parecidos a componentes de programas mixtos o múltiples con una frecuencia relativamente baja de reforzamiento. En este sentido, de acuerdo con Jwaideh y Mulvaney (1976) si el estímulo del eslabón inicial fuera aversivo, la tasa de respuestas de observación,

cercana a cero, durante la condición en la que solo se presentó el estímulo del eslabón inicial es congruente con la hipótesis del incremento en la demora, así como con el aumento en la tasa de respuestas de observación para los sujetos RP1 y RP3 en la condición en la que sólo se presentó el estímulo del eslabón terminal. Sin embargo, la tasa de respuestas de observación del sujeto RP1 fue muy parecida durante todas las condiciones a las que estuvo expuesto, sugiriendo que el estímulo del eslabón inicial no representaba un incremento en la demora que mantuviera menos respuestas de observación respecto a las respuestas para producir el estímulo del eslabón terminal.

Cuando se presentó por sí solo el estímulo del eslabón inicial, las respuestas de observación fueron cercanas a cero para los tres sujetos, según la hipótesis de la reducción de la demora este estímulo no representaba ninguna mejora en cuanto a la proximidad temporal a la entrega del reforzador primario y por ello no puede funcionar como reforzador condicionado. Sin embargo, tampoco parece haber adquirido una función como estímulo aversivo para todos los sujetos. Es posible que el estímulo del eslabón inicial haya adquirido una función discriminativa que señalaba la entrega de un reforzador condicionado en el siguiente eslabón.

Con base en los resultados presentados en este trabajo, se ha encontrado evidencia que apoya la utilidad del uso del concepto de reforzamiento condicionado para estudiar los patrones de respuesta en programas de reforzamiento encadenado y aporta evidencia de que el estímulo del eslabón terminal funciona como un reforzador condicionado. Si bien los datos sugieren que el estímulo del eslabón inicial podría adquirir propiedades aversivas es necesario seguir investigando bajo qué condiciones ocurre este efecto; por ejemplo, si se

agregaran eslabones en el programa de reforzamiento encadenado o si se manipulara la frecuencia de reforzamiento en cada uno de los eslabones.

Referencias

- Baum, W. M. (1973). The correlation based law of effect. *Journal of the Experimental Analysis of Behavior*, 20, 137-153.
- Bejarano, R., & Hackenberg, T. D. (2007). IRT-stimulus contingencies in chained schedules: Implications for the concept of conditioned reinforcement. *Journal of the Experimental Analysis of Behavior*, 88, 215-227
- Case, D. A., & Fantino, E. (1981). The delay–reduction hypothesis of conditioned reinforcement and punishment: observing behavior. *Journal of the Experimental Analysis of Behavior*, 35, 93-108.
- Catania, C. (1975). *Investigación contemporánea en conducta operante*. México: Trillas.
- Costa, C., & Cançado, C. (2012). Stability check: a program for calculating the stability of behavior. *Revista Mexicana de Análisis de la Conducta*, 38(1), 61-71.
- Davison, M., & Baum, W. (2006). Do conditional reinforcers count? *Journal of the Experimental Analysis of Behavior*, 86, 269-283.
- Dinsmoor, J. (1983). Observing and conditioned reinforcement. *Behavioral and Brain Science*, 6, 693-728.
- Escobar, R. (2016). *3D printing, Laser Cutting, and Behavior Analysis*. En Lattal, K. A. (coordinador). Behavior Watch. The Aubrey Daniels Institute. Recuperado de <http://aubreydaniels.com/institute/behavior-watch/3d-printing-laser-cutting-and-behavior-analysis>
- Escobar, R., & Pérez-Herrera, C. (2015). Low-Cost USB interface for operant research using Arduino and Visual Basic. *Journal of the Experimental Analysis of Behavior*, 103, 427-435.

- Fantino, E. (1977). Conditioned reinforcement: Choice and information. En W. Honig, & J. Staddon, *Handbook of operant behavior* (pp 313-339). Englewood Cliffs, NJ: Prentice-Hall.
- Fantino, E., Preston, R. A., & Dunn, R. (1993). Delay reduction: current status. *Journal of the Experimental Analysis of Behavior*, 60, 159-169.
- Ferster, C., & Skinner, B. F. (1957). *Schedules of reinforcement*. Englewood Cliffs, NJ: Prentice-Hall.
- Fleshler, M., & Hoffman, H. S. (1962). A progression for generating variable interval schedules. *Journal of the Experimental Analysis of Behavior*, 5, 529-530.
- Gaynor, S., & Shull, R (2002). The generality of selective observing. *Journal of the Experimental Analysis of Behavior*, 77, 171-187.
- Gollub, L. (1977). Conditioned reinforcement: schedule effects. En W. Honig, & J. Staddon, *Handbook of operant behavior* (pp. 288-312). Englewood Cliffs, NJ: Prentice-Hall.
- Hendry, D., & Dillow, P. (1966). Observing behavior during interval schedules. *Journal of the Experimental Analysis of Behavior*, 9, 337-349.
- Jwaideh, A., & Mulvaney, D. (1976). Punishment of observing by a stimulus associated with the lower of two reinforcement frequencies. *Learning and Motivation*, 7, 211-222.
- Kelleher, R., & Gollub, L. (1962). A review of positive conditioned reinforcement. *Journal of the Experimental Analysis of Behavior*, 5(4), 543-597.
- Kelleher, R., & Fry, W. (1962). Stimulus functions in chained fixed-interval schedules. *Journal of the Experimental Analysis of Behavior*, 5, 167-173.

- Keller, F. S., & Schoenfeld, W. N. (1950). *Principles of Psychology*. New York: Appleton-Century-Crofts.
- Kendall, S. B. (1972). Some effects of response-dependent clock stimuli in a fixed-interval schedule. *Journal of the Experimental Analysis of Behavior*, *17*, 161-168.
- Leung, J. P. (1994). Psychological distance to reward: The aversiveness of the first component stimulus in a chain. *Behavioural Processes*, *32*, 67-78.
- Lieberman, D. (1972). Secondary reinforcement and information as determinants of observing behavior in monkeys (*Macaca mulatta*). *Learning and Motivation*, *3* (3), 341-358.
- Malagodi, E. F., DeWeese, J., & Johnston, J. M. (1973). Second-order schedules: A comparison of chained, brief-stimulus, and tandem procedures. *Journal of the Experimental Analysis of Behavior*, *20*, 447-460.
- Mazur, J. E. (2006). *Learning and Behavior* (Sixth Edition). Upper Saddle River, NJ: Prentice Hall.
- Royalty, P., Williams, B., & Fantino, E. (1987). Effects of delayed conditioned reinforcement in chain schedules. *Journal of the Experimental Analysis of Behavior*, *47*, 41-56.
- Shoenfeld, W., Cumming, W., & Hearst, E. (1956). On the classification of reinforcement schedules. *Proceedings of the National Academy of Sciences*, *42*, 563 - 570.
- Skinner, B. F. (1934). The extinction of chained reflexes. *Proceedings of the National Academy of Sciences of the United States of America*, *20*, 234-237.
- Skinner, B. F. (1938). *The Behavior of Organisms*. New York: Appleton-Century-Crofts.

- Slezak, J., & Anderson, K. (2014). Observing of chain-schedule stimuli. *Behavioral Processes, 105*, 19-27.
- Staddon, J. E. R. (1983). *Adaptive learning and behavior*. New York: Cambridge University Press.
- Staddon, J. E. R., & Cerutti, D. T. (2003). Operant Conditioning. *Annual Review of Psychology, 54*, 115-144.
- Thomas, J. R. (1966). Avoidance of a return to the first component of a chain from the terminal component. *Journal of the Experimental Analysis of Behavior, 9*, 435-441.
- Wallace, F., Osborne, S., & Fantino, E. (1982). Conditioned reinforcement in two-link chain schedules. *Behaviour Analysis Letters, 2*, 335-344.
- Williams, B. A. (1994). Conditioned reinforcement: Experimental and theoretical issues. *The Behavior Analyst, 17*(2), 261-285.
- Williams, B. A., Ploog, B. O., & Bell, M. C. (1995). Stimulus devaluation and extinction of chain schedule performance. *Animal Learning & Behavior, 24*, 104-114.
- Wyckoff, L. B., Jr. (1952). The role of observing responses in discrimination learning. Part I. *Psychological Review, 59*, 431-442.
- Wyckoff, L. B., Jr. (1969). The role of observing responses in discrimination learning. Part II. In D. P. Hendry (Ed.), *Conditioned reinforcement* (pp. 237-260). Homewood, IL: Dorsey Press.