



**UNIVERSIDAD NACIONAL AUTÓNOMA DE  
MÉXICO**

---

---

**FACULTAD DE PSICOLOGÍA**

**PROGRAMA AJUSTABLE PARA EL  
ENTRENAMIENTO DE RESPUESTAS DE  
OBSERVACIÓN EN RATAS**

**T E S I S**

**QUE PARA OBTENER EL TÍTULO DE:  
LICENCIADO EN PSICOLOGÍA**

**P R E S E N T A:  
EDUARDO MEZA GARDUÑO**

**DIRECTOR: DR. ROGELIO ESCOBAR HERNÁNDEZ  
REVISORA: DRA. ALICIA ROCA COGORDAN**

**COMITÉ:  
DR. RAÚL ÁVILA SANTIBÁÑEZ  
DR. ÓSCAR ZAMORA ARÉVALO  
DR. LUIS RODOLFO BERNAL GAMBOA**



Apoyo del proyecto PAPIIT IN305819  
Ciudad Universitaria, CD.MX., 2023



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.

*“Those supposedly cold, emotionally sterile laboratories create plenty of heat.”*

— Murray Sidman

*Para mi mamá y mi papá.*

## Agradecimientos

Al Dr. Rodolfo Bernal Gamboa, al Dr. Óscar Zamora Arévalo y al Dr. Raúl Ávila Santibáñez, sinodales del presente trabajo, por su tiempo y valiosas aportaciones que me permitieron concluir esta tesis.

A la Dra. Alicia Roca Cogordan, revisora del presente trabajo, por su tiempo, atención y apoyo para la revisión de este trabajo. Gracias por sus inspiradoras clases, que me mostraron las aplicaciones del análisis de la conducta y lo mucho que se puede hacer por el mundo con nuestra ciencia.

A mi director de tesis y mentor, el doctor Rogelio Escobar Hernández. Gracias, Rogelio, por introducirme al análisis de la conducta, en donde encontré una forma diferente y fascinante para ver la psicología y también el mundo. Gracias por todas tus enseñanzas, tu guía, paciencia y apoyo, tanto en la realización de este trabajo, como en mi formación como analista de la conducta. Sobre todo, gracias por tu trato siempre amable y por generar un ambiente de confianza. Espero poder convertirme en un analista de la conducta tan excepcional como tú.

A mis compañeras y compañeros de laboratorio, Olga, Katya, Sandra, Brissa, Rodrigo, Sebastián, Alexis y Brenda, por todo lo compartido en los seminarios y clases en las que nos formamos juntos. Agradezco a Brissa y Rodrigo, por su guía desde que entré al laboratorio, y en especial a Brissa, por el apoyo que me brindó en los aspectos técnicos y teóricos para conducir los experimentos del presente trabajo. Igualmente agradezco a Moisés, Antonio, Sandra, Ángel, Maira y Olga por las charlas, risas y la compañía durante la conducción de mis experimentos.

A mis padres, Marisela y Eduardo, por todo su amor, apoyo y trabajo. Muchas gracias por procurarme y cuidarme, por esforzarse para brindarme todo lo que yo necesitara. Gracias por acompañarme en cada paso y cada logro. Todo esto es por ustedes. Los amo.

A mi hermano, Alexis, por ser mi mejor compañero desde niño y mi persona favorita. Gracias por todo lo que hemos vivido juntos compartiendo prácticamente todo, desde escuelas hasta nuestra afición por la música, la tecnología y el análisis de la conducta. Aprecio tu compañía y te extraño siempre. Me has inspirado, enseñado y ayudado más de lo que crees.

A mis tíos y tías, Guillermo, Guadalupe, Gerardo, Luz y Adolfo, por brindarme una familia tan alegre, cálida y unida. Les agradezco profundamente el espacio seguro y lleno de risas que crearon para los más jóvenes. Y por supuesto, agradezco a mi abuela, María de la Luz Ruiz Barrientos. Gracias, abuelita, por iluminar mi vida desde niño con tu amor inagotable. Te tengo presente en todo lo que hago y cada paso que doy. Ojalá estuvieras aquí para ver esto.

A las amistades que me han acompañado durante un largo tiempo. A Karen, Pamela, Diana, Lorena, Ceci y Rosa, por tantos años de amistad y por todos los momentos increíbles que siempre recuerdo con mucho cariño. A Areli, mi colega, por tu escucha, apoyo y el sarcasmo compartido que siempre nos provoca risas. Gracias, amigas, por seguir aquí. A Mich y Crash, por todo el tiempo en el que compartimos la música, así como nuestras mayores tristezas y alegrías.

A mis amistades de la facultad, Isabel, Diana, Alan, Diego y Johana. Gracias por ser un espacio tanto para las conversaciones más serias y académicas, como para las fiestas, el soporte emocional y el apoyo mutuo.

A Blanca, por ser mi amiga, mi confidente, mi espacio seguro, mi motivación, mi equipo, mi inspiración, mi colega y mi hogar. Gracias por hacer mi vida más feliz al elegirme cada día. Espero aportar lo mismo a tu vida y deseo que compartamos muchas cosas más. Te amo y admiro inmensamente.

## Tabla de contenido

Resumen .....	1
Introducción.....	2
Experimento 1.....	17
Método .....	19
Sujetos .....	19
Aparatos.....	19
Procedimiento .....	20
Resultados .....	24
Programa ajustable.....	24
Fase de prueba: Procedimiento de observación estándar .....	26
Discusión.....	27
Experimento 2.....	32
Método .....	33
Sujetos .....	33
Aparatos.....	33
Procedimiento .....	34
Resultados .....	35
Programa ajustable.....	35
Fase de prueba: Procedimiento de observación y prueba de función de los estímulos .....	37
Discusión.....	42
Discusión general.....	46
Referencias .....	57

**Lista de tablas y figuras**

<i>Figura 1.</i> Tasa de respuesta por agua, de RO y frecuencia de reforzamiento en el programa ajustable.....	25
<i>Figura 2.</i> Tasa de respuesta por agua, de RO y frecuencia de reforzamiento en la fase de prueba .....	27
<i>Figura 3.</i> Tasa de respuesta por agua, de RO y frecuencia de reforzamiento en el programa ajustable del Experimento 2 .....	36
<i>Figura 4.</i> Tasa de respuesta en la palanca de reforzamiento durante las tres condiciones de prueba .....	38
<i>Figura 5.</i> Tasa de RO durante las tres condiciones de prueba .....	40
<i>Figura 6.</i> Frecuencia de reforzamiento durante las tres condiciones de prueba .....	41

## Resumen

Las respuestas de observación (RO) son operantes que exponen a los organismos a estímulos discriminativos. Las RO han sido de interés en la literatura porque permiten estudiar la función reforzante y el control discriminativo de los estímulos. En los estudios previos de RO, sin embargo, se encuentran inconsistencias y problemas en el entrenamiento de RO. Algunos autores reportan un tiempo relativamente largo para entrenarlas y que no siempre se logran establecer en todos los sujetos. En el presente trabajo se probó un programa ajustable que facilite y reduzca el tiempo para establecer RO. En el Experimento 1 se expuso a cuatro ratas macho Wistar a un programa mixto ajustable IA5 s extinción. Las RO encendían solamente el E+. En la fase de prueba se expuso a los sujetos a un procedimiento de observación tradicional para evaluar si las RO se mantenían. En el Experimento 2 se expuso a cuatro ratas macho Wistar a un programa ajustable igual al del Experimento 1. Las RO encendían el E+ y el E-. En la fase de prueba se expuso a los sujetos a un procedimiento de observación. Posteriormente, con el fin de determinar la función de los estímulos, se expusieron a una fase con los estímulos invertidos; esto es, el E+ en la fase de prueba señalaba extinción y el E- el componente de reforzamiento. En la última fase se reestableció la fase de prueba. El programa ajustable estableció RO consistentemente en todos los sujetos. Invertir los estímulos en el Experimento 2 probó la función discriminativa y mostró evidencia del valor reforzante de los estímulos. El programa ajustable podría ser útil para entrenar RO en estudios posteriores.

*Palabras clave:* respuestas de observación, adquisición, reforzamiento condicionado, programa ajustable, discriminación

## Introducción

Las respuestas de observación (RO) son aquellas que exponen a los organismos a estímulos discriminativos sin alterar la disponibilidad del reforzador (Wyckoff, 1952, 1969). Las RO fueron descritas por primera vez por Wyckoff en 1952, en un estudio de discriminación de estímulos. En su procedimiento, Wyckoff utilizó dos grupos de palomas. Expuso a ambos grupos a un programa mixto intervalo fijo (IF) 30 s extinción, en el que los componentes alternaban de manera aleatoria y la tecla de respuesta se encontraba iluminada de color blanco. Debajo de la tecla, en el suelo de la cámara experimental, se encontraba un pedal que las palomas podían presionar. Para el primer grupo, las presiones al pedal cambiaban el color de la tecla a rojo si las palomas se encontraban en el componente de reforzamiento, y a verde, si se encontraban en el componente de extinción. De esta forma, las presiones al pedal (RO) cambiaban el programa mixto a un programa múltiple para ese grupo. Para el segundo grupo, las presiones al pedal cambiaban el color de la tecla a rojo o verde sin correlacionar con el componente vigente. Para ambos grupos, dejar de presionar el pedal regresaba la tecla a color blanco. Wyckoff registró la duración de las presiones al pedal y en sus resultados encontró que estas presiones ocurrieron durante más tiempo en el primer grupo respecto al segundo. El autor interpretó estos resultados como un fenómeno de reforzamiento condicionado, es decir, que las presiones al pedal en el primer grupo se mantuvieron debido a que los estímulos discriminativos adquirieron propiedades reforzantes después de ser emparejados con el reforzador (Wyckoff, 1952, 1969). Al procedimiento que Wyckoff utilizó en su primer grupo de sujetos se le conoce como procedimiento de RO (Dinsmoor, 1983). El procedimiento de RO permitió por primera vez estudiar de manera separada las funciones reforzantes de los estímulos discriminativos en un programa de reforzamiento, por lo tanto, el procedimiento de observación es relevante y

ampliamente utilizado en los estudios de reforzamiento condicionado (Case & Fantino, 1981; DeFulio & Hackenberg, 2008; Dinsmoor, 1983; Escobar & Bruner, 2009; Kendall & Gibson, 1965; Lieving et al. 2006; Shahan & Cunningham, 2015; Williams, 1994). El procedimiento de observación también puede ser importante para el estudio del control de estímulos debido a que las RO pueden interpretarse como respuestas que hacen que el organismo haga contacto o “atienda” estímulos discriminativos (Kelleher, 1958). El procedimiento de observación tuvo algunos problemas por las características del procedimiento de Wyckoff. Kelleher (1958), por ejemplo, señaló que podía haber una desventaja en utilizar la duración como medida de las RO, debido a que Wyckoff (1952, 1969) obtuvo una duración alta en estas respuestas incluso cuando se observaba una discriminación de estímulos pobre, lo que podía ser una razón para cuestionar el supuesto papel reforzante de los estímulos sobre las RO. Para Kelleher la frecuencia podía ser una mejor medida y facilitaría la programación de la presentación de los estímulos contingente a las RO, por lo que realizó un estudio con chimpancés para replicar el experimento de Wyckoff (1952, 1969) usando la frecuencia como medida de las RO. Los chimpancés se encontraron en un programa mixto en el que podían presionar una tecla que resultaba en la entrega de reforzadores (operando de reforzamiento), o presionar una tecla distinta que resultaba en la presentación de estímulos discriminativos durante un periodo de tiempo determinado (operando de observación). Los resultados del experimento replicaron los resultados de Wyckoff, indicando que el control de estímulos podía establecerse a través de las RO y que estas respuestas podían mantenerse con la presentación de los estímulos discriminativos (Kelleher, 1958).

Posteriormente, Hirota (1972) realizó una serie de experimentos con palomas como sujetos y con un arreglo idéntico al de la caja de Wyckoff (1952, 1969) para estudiar la posibilidad de que las RO en el experimento de Wyckoff pudieran explicarse en términos de

algún artefacto del procedimiento. Hirota mencionó que a pesar de que en los programas mixtos no hay estímulos correlacionados con los componentes del programa, también está presente cierto tipo de estimulación a la que llamó “estímulo ambiguo” o “estímulo del programa mixto”. Según esta idea, el estímulo del programa mixto tendría propiedades aversivas por su “ambigüedad”, es decir, por no señalar la disponibilidad del reforzador, y de ser así, las RO podrían estar mantenidas tanto por la producción de los estímulos discriminativos, como por la eliminación del estímulo del programa mixto. En ese sentido, Hirota diseñó sus experimentos para evaluar por separado el supuesto efecto reforzante de la eliminación del estímulo del programa mixto. En una de las fases de uno de sus experimentos, Hirota expuso a sus sujetos experimentales a un programa con tres condiciones, cada una señalada con una luz de color diferente en la tecla, y que alternaban a lo largo de la sesión. Dentro de cada condición ocurría un programa mixto IF 30 s extinción 30 s. Las luces de colores en la tecla correspondientes a cada condición fueron la manera de hacer explícito el estímulo del programa mixto. Bajo la primera condición, las presiones al pedal funcionaban como RO produciendo los estímulos discriminativos (líneas verticales u horizontales) según el componente vigente, reforzamiento o extinción. En la segunda condición, las presiones al pedal apagaban la luz de la tecla, es decir, eliminaban el estímulo del programa mixto. En la tercera condición, las presiones al pedal no tenían ninguna consecuencia. Hirota hipotetizó que, si las presiones al pedal eran mantenidas también por la eliminación del estímulo del programa mixto, entonces la frecuencia de presiones al pedal en la segunda condición debería ser menor a cuando las presiones en el pedal presentaban los estímulos discriminativos (primera condición), pero mayores a cuando las presiones al pedal no tenían consecuencias (tercera condición). El autor no encontró ninguna diferencia en las presiones al pedal entre las condiciones, ni siquiera entre la condición donde las

presiones al pedal producían estímulos discriminativos y la condición donde las presiones al pedal no tenían consecuencias. El autor argumentó que probablemente las RO que observó Wyckoff (1952, 1969) no se debieron a la producción de los estímulos discriminativos ni a la eliminación del estímulo del programa mixto, sino a que el pedal donde ocurrían las RO se encontraba debajo de la tecla en donde los picotazos se reforzaban. Es decir, era posible que las palomas presionaran el pedal al acercarse a la tecla, y no por las razones que había argumentado Wyckoff. Hirota sugirió para futuras investigaciones con procedimientos de observación separar espacialmente el operando de observación del operando de reforzamiento, o utilizar un operando distinto como otra tecla en lugar del pedal, tal como ya había hecho Kelleher (1958). Sin embargo, esto introdujo el problema de que las RO y las respuestas por reforzador no puedan ocurrir simultáneamente y, al contrario, puedan competir entre sí.

Dos de los problemas señalados por Kelleher (1958) y Hirota (1972) se solucionaron con las mismas propuestas de estos autores, que fueron 1) medir la frecuencia de las RO en lugar de su duración, y 2) separar espacialmente el operando de observación y el operando de reforzamiento. Estas soluciones junto con la descripción de los programas de reforzamiento múltiples y mixtos en el libro de Ferster y Skinner (1957) resultó en la estandarización del procedimiento de observación que se utiliza desde hace varias décadas y que consiste en usar dos teclas o dos palancas, una de observación y una de comida y utilizar un programa de reforzamiento mixto con un componente de extinción y uno de reforzamiento que las respuestas de observación convierten temporalmente en un programa múltiple. La incompatibilidad de las RO y las respuestas por reforzador persiste en los procedimientos de observación y podría seguir siendo un problema, sin embargo, numerosas investigaciones en el área han obtenido hallazgos muy similares a los de Wyckoff (1952, 1969) usando teclas o palancas como operandos y

estímulos que se presentan con una duración fija de manera contingente a las RO (véase Dinsmoor, 1983). Así, el procedimiento de observación sigue siendo útil para el estudio del reforzamiento condicionado.

En un procedimiento de RO, existen dos respuestas ocurriendo de manera concurrente en un programa mixto de dos componentes. Las respuestas se registran en dos operandos diferentes, uno de observación y uno de reforzamiento. Las RO, emitidas en el operando de observación, tienen como consecuencia la presentación de los estímulos asociados con los componentes del programa de reforzamiento, por ejemplo, durante 5 s. El estímulo positivo (E+), señala la vigencia del componente de reforzamiento y el estímulo negativo (E-), señala la vigencia del componente de extinción. Dado este procedimiento es que se señala que las RO convierten temporalmente al programa mixto en un programa múltiple. El establecimiento y posterior mantenimiento de las RO podría tomarse como evidencia de los efectos de los estímulos como reforzadores condicionados, pues se mantienen a pesar de que sólo producen estímulos y no alteran la disponibilidad del reforzador. Las respuestas por reforzador, emitidas en el operando de reforzamiento, tienen como consecuencia la entrega del reforzador según el criterio del programa de reforzamiento vigente. Las respuestas por reforzador revelan la función discriminativa de los estímulos, pues se observa que los organismos emiten estas respuestas diferencialmente en presencia del E+ y el E-, producidos por las RO (Dinsmoor, 1983; Escobar, 2011; Fernández et al., 2015; Hendry & Dillow, 1966).

La utilidad del procedimiento de RO para estudiar tanto efectos de reforzamiento condicionado como de control de estímulos en animales no humanos, sin embargo, está limitada por numerosas dificultades asociadas con establecer las RO (véase e.g., Fernández et al., 2015). Un problema es que el entrenamiento no sólo involucra moldear las dos respuestas, sino que

parece ser necesario establecer al E+ como estímulo discriminativo y reforzador condicionado para que aumente la frecuencia de las RO. Esto mientras el E- adquiere una función como estímulo delta, el cual podría tener una función aversiva (Kelleher, 1958; Shahan, 2002). Asimismo, el procedimiento de observación requiere que las RO se emitan de manera concurrente a las respuestas por reforzador, con las que compite por ocurrir en un operando separado. Estas dificultades han resultado en que el entrenamiento de las RO se organice en diferentes fases sucesivas que pueden dividirse en tres (Fernández et al, 2015; Villegas & Bruner, 2008a).

La primera fase consiste en establecer las respuestas en los operandos, lo cual involucra entrenamiento de aproximación a la fuente de entrega del reforzador, moldeamiento por aproximaciones sucesivas a la respuesta a los operandos, y el mantenimiento de las respuestas con reforzamiento intermitente. Algunos autores han moldeado las respuestas solamente en el operando de reforzamiento (e.g. DeFulio & Hackenberg, 2008; Escobar & Bruner 2009; Fernández et al., 2015; Kendall, 1965; Preston, 1985; Shahan, 2002; Villegas & Bruner, 2008a). Otro caso es el de Lieberman, (1972), quien moldeó solamente las respuestas en el operando de observación. Otros autores, en cambio, han moldeado las respuestas en ambos operandos (e.g. Dinsmoor et al. 1981; Kelleher, 1958; Raiff, 2008).

La segunda fase del entrenamiento consiste en establecer al E+ y al E- como estímulos discriminativos, lo que regularmente se hace exponiendo a los sujetos a un programa múltiple con dos componentes, uno de reforzamiento y uno de extinción. Cuando los sujetos han tenido historia experimental, también se ha pasado directamente al programa múltiple sin necesidad del entrenamiento de las respuestas a los operandos (Case, 1981). Todo el entrenamiento que se realiza antes de la tercera fase puede verse como entrenamiento preliminar, dado que ninguna

respuesta tiene como consecuencia programada la presentación de los estímulos (Fernández et al., 2015; Villegas & Bruner, 2008a). Es importante señalar que en esta fase un índice de discriminación relativamente alto (aunque no existe un valor recomendado) es necesario para aumentar la probabilidad de establecer las RO en la siguiente fase.

La tercera y última fase es propiamente el procedimiento de observación, en el que se expone a los sujetos a un programa mixto con componentes idénticos al programa múltiple de la segunda fase. Es en este punto en el que las RO convierten temporalmente el programa mixto en uno múltiple (RO). Algunos autores han omitido el programa múltiple de la segunda fase pasando directamente al procedimiento de observación de la tercera fase (Dinsmoor et al., 1981; Fernández et al., 2015; Gaynor & Shull, 2002; Villegas & Bruner, 2008a) aunque esto no siempre resulta en el establecimiento de RO en animales sin historia experimental en procedimientos similares o en otro tipo de procedimientos.

Pese a que los entrenamientos usados hasta ahora han permitido establecer con mayor o menor éxito RO y que existen similitudes entre ellos, estos no han sido consistentes a través de los diversos estudios. No todos los autores moldean las respuestas a los mismos operandos ni de la misma manera; algunos incluyen las tres fases del entrenamiento, pero otros omiten la primera fase, la segunda, o ambas; y también hay casos en los que los sujetos tienen una historia experimental, lo que, en suma, hace complicado conocer con precisión cuáles han sido las variables determinantes para obtener RO de forma efectiva en los sujetos experimentales (Villegas & Bruner, 2008a). Además, la mayoría de los estudios no reporta el tiempo requerido para el entrenamiento de las RO y solo algunos autores señalaron que fueron necesarias entre 31 y 70 sesiones sólo para el entrenamiento preliminar, lo que podría considerarse un periodo de tiempo relativamente largo (Lieberman, 1972; McMillan, 1974). Incluso en algunos estudios los

autores han señalado que a pesar de que se usen las tres fases de entrenamiento las RO no se establecen en todos los sujetos (Fernández et al., 2015; Slezak & Anderson, 2014).

A partir de las irregularidades en los procedimientos para establecer RO, Villegas y Bruner (2008a) sugirieron una aproximación que proponía simplificar el entrenamiento. Los autores argumentaron que probablemente un procedimiento de entrenamiento preliminar no es necesario para establecer las RO, por lo que realizaron un par de experimentos que tuvieron como objetivo evaluar el papel del entrenamiento preliminar en los procedimientos de observación. Su primer experimento consistió en exponer a sus sujetos experimentales a un procedimiento de RO sin ningún entrenamiento preliminar (e.g. entrenamiento de aproximación a la fuente de reforzamiento o moldeamiento de la respuesta a los operandos). Para ello, utilizaron a tres ratas macho experimentalmente ingenuas, una caja de condicionamiento operante con dos palancas en el panel frontal, una luz arriba de cada palanca y un generador de tonos. El procedimiento consistió en exponer a los sujetos a un programa mixto intervalo al azar (IA) 8 s ( $t = 2$  s,  $p = 0.25$ ) extinción durante 80 sesiones. En cada sesión se presentaban 30 componentes de reforzamiento con una duración de 32 segundos y 30 componentes de extinción con una duración de 64 segundos que alternaban al azar con la restricción de que ningún componente se presentara más de dos veces consecutivas. Las presiones a la palanca izquierda producían la entrega de comida en presencia del componente de reforzamiento IA 8 s. De manera concurrente, cada presión a la palanca derecha exponía a los sujetos a los estímulos discriminativos (E+ y E-) durante seis segundos. El E+ consistió en el apagado de la luz general y el encendido de las luces arriba de cada palanca con una alternación de un segundo. El E- consistió en la emisión constante del tono.

Villegas y Bruner (2008a) reportaron que las RO ocurrieron únicamente en una de las tres ratas debido a que sólo en el caso de esa rata encontraron un aumento en la frecuencia de RO y diferencias en las tasas de respuestas por comida entre el componente de reforzamiento y el componente de extinción que se esperaría si el programa mixto funciona como un múltiple al presentarse el E+ y el E-. La interpretación de estos hallazgos fue que la adquisición de las RO sin entrenamiento preliminar es posible, pero no sucede de manera homogénea en los sujetos. Observaron, además, una variación sistemática entre las respuestas por comida, la frecuencia de reforzamiento y las RO de los sujetos, por lo que sugirieron que el establecimiento de las RO podría depender de una ocurrencia consistente de las respuestas por comida. Esto es, que el aumento en la frecuencia de reforzamiento, causado por las respuestas por comida, aumentaría también la probabilidad de que los estímulos presentados por las RO se emparejaran con el reforzador primario, de tal forma que los estímulos discriminativos pudieran adquirir propiedades reforzantes y mantener a las RO por sí mismos. Por último, los autores notaron que la frecuencia de reforzamiento en su experimento pudo haber sido afectada por la duración de los intervalos entre reforzadores. Debido a que era posible que los componentes de extinción se presentaran dos veces consecutivas, pudieron haberse presentado intervalos entre reforzadores de hasta 128 segundos, dificultando el establecimiento de las respuestas por comida y las RO.

Para evaluar el posible efecto de los intervalos entre reforzadores sobre la adquisición de RO, los investigadores diseñaron un segundo experimento en el que expusieron a dos grupos de tres ratas a entrenamientos preliminares con diferentes frecuencias de reforzamiento, uno con valores cercanos al reforzamiento continuo, y el otro con intervalos entre reforzadores cercanos a 60 s. El entrenamiento preliminar para ambos grupos fue sin entrenar la aproximación a la fuente de reforzamiento ni moldear las presiones a la palanca. Uno de los grupos estuvo expuesto a un

programa IA 6 s ( $t = 1.5$  s,  $p = 0.25$ ) y el otro a un programa IA 52 s ( $t = 5.2$  s,  $p = 0.1$ ), en los cuales la palanca izquierda producía la entrega de reforzadores, y las presiones a la palanca derecha sólo eran registradas. Esta condición estuvo vigente durante 10 sesiones. Posteriormente, ambos grupos de ratas fueron colocados en un programa IA 8 s extinción 64 s idéntico al del primer experimento durante 60 sesiones. Los autores encontraron RO en todos los sujetos experimentales con diferencias en el número de sesiones en el que se establecían, dependiendo del entrenamiento preliminar en el que se encontraron anteriormente. En este experimento, los autores determinaron que las RO se establecían si las presiones en la palanca de reforzamiento obtenían un valor mayor a 0.5 en el índice de discriminación (ID), que resulta de la división de la tasa de respuesta por comida en el componente de reforzamiento entre la tasa de respuesta por comida global. Los autores utilizaron el ID porque se utiliza como medida del control de estímulos. Villegas y Bruner observaron que las ratas que estuvieron expuestas al programa IA 6 s, con una frecuencia de reforzamiento más alta, no tardaron más de 6 sesiones en emitir RO. En cambio, la adquisición de RO tardó más en el grupo que estuvo en el programa IA 52 s, el cual tenía una frecuencia de reforzamiento más baja. En este grupo una de las ratas tardó hasta 36 sesiones para cumplir con el criterio del ID.

Los investigadores concluyeron que una frecuencia de reforzamiento cercana al reforzamiento continuo para las respuestas por el reforzador primario es suficiente para establecer RO de manera exitosa y en seis sesiones en promedio, contrario a una frecuencia de reforzamiento baja. Villegas y Bruner (2008a; Experimento 2) reportaron que la cantidad mínima de reforzadores entregada a un sujeto durante el entrenamiento preliminar fue de 13, por lo que sugirieron que las respuestas por comida y las RO pueden establecerse con relativamente pocos reforzadores.

Villegas y Bruner (2008b) realizaron un segundo estudio para observar el efecto del número de reforzadores entregados en el entrenamiento preliminar sobre la adquisición de las RO. En este estudio se expuso a cuatro grupos de 3 ratas a un entrenamiento preliminar que consistía en un programa IA 4 s ( $t = 1$  s,  $p = 0.25$ ), el cual finalizaba cuando los sujetos obtuvieran la cantidad de reforzadores establecida para su grupo (24, 8, 4 o 2). Para ninguno de los sujetos se entrenó la aproximación al comedero ni se hizo moldeamiento de la respuesta a los operandos. Una vez finalizado el entrenamiento preliminar, todos los sujetos se colocaron durante 30 sesiones en un procedimiento de respuestas de observación IA 8 s extinción idéntico al que utilizaron en el estudio anterior (Villegas & Bruner, 2008a; Experimento 1).

Los autores encontraron que un mayor número de reforzadores facilita la adquisición de RO, pues a excepción de dos sujetos del grupo que recibió dos reforzadores en el entrenamiento preliminar, se observó adquisición de RO en la mayoría de los sujetos. Con los datos obtenidos en su segundo estudio, Villegas y Bruner (2008b) concluyeron nuevamente que no es necesario un entrenamiento explícito de las RO, y que incluso reforzar solamente cuatro veces las respuestas por comida era suficiente para observar eventualmente la adquisición de RO. Además, el tiempo invertido para el entrenamiento y finalmente observar RO en los grupos con mejores resultados (24 y 8 reforzamientos) era menor a 15 sesiones. En ambos estudios, Villegas y Bruner (2008a, 2008b), destacaron la importancia del establecimiento de tasas consistentes de las respuestas por reforzador, argumentando que al lograrlo se favorece la ocurrencia de relaciones temporales entre el reforzador primario y los estímulos producidos por las RO. De este modo, los estudios de estos autores aportaban un procedimiento breve para la adquisición de RO e información sobre las variables determinantes para su establecimiento. Sin embargo, los resultados de otros experimentos sobre RO realizados con ratas como sujetos (Escobar, 2010;

Fernández et al. 2015; Gaynor & Shull, 2002) sugieren que los hallazgos de Villegas y Bruner podrían no ser generales.

Fernández et al. (2015) se interesaron en estudiar las relaciones temporales entre los estímulos y el reforzador en los procedimientos de observación. Una manera de hacerlo era manipular la duración de los estímulos producidos al responder en el operando de observación. Para su estudio, Fernández et al. diseñaron un procedimiento muy similar al de Villegas y Bruner (2008a) en el que utilizaron a doce ratas experimentalmente ingenuas. El arreglo de las cámaras experimentales era prácticamente igual al de Villegas y Bruner: dos palancas con una luz arriba de cada una, un generador de tonos, y una luz de iluminación general. La única diferencia notable era que Fernández et al. utilizaron palancas retráctiles.

Las ratas estuvieron en alimentación libre, privadas de agua (la cual fue usada como reforzador) durante 23 horas y 30 minutos, y al término de cada sesión tenían acceso a agua durante 30 minutos. Para el entrenamiento preliminar primero se introdujo sólo la palanca izquierda, que se desplegada durante 15 segundos y dentro de ese lapso se reforzaba cualquier presión. Transcurridos los 15 segundos, la palanca se retraía y se entregaba una dosis de agua, a partir de lo cual transcurrían otros 60 segundos hasta que la palanca volviera a desplegarse. Las ratas estuvieron en esta condición hasta que ocurrieran tres sesiones consecutivas con 80 reforzadores entregados. Después se introdujo la palanca derecha sin ninguna consecuencia programada, la palanca izquierda permaneció siempre desplegada y se reforzaron todas las respuestas en ella. Esta segunda condición se mantuvo hasta que ocurrieran tres días consecutivos con entrega de 100 reforzadores. Por último, se expuso a las ratas a un programa IA 6 s ( $t = 1.5$ ,  $p = 0.25$ ) durante 10 sesiones más.

Terminado el entrenamiento preliminar, los investigadores formaron tres grupos de cuatro ratas cada uno. Con base en el desempeño de las ratas en el programa IA 6 s y de la inspección visual de los datos, los investigadores se aseguraron de que en cada uno de los grupos hubiera ratas con tasas de respuesta altas, intermedias y bajas. Conformar a los tres grupos de esta manera, les permitió controlar las tasas de respuestas por reforzador en todos los grupos y, por lo tanto, también la frecuencia de reforzamiento (variables que Villegas y Bruner, 2008a; 2008b, destacaban como determinantes en el establecimiento de RO). Esta distribución de los sujetos ayudaría a observar de manera más clara la manipulación de su variable independiente.

Una vez formados los grupos, cada uno fue colocado en un procedimiento de RO muy similar al de Villegas y Bruner (2008a), con diferencia en la duración de los estímulos producidos por las RO. Las duraciones de los estímulos fueron 0.5 s, 6 s y 10 s. Los sujetos estuvieron en el programa correspondiente a su grupo durante 40 sesiones. En sus resultados, los autores encontraron la adquisición de RO solamente en cuatro sujetos, dos de ellos parte del grupo donde los estímulos duraban 10 s, y los otros dos del grupo en donde los estímulos duraban 6 s. Los autores analizaron la relación de contigüidad entre el E+ y el reforzador primario, y encontraron que los sujetos en los que fue posible observar establecimiento de RO estuvieron expuestos a más relaciones temporales en las que la entrega del reforzador primario ocurría en presencia del E+ o poco después de que el E+ se apagara, respecto a los sujetos en los que no se observó adquisición de RO. Fernández et al. (2015) concluyeron que una mayor duración de los estímulos producidos favorecía la adquisición de RO debido a que duraciones largas de los estímulos hacían más probable la relación de contigüidad entre el E+ y el reforzador primario, lo cual eventualmente haría que el E+ adquiriera propiedades reforzantes.

Un hecho a destacar es que el experimento de Fernández et al. (2015) no replicó los resultados de Villegas y Bruner (2008a), pese a que dos grupos, uno de cada estudio, estuvieron bajo casi los mismos parámetros (IA 8 s extinción 64 s con 5 o 6 s de duración de los estímulos). Además, las tasas de RO observadas fueron menores en el estudio de Fernández et al. respecto a las del estudio de Villegas y Bruner. En su discusión, Fernández et al. mencionaron que las tasas de respuesta por reforzador y, por lo tanto, la frecuencia de reforzamiento, eran en principio las mismas en todos los grupos debido a que cada grupo tenía sujetos con tasas altas, intermedias y bajas de respuestas por reforzador. Por lo tanto, no parecía que la frecuencia de reforzamiento o la tasa de respuesta por reforzador hicieran alguna diferencia en la adquisición de RO. De hecho, tres de las ratas en las que fue posible encontrar adquisición de RO no tenían una tasa más alta que el resto al final del entrenamiento preliminar. El único factor común entre ellas parecía ser la relación de contigüidad entre el E+ y el reforzador primario. Los autores sugirieron que es probable que el establecimiento de las RO sin entrenamiento previo dependa en gran medida de la proximidad temporal accidental entre el E+ y el reforzador primario. Esta relación accidental podría explicar la falta de homogeneidad en el establecimiento de las RO en la mayoría de los estudios. A partir de una comunicación personal, T. Villegas (comunicación personal, Mayo 24, 2008), pudo determinarse que la selección de los sujetos experimentales en los estudios de Villegas y Bruner (2008a, 2008b) se realizó a partir de la tasa de respuesta observada antes de iniciar el experimento. Esto sugiere un sesgo importante en ambos estudios donde se reportó establecimiento de RO en todos los sujetos.

El hecho de que los resultados de Villegas y Bruner (2008a, 2008b) no se replicaron con un procedimiento prácticamente idéntico, y que al momento no se encuentre ningún otro estudio que haya podido replicarlos, plantea de nuevo la pregunta sobre cómo tendría que ser un

entrenamiento para establecer RO consistentemente entre sujetos. Los datos obtenidos por Fernández et al. respecto al papel de las relaciones temporales entre el E+ y el reforzador para que el E+ adquiriera propiedades reforzantes, brindan información sobre las variables que deberían considerarse.

Una opción que podría facilitar el entrenamiento de las RO es utilizar un programa ajustable. En los programas ajustables, los valores o parámetros de una variable se modifican en función del desempeño de los organismos en el programa. A su vez, el desempeño de los organismos está influenciado por el contacto que tienen con los estímulos o condiciones presentes. Originalmente, este tipo de procedimientos fueron diseñados para que los sujetos experimentales regularan con su conducta la cantidad de estimulación que recibían, o para estudiar fenómenos como los umbrales sensoriales en el campo de la psicofísica (Békèsy, 1947; Evans, 1963; Oldfield, 1949). En el análisis de la conducta, estos procedimientos han sido útiles para estudiar límites en la duración de demoras de reforzamiento (García-Leal et al., 2019; Mazur, 1988), de intervalos de retención en procedimientos de igualación a la muestra (Colares et al., 2020; Kangas et al., 2010; Kangas & Branch, 2012), y puntos de corte en estudios de elección con diferentes demoras y magnitudes del reforzamiento (Green et al., 2007). Utilizar un programa ajustable para el entrenamiento de las RO, permitiría garantizar un contacto constante del organismo con los estímulos discriminativos y la formación de relaciones de contigüidad del E+ con el reforzador, lo que podría facilitar el establecimiento de las RO. Por lo tanto, el propósito de este trabajo será probar la efectividad de un programa ajustable para el entrenamiento de RO en ratas como sujetos.

## Experimento 1

En la literatura de reforzamiento condicionado con procedimientos de observación, existe un debate respecto a las razones por las que el E+ tiene efectos reforzantes y cuál es el papel del E-. Uno de los hallazgos encontrados en la literatura, es que el E- por sí solo no funciona como reforzador condicionado y que incluso puede funcionar como un estímulo aversivo condicionado para las RO (Escobar & Bruner, 2009; Gaynor & Shull, 2002; Shahan & Cunningham, 2015). Bajo esta lógica, llevar a cabo el entrenamiento de las RO sólo con el E+ podría ser más conducente a su establecimiento, pues se evitaría el posible castigo accidental de las RO por la presentación del E-. Además, acorde con los hallazgos de Fernández et al. (2015), sólo el E+ es el que adquiriría funciones de reforzador condicionado. Por lo tanto, en este primer experimento las RO sólo producían el E+.

Para que el procedimiento ajustable funcione como tal, los sujetos deberán comenzar en un punto donde las variables del programa tengan valores determinados que se modifiquen gradualmente en función del desempeño de los sujetos, hasta llegar a un punto donde el programa sea similar al de un procedimiento de RO (i.e. programa mixto con dos componentes, uno de reforzamiento y uno de extinción). De esta forma, habría un continuo de valores en las variables del programa desde el punto inicial hasta el procedimiento de RO. Con base en los resultados de Fernández et al. (2015), establecer una relación de contigüidad entre el E+ y el reforzador desde el punto inicial del programa ajustable favorecería el establecimiento y futuro mantenimiento de las RO. Una manera de hacerlo podría ser elaborar un programa ajustable en donde la duración del componente de extinción y la probabilidad de reforzamiento en presencia del E+ se modifiquen gradualmente en un continuo. El componente de extinción se modificaría desde una duración de 0 s a una duración de 60 s. Por otro lado, la probabilidad de reforzamiento

en presencia del E+ empezaría con un valor de uno y se acercaría gradualmente al valor de cero. Así, la entrega del reforzador ocurriría sólo si el E+ producido por una respuesta está presente, es decir, que sea necesario emitir una respuesta en el operando de observación, e inmediatamente después emitir respuestas en el operando de reforzamiento mientras el E+ esté activo para que se produzca un reforzador. Esto garantizaría relaciones temporales entre el E+ y el reforzador en cada entrega del reforzador, sobre todo en el punto inicial del programa ajustable, donde el componente de extinción tendría una duración de 0 s.

Para que el programa cambie en el continuo de valores, la duración del componente de extinción deberá aumentar al mismo tiempo que disminuya la probabilidad de la entrega del reforzador en presencia del E+, acercándose al procedimiento de RO tradicional. Estos cambios dependerían del desempeño de los sujetos, el cual se determina por la emisión de respuestas y la obtención de reforzadores. Si la conducta de los sujetos deja de producir reforzadores, el programa deberá moverse a valores anteriores, acercándose al punto inicial del programa ajustable. De esta forma, se podría asegurar una relación de contigüidad constante entre el E+ y el reforzador, una emisión consistente de respuestas en ambos operandos y el control de estímulos sobre las respuestas por reforzador.

El programa ajustable en su punto final tendría características similares a las de un programa mixto de dos componentes, uno de reforzamiento y uno de extinción, en el que las respuestas al operando de observación produzcan el E+ durante el componente de reforzamiento y las respuestas al operando de reforzamiento produzca la entrega de reforzador según el programa de reforzamiento determinado. Una vez alcanzado el punto final del programa ajustable, exponer a los sujetos a una fase de prueba que consista en un procedimiento de observación en el que los valores de las variables del programa se mantengan constantes,

mostraría la efectividad del programa ajustable para el entrenamiento de las RO. Por lo tanto, el propósito de este primer experimento fue probar si un programa ajustable empleando únicamente el E+ es efectivo para establecer RO.

## **Método**

### *Sujetos*

Cuatro ratas macho Wistar experimentalmente ingenuas de tres meses de edad al inicio del experimento se utilizaron como sujetos. Cada una de las ratas se mantuvo en su propia caja habitación durante el tiempo fuera de la sesión con acceso libre a comida. Al término de cada sesión tuvieron acceso a agua durante una hora, y estuvieron privadas durante el resto del día.

### *Aparatos*

Se utilizó una cámara de condicionamiento operante de 30 x 23 x 20.5 cm que se fabricó para la presente tesis. Los paneles frontal y trasero se imprimieron en 3D con material PLA y los otros dos paneles (laterales) se construyeron con acrílico (véase Escobar et al., 2022). En la parte inferior central del panel frontal se encontraba una charola metálica dentro de un espacio de 50 x 50 mm. A cada lado de la charola se encontraban dos palancas de 4 x 4.7 cm ubicadas a 61 mm de distancia del suelo, cada una con un interruptor Honeywell modelo 311SM4-T, que necesitaba una fuerza mínima de 0.15 N para activarse. Arriba de cada palanca, se encontraba un LED que emitía luz blanca con una cubierta traslúcida de 29 mm de diámetro a una distancia de 40 mm desde cada palanca a los bordes de las cubiertas traslúcidas. En la parte exterior de la caja, se fijó un sonalert que emitía un tono constante a 70 dB y 2900 Hz, los dB se regularon a través de un potenciómetro. En la parte posterior del panel se colocó una bomba peristáltica que dejaba caer 0.04 ml de agua dentro de la charola metálica en cada activación. En la parte superior central del

panel trasero, se encontraba un LED que emitía luz blanca con una cubierta traslúcida de 17 mm de diámetro para iluminación general.

Los eventos dentro de la caja experimental se controlaron con una laptop mediante una interfaz Arduino-Visual Basic descrita por Escobar y Pérez-Herrera (2015). Cada sesión tuvo lugar en un cubículo de madera sonoamortiguado. Se reprodujo un archivo mp3 con ruido blanco en una bocina a 60 dB para evitar la influencia de sonidos externos.

### ***Procedimiento***

El entrenamiento de las RO en el presente trabajo consistió en tres fases. La primera fase consistió en el entrenamiento de las respuestas a la palanca. En la segunda fase se expuso a los sujetos a un programa ajustable con el que se buscó que, en función del desempeño de los sujetos, el programa avanzara gradualmente de un punto A hasta un punto B. En el punto A sólo había un componente de reforzamiento y la probabilidad de entrega del reforzador en presencia del E+ era de uno. En el punto B, los sujetos se encontraban en un programa mixto en el que había un componente de reforzamiento, un componente de extinción, y la probabilidad de reforzamiento en presencia del E+ era de .1. La tercera fase consistió en un procedimiento de observación con un programa mixto similar al del programa ajustable en su punto final. Esta fase también se nombró fase de prueba dado que fue una manera de probar si el entrenamiento en el programa ajustable fue suficiente para que las RO se mantuvieran en un procedimiento de observación tradicional. Las sesiones de todas las fases comenzaron con el encendido de la luz general y ésta se apagaba cuando las sesiones terminaban.

**Establecimiento de las presiones a la palanca.** Para entrenar las presiones a la palanca, solamente la palanca izquierda se encontraba en el interior de la caja experimental. La palanca izquierda fue más adelante la palanca de reforzamiento. Se realizó un entrenamiento por

aproximaciones sucesivas de las respuestas de acercamiento a la fuente de reforzamiento (la charola metálica de la caja), y posteriormente se moldearon las presiones a la palanca igualmente por aproximaciones sucesivas. Los reforzadores consistieron en la entrega de 0.04 ml de agua. Posteriormente, para asegurar una tasa consistente de respuestas, se expuso a los sujetos a programas de reforzamiento cuyos requisitos de respuesta aumentaban. Se empezó con un programa de razón fija (RF) 1, seguido por un programa de razón variable (RV) 2, a partir del cual, el valor de razón aumentaba de dos en dos hasta llegar a un programa RV 10. Los sujetos pasaron al siguiente programa de RV (o de RF 1 a RV 2) si obtenían 30 reforzadores en dos sesiones consecutivas bajo el mismo valor de razón. Cada sesión terminaba después de la entrega de 30 reforzadores o transcurrido el tiempo de una hora. El criterio para que los sujetos pasaran a la segunda fase fue que obtuvieran 30 reforzadores en dos sesiones consecutivas bajo el programa RV 10.

**Programa ajustable.** En el programa ajustable se introdujo la palanca derecha a la cámara experimental, que funcionó como operando de observación. Se utilizó un programa ajustable que, con base en el desempeño de los sujetos, modificaba dos parámetros del programa dentro de un continuo de dos extremos: inicial, en el que se evitaron las respuestas en extinción y final en el que el procedimiento se aproximó a uno de RO convencional con un periodo de reforzamiento y uno de extinción. Uno de los parámetros que se ajustaban a partir del desempeño de los sujetos fue la duración del componente de extinción, cuyo valor inicial fue de 0 s y cuyo valor final fue de 60 s. El otro parámetro ajustable en el programa fue la probabilidad de reforzamiento dentro de un programa construido de manera similar a un IA. Un IA se genera con dos parámetros,  $t$  o duración del ciclo y  $p$  o probabilidad de reforzamiento para la primera respuesta en el ciclo (véase Schoenfeld & Cole, 1972). El programa inició con  $t = 0.5$  s  $p = 1$ , lo

cual hacía al programa funcionalmente similar a un programa de reforzamiento continuo. El valor de  $p$  se redujo en pasos de .01, hasta que llegó a 0.1, lo cual generó un IA de 5 s. Este programa se encontraba vigente en un componente de 30 s (componente de reforzamiento).

Con este arreglo, cuando el programa ajustable estaba en los valores iniciales de ambos parámetros, los sujetos se encontraban en un programa de reforzamiento continuo. En cambio, si los valores eran diferentes a los iniciales, un componente de extinción de  $n$  segundos alternaba semialeatoriamente con el componente de reforzamiento con la restricción de que ningún componente podía presentarse más de dos veces consecutivas. En el extremo con los valores finales de ambos parámetros, el programa ajustable era similar al de un procedimiento de observación tradicional (i.e. Un programa mixto IA 5 s EXT con componentes de 30 y 60 s, respectivamente). Las presiones en la palanca de observación durante el componente de reforzamiento encendían el E+ durante 5 s. El E+ que consistió en apagar la luz general, encender las luces colocadas arriba de cada palanca y activar un tono constante de 2900 Hz y 70 dB. Las presiones a la palanca de observación en el componente de extinción se registraron, pero no tuvieron consecuencias programadas. Si ocurría un cambio de componente de reforzamiento a extinción el E+ se apagaba.

El desempeño de los sujetos se determinó por la cantidad de reforzadores que producían las presiones a la palanca de reforzamiento. Los parámetros ajustables del programa cambiaban juntos hacia el extremo inicial o final según los siguientes criterios: 1) la entrega de tres reforzadores dentro de un periodo de seis componentes aumentaba cinco segundos la duración del componente de extinción y disminuía 0.1 la probabilidad de reforzamiento, lo que acercaba al programa ajustable hacia el extremo final. 2) Periodos de seis componentes sin entrega de reforzadores resultaban en una disminución de cinco segundos la duración del componente de

extinción y un aumento de 0.1 de la probabilidad de reforzamiento, lo que acercaba a los sujetos al extremo inicial del programa ajustable.

En la primera sesión del programa ajustable, el programa comenzaba con los valores iniciales para todos los sujetos (i.e. 0 s de duración del componente de extinción y probabilidad de 1 del programa IA 5 s). A partir de la segunda sesión, el programa comenzaba para cada sujeto con los valores con los que había terminado la sesión previa. Los ajustes del programa ocurrían durante cada sesión en función al desempeño de los sujetos. Cada sesión tuvo una duración de 50 minutos. Se registraron las presiones a ambas palancas en cada componente, así como el número de reforzadores obtenidos en toda la sesión.

Debido a que hay más valores de duración del componente de extinción de cero segundos a 60 s respecto a los valores de probabilidad de reforzamiento de uno a 0.1, era posible que el valor de probabilidad de reforzamiento alcanzara su valor final sin que la duración del componente de extinción alcanzara su valor final. Cuando esto sucedía, los siguientes ajustes hacia el extremo final del programa consistían en el aumento de la duración del componente de extinción, manteniendo constante el valor final de probabilidad de reforzamiento. El programa ajustable terminó después de 30 sesiones para todos los sujetos.

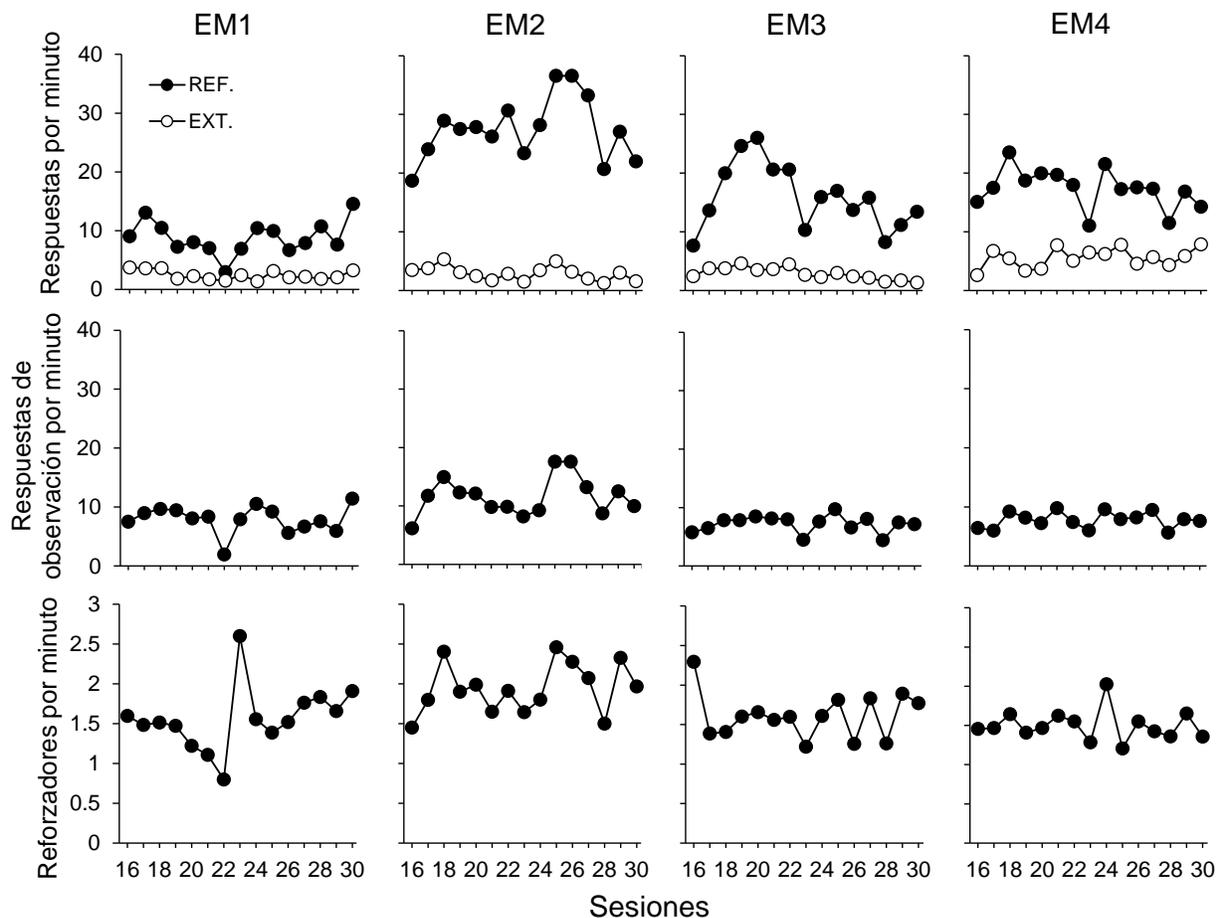
**Fase de prueba: Procedimiento de observación estándar.** El procedimiento de observación estándar consistió en un programa mixto IA 5 s ( $t = 1$  s,  $p = 0.2$ ) extinción. La duración del componente de reforzamiento fue de 30 s y el de extinción de 60 s. Treinta componentes de extinción y 30 de reforzamiento alternaban de manera semialeatoria con la condición de que ninguno se presentara más de dos veces consecutivas. Las presiones a la palanca de observación durante el componente de reforzamiento encendían las luces colocadas arriba de ambas palancas, apagaban la iluminación general y encendían el tono a 70 dB durante

cinco segundos. Las presiones a la palanca de reforzamiento durante el componente de reforzamiento producían la entrega del reforzador según el criterio del programa. Las presiones a ambas palancas en el componente de extinción se registraron, pero no tenían consecuencias programadas. La diferencia con el programa ajustable es que en la Fase de prueba todos los valores se mantenían constantes independientemente del desempeño de los sujetos y la entrega del reforzador no dependía de que el requisito de reforzamiento se cumpliera en presencia del E+. Cada sesión tuvo una duración total de 45 minutos. Se registraron las presiones de ambas palancas en cada componente, así como el número de reforzadores obtenidos. El criterio para finalizar esta fase fue de 15 sesiones.

## **Resultados**

### ***Programa ajustable***

La Figura 1 muestra la tasa de respuesta en la palanca de reforzamiento (panel superior), la tasa de RO (panel intermedio) y la frecuencia de reforzamiento (panel inferior) de todos los sujetos durante las últimas 15 sesiones del programa ajustable. Debido a un error en el registro de la duración de los componentes, no se muestran los datos de las primeras 15 sesiones del programa ajustable. Para las respuestas en la palanca de reforzamiento, los marcadores negros señalan las respuestas durante el componente de reforzamiento y los marcadores blancos señalan las respuestas durante el componente de extinción. En el caso de las RO se contó como RO solamente a las respuestas a la palanca de observación que encendieron el E+. La tasa de RO se calculó dividiendo el total de RO, entre el resultado del tiempo en que el componente de reforzamiento estuvo vigente menos el tiempo que el E+ estuvo encendido.



*Figura 1.* Tasa de respuesta por agua, de RO y frecuencia de reforzamiento en el programa ajustable. Tasa de respuesta en cada palanca y reforzadores por minuto a través de las sesiones para los cuatro sujetos en el programa ajustable del Experimento 1. Los marcadores negros y blancos corresponden a los componentes de reforzamiento y de extinción respectivamente.

La tasa de respuesta en la palanca de reforzamiento se mantuvo relativamente estable en ambos componentes para todos los sujetos. La tasa de respuesta en la palanca de reforzamiento fue mayor durante el componente de reforzamiento que en el componente de extinción durante todas las sesiones para todos los sujetos. Las RO de todas las ratas fueron relativamente estables a lo largo de las últimas 15 sesiones del programa ajustable, y se mantuvieron por encima de las

cinco respuestas por minuto en la mayoría de las sesiones. Igualmente, la frecuencia de reforzamiento fue relativamente estable para todos los sujetos.

***Fase de prueba: Procedimiento de observación estándar***

La Figura 2 muestra la tasa de respuesta en la palanca de reforzamiento (panel superior), la tasa de RO (panel intermedio) y la frecuencia de reforzamiento (panel inferior) de todos los sujetos durante las 15 sesiones del procedimiento de observación. La última sesión del sujeto EM4 se eliminó debido a un error en el procedimiento en esa sesión.

Las respuestas en la palanca de reforzamiento se mantuvieron relativamente estables en ambos componentes durante toda la fase de prueba para todos los sujetos. La tasa de respuesta en la palanca de reforzamiento fue más alta en el componente de reforzamiento que en el componente de extinción durante todas las sesiones para todos los sujetos. La tasa de respuesta por reforzador en el componente de reforzamiento tuvo una ligera tendencia ascendente en el sujeto EM2. Por otro lado, la tasa de RO se mantuvo relativamente estable para tres de los cuatro sujetos (EM1, EM3 y EM4) a lo largo de toda la fase de prueba. En el caso del sujeto EM2, la tasa de RO tuvo una tendencia descendente en las últimas tres sesiones, pero sin llegar a valores cercanos a cero. La frecuencia de reforzamiento fue relativamente estable para todos los sujetos en la fase de prueba.

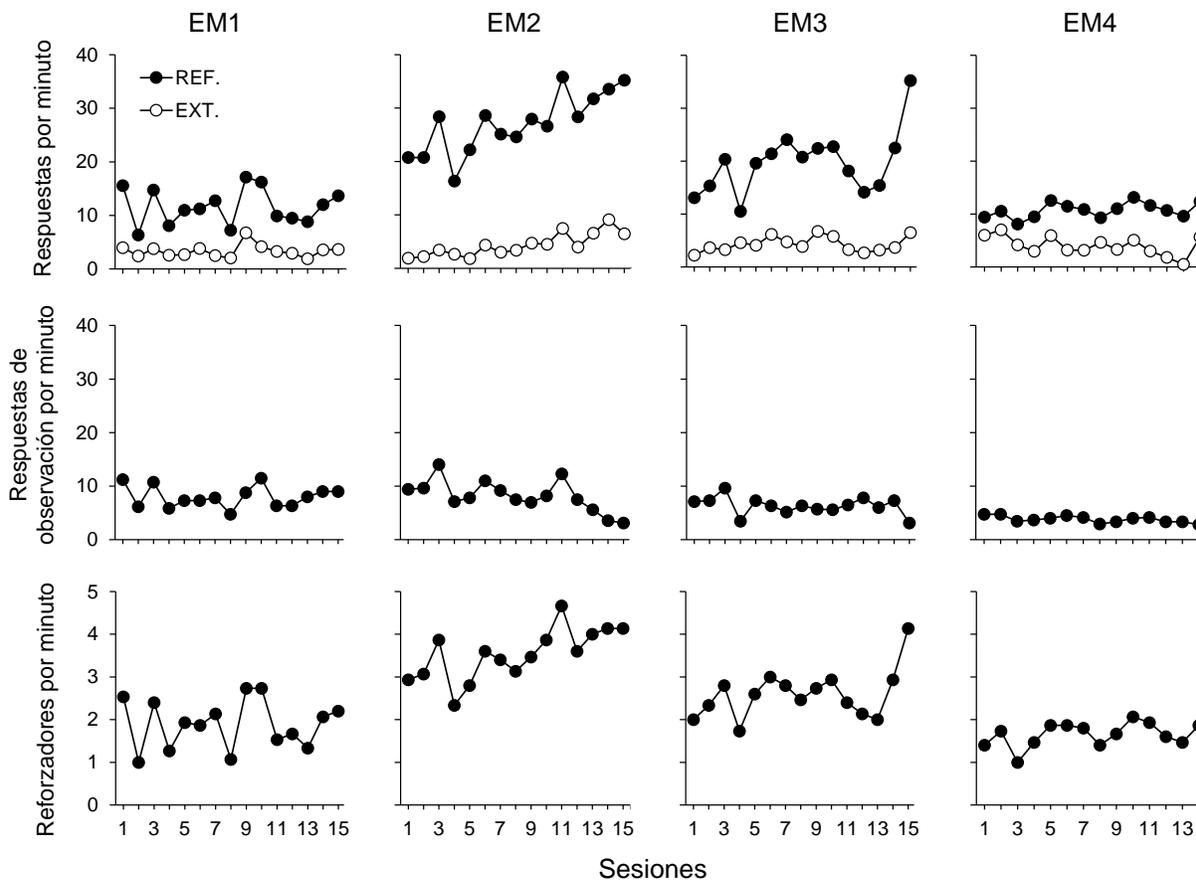


Figura 2. Tasa de respuesta por agua, de RO y frecuencia de reforzamiento en la fase de prueba.

Tasa de respuesta en cada palanca y reforzadores por minuto a través de las sesiones para los cuatro sujetos en el procedimiento de observación del Experimento 1. Los marcadores negros y blancos corresponden a los componentes de reforzamiento y de extinción respectivamente.

## Discusión

El propósito del primer experimento fue probar la efectividad de un programa ajustable para entrenar RO en ratas como sujetos. Tradicionalmente, después de entrenar a los sujetos a responder a el o los operandos de la cámara experimental, se expone a los sujetos a un programa múltiple para establecer control discriminativo de los estímulos en cada componente, o se expone a los sujetos directamente a un procedimiento de observación (programa mixto de dos componentes, reforzamiento y extinción). En este experimento, después del entrenamiento de

respuestas a la palanca, se expuso a los sujetos a un programa ajustable, que cambiaba durante las sesiones según el desempeño de los sujetos, lo que los llevaba gradualmente a un procedimiento de observación. Además, el programa ajustable aseguraba relaciones temporales constantes entre el E+ y el reforzador, lo que según los hallazgos de Fernández et al. (2015), podría ser determinante para el establecimiento de RO.

Los resultados del programa ajustable muestran que la tasa de respuesta en la palanca de reforzamiento durante el componente IA5 s fue mayor que en el componente de extinción desde la sesión 16 de esa condición en todos los sujetos, lo que es evidencia de control de estímulos. Las RO también ocurrían de manera consistente desde la sesión 16 para todos los sujetos. Sin embargo, en el programa ajustable, la entrega de los reforzadores ocurría solamente si el E+ se encontraba encendido, por lo que emitir RO era un requisito del programa y no era posible tener la certeza de que el E+ mantenía las RO.

En la fase de prueba se observó control de estímulos en todas las sesiones para todos los sujetos. En general, las tasas de RO se mantuvieron en los mismos niveles que en el programa ajustable para tres de los sujetos (EM1, EM2 y EM3). La tasa de RO de EM4 disminuyeron ligeramente en la fase de prueba en comparación con el programa ajustable, pero sin llegar a tasas cercanas a cero. Debido a que en la fase de prueba las RO sólo producían el E+ sin alterar la disponibilidad del reforzador, estos resultados pueden interpretarse en términos de reforzamiento condicionado, es decir, que las RO se mantenían por la presentación del E+.

En uno de los experimentos de Villegas y Bruner (2008a) y en el experimento de Fernández et al. (2015) hubo una condición en el que los sujetos estuvieron en un procedimiento de observación similar al del Experimento 1 de este trabajo. En el experimento de Villegas y Bruner (2008a), tres ratas tuvieron un entrenamiento preliminar en un programa mixto IA6 s ( $t =$

1.5 s,  $p = 0.25$ ) durante 10 sesiones de 30 minutos. Posteriormente se expuso a los sujetos a un procedimiento de observación en un programa mixto IA8 s ( $t = 1.5$  s,  $p = 0.25$ ) extinción 64 s, donde los componentes de reforzamiento duraban 32 s y las RO encendían los estímulos durante 6 s. En el experimento de Fernández et al. (2015), se expuso a cuatro ratas a un entrenamiento preliminar de moldeamiento de respuestas a la palanca, y después un programa IA6 s idéntico al de los sujetos de Villegas y Bruner. Luego se expuso a los sujetos a un procedimiento de observación que también fue igual al de los sujetos de Villegas y Bruner, con la única diferencia de que las RO encendían los estímulos durante 5 s.

Los resultados de Villegas y Bruner mostraron establecimiento de RO en todos los sujetos, reportaron tasas de RO cercanas a las dos respuestas por minuto en el componente de reforzamiento, y tasas de RO mayores a seis respuestas por minuto en el componente de extinción en todos los sujetos. Sin embargo, en los resultados de Fernández et al. se encontró que las RO se establecieron solamente en uno de los cuatro sujetos donde se replicó el entrenamiento y el procedimiento de observación de Villegas y Bruner. Además, las tasas de RO de los sujetos de Fernández et al. se mantuvieron cercanas a cero durante todas las sesiones para todos los sujetos. Los resultados de Villegas y Bruner no se replicaron con un procedimiento prácticamente idéntico, y dado que Fernández et al. controlaron la tasa de respuesta en la palanca de reforzamiento de sus sujetos para controlar también la frecuencia de reforzamiento, los autores descartaron que la frecuencia de reforzamiento fuera determinante para el establecimiento de RO. Por otro lado, Fernández et al. argumentaron que una variable determinante para entrenar RO exitosamente, podría ser la cantidad de emparejamientos de tipo pavloviano entre el E+ y el reforzador.

Los resultados del procedimiento de observación en el presente Experimento 1 muestran que las RO se mantuvieron consistentemente en todos los sujetos a lo largo de todas las sesiones. Las tasas de RO de todos los sujetos de este experimento fueron mayores a las reportadas por Villegas y Bruner (2008a) y Fernández et al. en los grupos de ratas bajo condiciones similares. También se observó control de estímulos sobre las respuestas en la palanca de reforzamiento durante todas las sesiones de los cuatro sujetos. Observar una emisión constante de RO y control de estímulos en las respuestas por reforzador son indicadores de que se establecieron RO. Por lo tanto, el uso de un programa ajustable como entrenamiento fue efectivo para el establecimiento de estas respuestas en todos los sujetos.

Las características del programa ajustable utilizado en el Experimento 1 de este trabajo consideraron los hallazgos de Fernández et al. (2015) respecto a las relaciones temporales entre el E+ y el reforzador. Dado que la entrega de los reforzadores en el programa ajustable siempre ocurría en presencia del E+, el programa garantizaba el emparejamiento de tipo pavloviano entre el E+ y el reforzador, lo que pudo haber facilitado el control discriminativo del E+, al mismo tiempo que adquiriría propiedades reforzantes. Estas características resultaron en el mantenimiento de RO en un procedimiento de observación y tasas más altas que las reportadas en otros estudios con parámetros similares. Los resultados de este primer experimento brindan evidencia que sustenta la idea de que una variable determinante para el entrenamiento de las RO es la relación temporal que se establece entre el E+ y el reforzador, tal como sugirieron Fernández et al. (2015).

En dos estudios diferentes, Villegas y Bruner (2008a, 2008b) argumentaron que la tasa de respuesta en la palanca de reforzamiento y la frecuencia de reforzamiento eran determinantes para la adquisición de RO. En específico, señalaron que mantener una tasa de respuesta

consistente en el operando de reforzamiento conduce eventualmente a la adquisición de RO, y que esto se logra con una tasa de reforzamiento cercana al reforzamiento continuo, aunque también puede lograrse reforzando sólo cuatro respuestas en el operando de reforzamiento. Fernández et al. reportaron una frecuencia de reforzamiento promedio de entre tres y cuatro reforzadores por minuto durante el entrenamiento preliminar en el grupo de ratas con condiciones similares al experimento de Villegas y Bruner (2008). También encontraron que tres de las cuatro ratas que mostraron adquisición RO no tenían una tasa de respuesta mayor al resto. Estos hallazgos llevaron a la conclusión de que la tasa de respuesta ni la frecuencia de reforzamiento son determinantes para el establecimiento de RO. En los resultados del Experimento 1 del presente trabajo se observó adquisición de RO a pesar de que la frecuencia de reforzamiento durante el programa ajustable (de 1 a 2.5 reforzadores por minuto) fue menor que la de los sujetos de Fernández et al. durante el entrenamiento preliminar. Las RO se establecieron incluso en el sujeto EM1, con la menor tasa de respuesta en la palanca de reforzamiento. Estos datos son contrarios a lo reportado por Villegas y Bruner (2008a, 2008b), pues las RO pueden establecerse con una baja tasa de reforzamiento y de respuestas en la palanca de reforzamiento si se mantiene una relación de contigüidad entre el E+ y el reforzador.

Uno de los problemas para el entrenamiento de las RO, es el periodo de tiempo relativamente largo que toma el entrenamiento preliminar, antes de exponer a los sujetos a un procedimiento de observación, que ronda entre las 31 y 71 sesiones en los reportes existentes (Lieberman, 1972; McMillan, 1974; Raiff, 2008). En el Experimento 1, se requirieron 43 sesiones en total para el entrenamiento preliminar (una sesión de entrenamiento a comedero y moldeamiento, 12 sesiones del programa RF1 al RV10 y 30 sesiones del programa ajustable). No obstante, para la sesión 16 del programa ajustable, ya ocurrían RO consistentemente y las

respuestas en la palanca de reforzamiento ya estaban diferenciadas, lo que podría sugerir que las RO podrían haberse establecido con menos sesiones del programa ajustable, reduciendo el tiempo de entrenamiento.

El Experimento 1 aportó información y un procedimiento efectivo para el entrenamiento de RO consistentes en todos los sujetos. El entrenamiento se realizó utilizando sólo un estímulo que señalaba el componente de reforzamiento (E+), lo que sirvió para conocer la importancia de su relación temporal con el reforzador.

## **Experimento 2**

Los estudios con procedimientos de observación utilizan tradicionalmente dos estímulos para señalar los componentes del programa mixto (E+ y E-) (Dinsmoor, 1983). Introducir el E- en el entrenamiento de RO con un programa ajustable, aunque pueda tener una función aversiva (Gaynor & Shull, 2002; Shahan & Cunningham, 2015), ayudaría a evaluar de manera directa la utilidad del programa ajustable en comparación con todos los trabajos anteriores en los que el E- está presente durante el entrenamiento.

Además, al contar con estímulos que señalen tanto el componente de reforzamiento como el componente de extinción del programa mixto, será posible evaluar la función discriminativa y reforzante de los estímulos de una forma más directa. Una forma de hacerlo, sería introducir una condición después del procedimiento de observación, en la que los estímulos de ambos componentes estén invertidos. Wyckoff (1952, 1969) realizó esta modificación en uno de sus experimentos con el mismo propósito. Los sujetos de su experimento mostraron tasas de respuestas diferenciadas en la tecla de reforzamiento en función del componente vigente (control de estímulos) y RO consistentes en el procedimiento de observación. Cuando Wyckoff invirtió los estímulos de ambos componentes, las respuestas en la

tecla de reforzamiento dejaron de estar diferenciadas y las RO disminuyeron. Finalmente, Wyckoff restauró los estímulos como estaban en un principio, y observó de nuevo tasas diferenciadas en la tecla de reforzamiento, así como emisión consistente de RO. Estos cambios probaron que los estímulos producidos por las RO tenían control discriminativo sobre las respuestas en la tecla de reforzamiento, y que también funcionaban como reforzadores condicionados para las RO.

Por lo tanto, el primer propósito de este segundo experimento será probar la efectividad del programa ajustable para entrenar RO con el E+ y el E-. El segundo propósito será probar la función discriminativa y reforzante de los estímulos realizando una manipulación similar a la de Wyckoff (1952, 1969).

## **Método**

### ***Sujetos***

Se utilizaron como sujetos a cuatro ratas macho Wistar experimentalmente ingenuas de tres meses de edad al inicio del experimento. Cada una de las ratas se mantuvo en su propia caja habitación durante el tiempo fuera de la sesión con acceso libre a comida. Al término de cada sesión tenían acceso a agua durante una hora, y estaban privadas durante el resto del día.

### ***Aparatos***

Se utilizó una cámara de condicionamiento operante de 30 x 23 x 20.5 cm idéntica al primer experimento. Se ubicaron dos palancas en el panel frontal, cada una con un interruptor, que necesitaban una fuerza mínima de 0.15 N para ser activados. Se encontraba un LED arriba de cada palanca que emitía una luz blanca. En la parte superior central al exterior de la caja, se fijó un sonalert que emitía un tono constante de 70 dB a 2900 Hz. En la parte inferior y centrado en el panel frontal, se encontraba un recipiente de agua. El agua se entregaba en cantidades de

0.1 ml a este recipiente por medio de una bomba peristáltica. En la parte superior central del panel trasero, se encontraba un LED que emitía luz blanca para iluminación general.

Los eventos dentro de la caja experimental se controlaron con una laptop mediante una interfaz Arduino-Visual Basic descrita por Escobar y Pérez-Herrera (2015). Cada sesión tuvo lugar en un cubículo de madera sonoamortiguado y durante las sesiones se reprodujo un archivo mp3 con ruido blanco en una bocina a 60dB para evitar la influencia de sonidos externos.

### ***Procedimiento***

**Establecimiento de las presiones a la palanca.** El entrenamiento de las presiones a la palanca fue igual que en el primer experimento. Se realizó un entrenamiento por aproximaciones sucesivas de las respuestas de aproximarse a la fuente de reforzamiento, moldeamiento por aproximaciones sucesivas de presiones a la palanca izquierda, y se mantuvieron las respuestas con programas de RV cuya razón aumentaba hasta llegar a un RV 10.

**Programa ajustable.** El entrenamiento en esta fase fue con un programa ajustable igual al del experimento anterior con sólo una modificación: las respuestas en la palanca de observación durante el componente de reforzamiento apagaban la luz general y encendían las dos luces sobre las palancas durante cinco segundos (E+). Las respuestas en la palanca de observación en el componente de extinción encendían el tono a 70 dB durante cinco segundos (E-). Los cambios en la duración del componente de extinción y en la probabilidad de reforzamiento en función al desempeño de los sujetos fueron exactamente los mismos a los del Experimento 1. El criterio para que los sujetos pasaran a la siguiente fase fue que alcanzaran los valores del extremo final del programa ajustable durante seis sesiones consecutivas.

**Fase de prueba: Procedimiento de observación estándar.** En la fase de prueba, se expuso a los sujetos a un procedimiento de observación que consistió en un programa mixto IA 5

s ( $t = 1$  s,  $p = 0.2$ ) extinción. El componente de reforzamiento tuvo una duración de 30 s y el de extinción de 60 s. A diferencia del Experimento 1, se añadió el E- de tal forma que las presiones en la palanca de observación durante el componente de reforzamiento apagaban la luz general y encendían las dos luces sobre las palancas durante cinco segundos, y las respuestas en la palanca de observación en el componente de extinción activaban el tono a 70 dB durante cinco segundos. El criterio para finalizar esta condición fue de 15 sesiones.

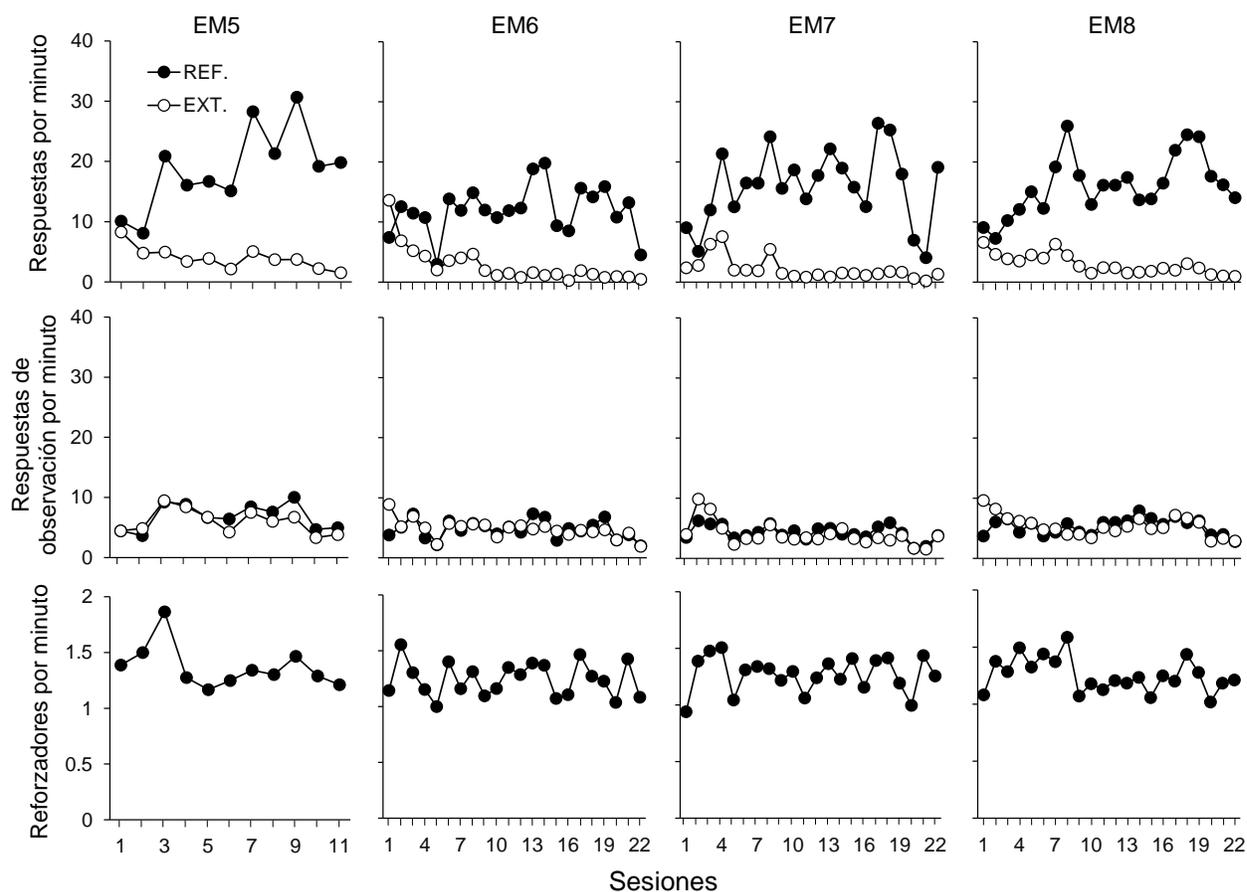
**Prueba de función de los estímulos: Estímulos invertidos y redeterminación.** Para probar la función del E+ y el E-, se colocó a los sujetos en un procedimiento similar al de la fase de prueba con la única diferencia de que los estímulos de cada componente estaban invertidos. Es decir, las presiones a la palanca de observación en el componente de reforzamiento encendían el tono a 70 dB durante cinco segundos, y las respuestas a la palanca de observación en el componente de extinción apagaban la luz general y encendían las dos luces sobre las palancas durante cinco segundos. Para evitar la extinción de las respuestas a las palancas, las ratas pasaron de esta condición si transcurrían cinco sesiones consecutivas con una frecuencia de reforzamiento en niveles cercanos a cero o hasta que transcurrieran 15 sesiones. Finalmente, se regresó a los sujetos a la fase de prueba durante 15 sesiones.

## **Resultados**

### ***Programa ajustable***

La Figura 3 muestra la tasa de respuesta en la palanca de reforzamiento y la tasa de RO durante los componentes de reforzamiento y extinción en todas las sesiones del programa ajustable del Experimento 2. También se muestra la frecuencia de reforzamiento en cada sesión del programa. Las tres variables se calcularon de la misma forma que en el Experimento 1. Se muestran menos sesiones del sujeto EM5 porque, cuando se expuso al programa ajustable, dejó

de presionar la palanca. Para garantizar que el sujeto EM5 presionara las dos palancas de la cámara experimental y sus respuestas produjeran reforzadores en el programa ajustable, recibió de nuevo el entrenamiento de presiones a la palanca izquierda, y al terminar, se entrenaron las respuestas a la palanca derecha de la misma forma. Luego de entrenar las respuestas a ambas palancas, se le introdujo de nuevo al programa ajustable y alcanzó el criterio para pasar a la siguiente fase en menos sesiones que el resto de los sujetos.



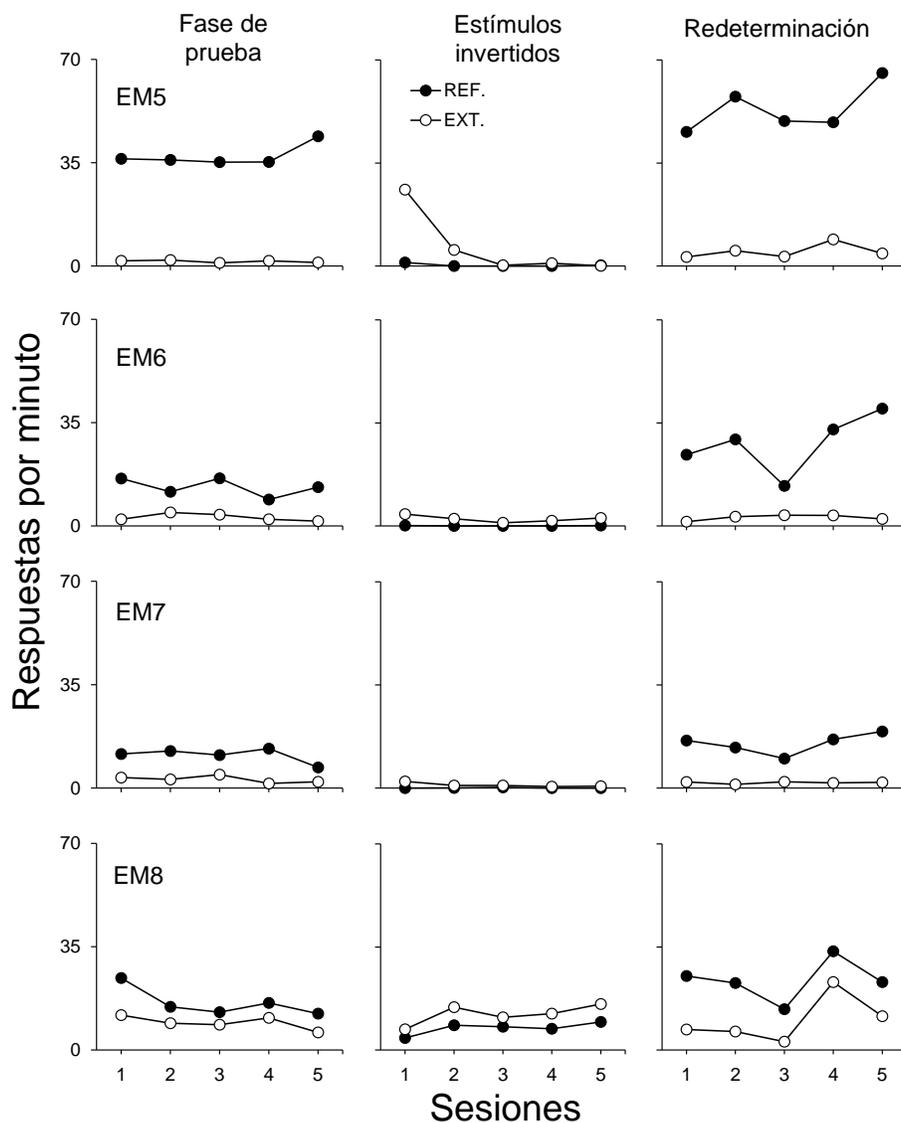
*Figura 3.* Tasa de respuesta por agua, de RO y frecuencia de reforzamiento en el programa ajustable del Experimento 2. Tasa de respuesta en cada palanca y reforzadores por minuto a través de las sesiones para los cuatro sujetos en el programa ajustable en el Experimento 2. Los marcadores negros y blancos corresponden a los componentes de reforzamiento y de extinción respectivamente.

La tasa de respuesta a la palanca de reforzamiento fue relativamente estable para todos los sujetos durante esta fase. La tasa de respuesta a la palanca de reforzamiento fue mayor durante el componente de reforzamiento que en el componente de extinción en todas las sesiones para tres de los sujetos (EM5, EM7 y EM8). Para el sujeto EM6 la tasa de respuesta a la palanca de reforzamiento fue más alta en el componente de extinción únicamente en la primera sesión.

Las RO en ambos componentes ocurrieron de manera relativamente estable para todas las ratas durante todas las sesiones del programa ajustable. En las primeras sesiones, la tasa de RO fue mayor en el componente de extinción que en el componente de reforzamiento para todas las ratas. Posteriormente, las RO no estuvieron claramente diferenciadas entre componentes para tres de los cuatro sujetos (EM6, EM7 y EM8). En el caso de EM5, las RO ocurrieron más frecuentemente en el componente de reforzamiento durante las últimas seis sesiones. La frecuencia de reforzamiento fue estable a través de las sesiones para todos los sujetos en el programa ajustable.

#### ***Fase de prueba: Procedimiento de observación y prueba de función de los estímulos***

La Figura 4 muestra la tasa de respuesta en la palanca de reforzamiento de todos los sujetos durante las últimas cinco sesiones del procedimiento de observación, las últimas cinco sesiones de la condición con estímulos invertidos, y las últimas cinco sesiones de la redeterminación. En todos los sujetos se observó una tasa de respuesta relativamente estable que fue más alta durante el componente de reforzamiento respecto al componente de extinción en la fase de prueba. Al pasar a la condición donde los estímulos de la fase de prueba se invirtieron, la tasa de respuesta disminuyó a niveles cercanos a cero para tres de los cuatro sujetos (EM5, EM6 y EM7). El sujeto EM5 estuvo en la condición de estímulos invertidos durante cinco sesiones, EM6 durante seis sesiones y EM7 durante siete sesiones.



*Figura 4.* Tasa de respuesta en la palanca de reforzamiento durante las tres condiciones de prueba. Respuestas por minuto en la palanca de reforzamiento en las últimas cinco de las tres condiciones para los cuatro sujetos del Experimento 2. Los marcadores negros y blancos señalan al componente de reforzamiento y al componente de extinción respectivamente.

En el caso del sujeto EM8, las respuestas en la palanca de reforzamiento no dejaron de ocurrir y fue el único en pasar 15 sesiones en la condición con estímulos invertidos. Sin embargo, la tasa de respuesta de EM8 aumentó en el componente de extinción y disminuyó en el componente de reforzamiento, de tal modo que la tasa de ambos componentes se invirtió

respecto a la fase de prueba. En la fase de redeterminación, las respuestas por agua aumentaron en ambos componentes y volvieron a estar diferenciadas como en la fase de prueba para todos los sujetos.

En la Figura 5 se muestra la tasa de RO de todos los sujetos durante las últimas cinco sesiones de la fase de prueba, las últimas cinco sesiones de la condición con estímulos invertidos, y las últimas cinco sesiones de la redeterminación. Durante la fase de prueba, la tasa de RO fue ligeramente más alta en el componente de extinción respecto al componente de reforzamiento para tres de los cuatro sujetos (EM6, EM7 y EM8). Para el sujeto EM5, la tasa de RO fue claramente más alta durante el componente de reforzamiento respecto al de extinción en la fase de prueba. En la fase con estímulos invertidos, la tasa de RO disminuyó para todos los sujetos y llegó a niveles cercanos a cero para tres de los cuatro sujetos (EM5, EM7 y EM8). La tasa de RO durante la fase con estímulos invertidos no estuvo claramente diferenciada entre componentes para ninguno de los sujetos, salvo las últimas dos sesiones de EM7, donde la tasa de RO fue más alta en el componente de extinción. Al pasar a la fase de redeterminación, las RO volvieron a ocurrir de manera consistente en ambos componentes para todos los sujetos. La tasa de RO de todos los sujetos fue más alta durante la condición de redeterminación respecto a las dos condiciones anteriores. Las RO estuvieron claramente diferenciadas entre componentes para dos de los cuatro sujetos (EM5 y EM7). En el caso de EM5, las RO ocurrieron más durante el componente de reforzamiento en comparación con el componente de extinción. En el caso de EM7, la tasa de RO fue más alta durante el componente de extinción respecto al componente de reforzamiento.

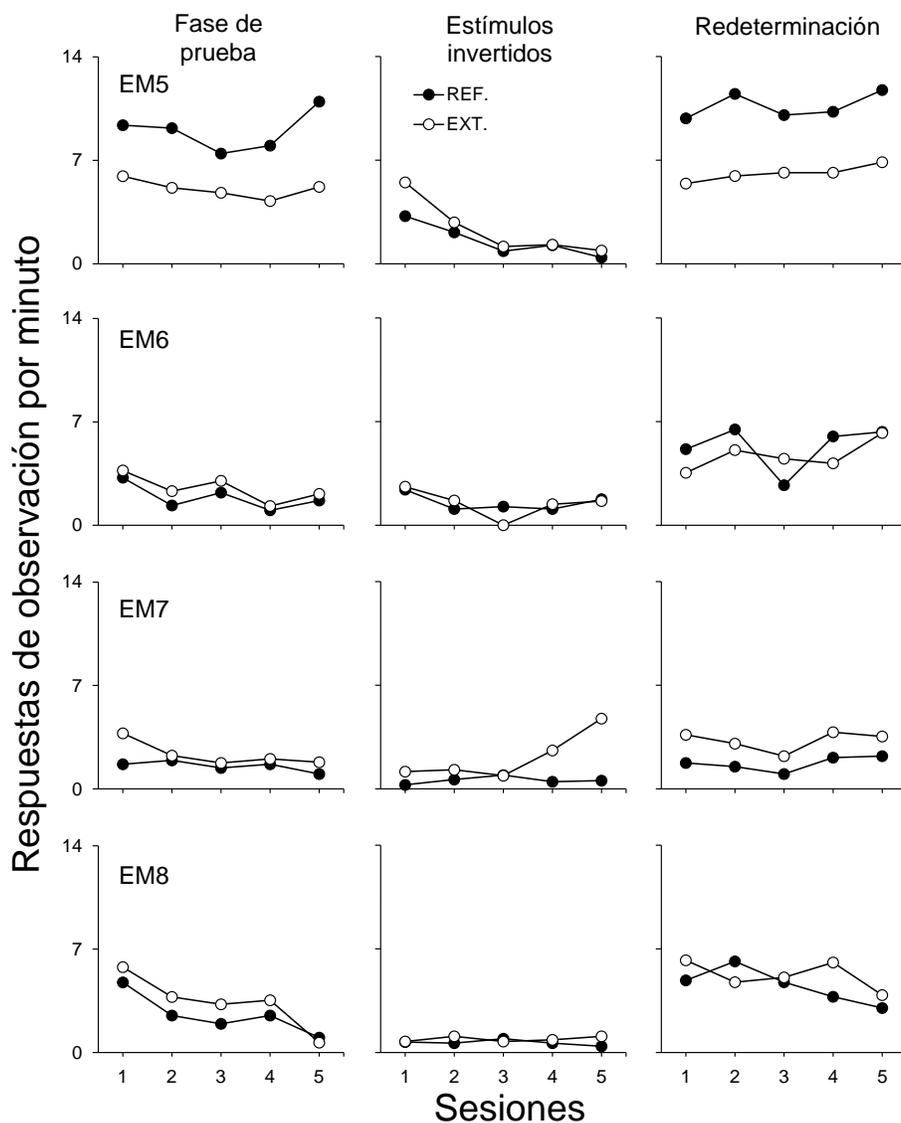
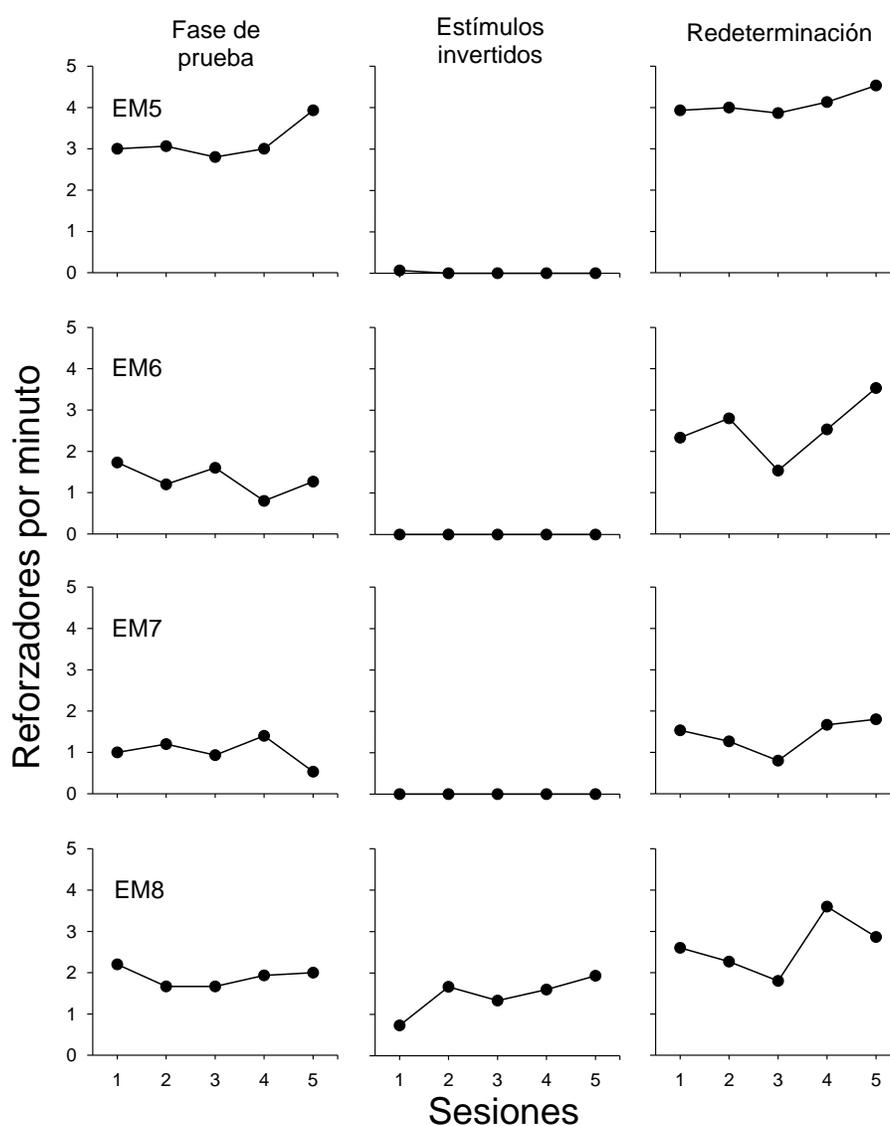


Figura 5. Tasa de RO durante las tres condiciones de prueba. Respuestas por minuto en la palanca de observación en las últimas cinco de las tres condiciones de prueba para los cuatro sujetos del Experimento 2. Los marcadores negros y blancos señalan al componente de reforzamiento y al componente de extinción respectivamente.

La Figura 6 muestra la frecuencia de reforzamiento de todos los sujetos durante las últimas cinco sesiones de la fase de prueba, las últimas cinco sesiones de la condición con estímulos invertidos, y las últimas cinco sesiones de la redeterminación. La frecuencia de reforzamiento fue relativamente estable en la fase de prueba para todos los sujetos. Cuando se

invirtieron los estímulos, la frecuencia de reforzamiento disminuyó para todos los sujetos y llegó a niveles de cero para tres de ellos (EM5, EM6 y EM7). En la fase de redeterminación la frecuencia de reforzamiento volvió a aumentar y fue más alta en comparación con la fase de prueba para todos los sujetos.



*Figura 6.* Frecuencia de reforzamiento durante las tres condiciones de prueba. Reforzadores por minuto en las últimas cinco de las tres condiciones de prueba para los cuatro sujetos del Experimento 2.

El tiempo del entrenamiento preliminar en el Experimento 2 fue de 35 sesiones para tres de los sujetos (EM6, EM7 y EM8): una sesión para el entrenamiento de aproximación a la fuente de reforzamiento y moldeamiento de las presiones a la palanca, 12 sesiones del programa RF1 al programa RV10, y 22 sesiones de programa ajustable. Para el sujeto EM5 el entrenamiento preliminar tomó un total de 48 sesiones: recibió el entrenamiento de aproximación a la fuente de reforzamiento y moldeamiento de la presión a la palanca (una sesión), recibió dos veces el entrenamiento del programa RF1 al programa RV10 en la palanca izquierda (24 sesiones), recibió una vez el entrenamiento del programa RF1 al programa RV10 en la palanca de reforzamiento (12 sesiones), y se expuso durante 11 sesiones en el programa ajustable.

### **Discusión**

El Experimento 2 replicó los resultados del Experimento 1, pues al pasar del programa ajustable al procedimiento de observación, en todos los sujetos se observó control de estímulos sobre las respuestas por agua y las respuestas en la palanca de observación se mantuvieron pese a que no alteraban la disponibilidad del reforzador. El Experimento 2 probó la efectividad del programa ajustable para el establecimiento de RO en todos los sujetos incluyendo los dos estímulos, positivo y negativo. Considerando las inconsistencias actuales que existen en el entrenamiento de las RO, el programa ajustable de este experimento podría aportar una opción más práctica y efectiva para las investigaciones que usen procedimientos de observación con ratas como sujetos.

El otro propósito del segundo experimento fue probar de manera directa la función del E+ y el E- en el procedimiento de observación. Para ello se incluyó una condición con los estímulos del procedimiento de observación invertidos, y posteriormente se regresó a los sujetos a la

condición anterior. Esta manipulación fue muy similar a una que hizo Wyckoff (1952, 1969) en uno de sus experimentos.

En la condición donde los estímulos estuvieron invertidos, las respuestas a la palanca de reforzamiento de todos los sujetos fueron más altas en el componente de extinción, aun cuando en el procedimiento de observación ocurrían de manera contraria. Esto sugiere que las respuestas a la palanca de reforzamiento estaban controladas por el estímulo que en el pasado señalaba el componente de reforzamiento en el procedimiento de observación, por lo que, al invertirse los estímulos, las respuestas ocurrían más en presencia de dicho estímulo aunque el componente vigente fuera el de extinción. Al contrario, las respuestas ocurrían menos durante el componente de reforzamiento porque estaba presente el estímulo que señalaba el componente de extinción en la fase de prueba. Finalmente, cuando se restauraron los estímulos en la fase de redeterminación, las respuestas en la palanca de reforzamiento volvieron a ser más altas durante el componente de reforzamiento. Estos resultados replicaron los hallazgos de Wyckoff (1952, 1969), probando la función discriminativa del E+ y del E-.

Otro hallazgo fue que las RO disminuyeron en todos los sujetos cuando se invirtieron los estímulos de la fase de prueba. Esto puede explicarse porque el E+ dejó de ser un predictor de la disponibilidad del reforzador. Para todos los sujetos en la redeterminación, las RO aumentaron, incluso a niveles mayores que la fase de prueba. La disminución de las RO durante la condición de estímulos invertidos es evidencia de que las RO estuvieron mantenidas por la presentación de estímulos que señalaban la disponibilidad o no disponibilidad del reforzador. El efecto de reforzamiento condicionado del E+ y el E- puede observarse especialmente en el sujeto EM8, que nunca dejó de emitir respuestas en la palanca de reforzamiento ni dejó de producir reforzadores. Si las RO se hubieran mantenido por la entrega de reforzadores primarios, las RO

se hubieran emitido de manera consistente en la condición de estímulos invertidos junto con las respuestas a la palanca de reforzamiento. Sin embargo, las RO se mantuvieron cercanas a cero porque los estímulos ya no señalaban la disponibilidad del reforzador. En cambio, las respuestas en el componente de reforzamiento se mantuvieron porque seguían produciendo entregas del reforzador primario. Los resultados de la condición de estímulos invertidos y la redeterminación del Experimento 2 replicaron lo obtenido por Wyckoff (1952, 1969) con una manipulación idéntica.

Un caso particular de las RO en el Experimento 2 fue el sujeto EM5, pues a diferencia de los demás sujetos, la tasa de RO en el componente de reforzamiento fue mayor a la del componente de extinción. Es relevante recordar que con el sujeto EM5 se realizó entrenamiento de respuestas a ambas palancas. Una posible explicación a sus resultados es que las RO se hayan mantenido tanto por la producción del E+, como por la entrega del reforzador primario, pues las presiones a la palanca de observación estuvieron asociadas al reforzador primario durante el entrenamiento de presiones a la palanca. No obstante, es difícil saber con certeza la razón de los resultados obtenidos en las RO de este sujeto. Estudios posteriores que utilicen el programa ajustable en el entrenamiento podrían hacer comparaciones directas entre sujetos entrenados a presionar en una sola palanca y sujetos entrenados a presionar en ambas.

Un hallazgo para todos los sujetos fue que las respuestas en ambas palancas fueron mayores en la redeterminación respecto a la fase de prueba. Es posible que la diferencia entre ambas condiciones haya sido resultado del tiempo de exposición al procedimiento de observación, pues entre la fase de prueba y la fase de redeterminación, los sujetos estuvieron un total de 30 sesiones en contacto con el mismo programa (el doble de tiempo que los sujetos del Experimento 1). El número de sesiones en el procedimiento pudo haber favorecido que los

sujetos tuvieran más contacto con las contingencias del programa, llevándolos a producir una mayor cantidad de reforzadores. El aumento en la frecuencia de reforzamiento pudo incrementar las respuestas en la palanca de reforzamiento y aumentar así el número de emparejamientos entre el E+ y el reforzador, lo que posiblemente llevó a un incremento de las RO. Sin embargo, es difícil asegurar que el tiempo de exposición fue lo que causó este incremento en las respuestas a ambas palancas porque el procedimiento de observación se interrumpió por la condición de estímulos invertidos durante algunas sesiones.

El tiempo del entrenamiento preliminar para tres de los sujetos del Experimento 2 fue de 35 sesiones. Sin embargo, se observaron tasas de respuestas diferenciadas en la palanca de reforzamiento y tasas de RO consistentes desde las primeras 10 sesiones del programa ajustable. El criterio para pasar del programa ajustable al procedimiento de observación en el Experimento 2 fue que los sujetos alcanzaran los valores finales del programa ajustable durante seis sesiones consecutivas. Si el criterio se redujera a sólo una sesión en lugar de seis, el entrenamiento podría ser aún más breve. El caso del sujeto EM5 muestra que no todos los sujetos emiten las respuestas necesarias que producen reforzadores en el programa ajustable inmediatamente. En este caso, entrenar las respuestas en la palanca de observación condujo al establecimiento RO con el programa ajustable, tal como han hecho otros estudios con procedimientos de observación (Dinsmoor et al. 1981; Kelleher, 1958; Raiff, 2008). Para futuras investigaciones, entrenar las presiones a la palanca con programas de reforzamiento con un requisito más alto al de los utilizados en este trabajo (e.g. RV 15 o RV 20) podría facilitar que las respuestas produzcan reforzadores en el programa ajustable.

## Discusión general

Los procedimientos de observación son relativamente complejos porque involucran el entrenamiento de dos respuestas ocurriendo de manera concurrente en operandos distintos, una de las respuestas produce la entrega de reforzadores y la otra únicamente la producción de estímulos discriminativos, lo que se ha interpretado en términos de competencia de respuestas a ambos operandos (Escobar, 2010; Shahan, 2002). La competencia de las respuestas ha resultado en dificultades para el entrenamiento y mantenimiento de las RO (Gaynor & Shull, 2002; Slezak & Anderson, 2014), amplios periodos de tiempo para el entrenamiento (Lieberman, 1972; McMillan, 1974) y problemas para encontrar establecimiento de RO en todos los sujetos (Fernández et al., 2015; Slezak & Anderson, 2014). No existe un procedimiento estándar para el entrenamiento de RO, lo que ha obstaculizado conocer las variables necesarias para establecer este tipo de respuestas (Villegas & Bruner, 2008a). Sin embargo, la posibilidad que brinda el procedimiento de observación para estudiar separadamente los efectos discriminativos y reforzantes de los estímulos lo hacen un procedimiento útil en el análisis de la conducta (Dinsmoor, 1983; Williams, 1994).

El presente trabajo tuvo como propósito central evaluar la efectividad de un programa ajustable para entrenar RO. En el Experimento 1 se observó que un programa ajustable resultó en la adquisición de RO en todos los sujetos utilizando solamente el E+ durante el entrenamiento. Los resultados del Experimento 1 se relacionan directamente con los resultados de Fernández et al. (2015). El Experimento 1 brindó evidencia que apoya la importancia de establecer relaciones de contigüidad consistentes entre el E+ y el reforzador desde el inicio del entrenamiento de las RO. Los datos del Experimento 1, al contrario que lo reportado por Villegas y Bruner (2008a, 2008b), descartaron un papel determinante de la tasa de respuesta en el operando de

reforzamiento y de la frecuencia de reforzamiento para establecer RO. A partir de una comunicación personal, T. Villegas (comunicación personal, Mayo 24, 2008), pudo determinarse que la selección de los sujetos experimentales en los estudios de Villegas y Bruner (2008a, 2008b) se realizó a partir de la tasa de respuesta de los sujetos antes del experimento. Esto sugiere una variable importante no descrita en ambos estudios en los que se reportó establecimiento de RO en todos los sujetos. Es probable que Villegas y Bruner encontraran RO en sus experimentos debido a que la conducta de los sujetos durante el entrenamiento produjo accidentalmente proximidad temporal entre el E+ y el reforzador.

El Experimento 2 replicó los resultados del Experimento 1. En el Experimento 2 se observó que el programa ajustable también estableció RO en todos los sujetos cuando se incluyó al E- en el entrenamiento. Además, en los resultados de la fase de prueba del Experimento 2 se observó que, después del programa ajustable, en el procedimiento de observación los estímulos producidos por las RO adquirieron las dos funciones características de estos procedimientos. La primera función es el control discriminativo sobre las respuestas en el operando de reforzamiento, de tal modo que estas respuestas se emiten más en presencia del estímulo que ha señalado el componente de reforzamiento y dejan de ocurrir en presencia del estímulo que ha señalado el componente de extinción. La segunda función es la de reforzadores condicionados para las RO, después de que los estímulos se emparejaron con el reforzador primario. Los resultados de la condición con estímulos invertidos y la redeterminación en el Experimento 2 replicaron los datos de uno de los experimentos de Wyckoff (1952, 1969) y son congruentes con el concepto de reforzamiento condicionado.

Una limitación que se encontró en el Experimento 2, fue que uno de los sujetos requirió que se entrenaran las respuestas en la palanca de observación para que las respuestas produjeran

reforzadores en el programa ajustable. El entrenamiento de las respuestas a ambas palancas alargó el tiempo de entrenamiento preliminar para ese sujeto. Una alternativa para asegurar que las respuestas produzcan reforzadores en el programa ajustable, sería utilizar programas de reforzamiento con un requisito más alto a los utilizados en este trabajo (e.g. RV 15 o RV 20) para el entrenamiento de presiones a la palanca.

En resumen, el presente trabajo aporta un procedimiento para entrenar RO en ratas como sujetos utilizando un programa ajustable. El programa ajustable propuesto en este trabajo pudo lidiar con algunos de los retos del entrenamiento de RO que se encuentran en la literatura, pues 1) funcionó para entrenar RO que se mantienen en un procedimiento de observación, 2) funcionó para todos los sujetos y 3) brindó información sobre las variables determinantes para la adquisición de RO. A pesar de que el tiempo de entrenamiento con el programa ajustable no fue menor al que se ha reportado en la literatura, los resultados sugieren que es posible reducir su duración si se disminuye el criterio para que termine. El programa ajustable propuesto en este trabajo puede servir como guía para un entrenamiento estándar de RO que sea más breve y eficaz. Un entrenamiento estándar puede ser relevante para futuros estudios en líneas de investigación dentro del análisis de la conducta como la del reforzamiento condicionado y la del control de estímulos.

Una limitación del presente trabajo es que no fue posible hacer una comparación directa con otros procedimientos para entrenar RO. Los datos del Experimento 1 y del Experimento 2 sugieren que las características del programa ajustable utilizado son suficientes para la adquisición de RO en todos los sujetos. Sin embargo, los datos del presente trabajo no descartan que existan otros procedimientos cuyas características consideren otras variables determinantes para establecer RO para todos los sujetos y en menos sesiones. Futuros estudios podrían

comparar diferentes procedimientos para entrenar RO incluyendo al programa ajustable. Una comparación entre los distintos procedimientos podría brindar más datos sobre las variables determinantes para la adquisición de RO y para que los estímulos adquirieran una función de reforzador condicionado.

El concepto de reforzamiento condicionado es un concepto que fue central en el análisis de la conducta y ha estado rodeado de controversia desde hace décadas (Escobar, 2011; Davison & Baum, 2006; Shahan, 2010; Shahan, 2013). Numerosos estudios dentro de esta área utilizaron programas de reforzamiento encadenados (Williams, 1994). Los programas encadenados se usan para evaluar la idea de que los estímulos de cada eslabón funcionan como reforzadores condicionados para las respuestas del eslabón anterior, pues de acuerdo con el concepto de reforzamiento condicionado, el reforzador primario no podría mantener las respuestas de los eslabones iniciales si están muy alejados temporalmente (Kelleher & Gollub, 1962). Un problema que se encontró al usar programas encadenados fue la dificultad para aislar experimentalmente los efectos del reforzador primario de aquellos atribuidos al reforzador condicionado. Por ejemplo, una manipulación para variar la frecuencia de los reforzadores condicionados fue variar la duración del eslabón inicial (a menor duración, mayor frecuencia de reforzamiento condicionado), pero esta manipulación también modificaba la frecuencia de entrega del reforzador primario. Para evitar los efectos de las variaciones de la frecuencia de reforzamiento, las RO se incluyeron en los programas encadenados para estudiar reforzamiento condicionado (Dinsmoor, 1983, Shahan, 2010; Williams, 1994). Los procedimientos de observación son una buena alternativa para usarse en programas encadenados porque las RO no alteran la entrega del reforzador primario, por lo que revelan de forma separada la función reforzante de los estímulos de cada eslabón. Sin embargo, se han reportado dificultades para

establecer y mantener RO en los programas encadenados, incluso en condiciones donde no hay un requisito de respuestas en el operando de reforzamiento (Slezak & Anderson, 2014).

Dado que las RO se mantienen por la presentación de estímulos asociados al reforzador primario, puede establecerse que estudiar la adquisición de RO es igual a estudiar las condiciones bajo las cuales un estímulo adquiere la función de reforzador condicionado y mantiene una respuesta. Es posible que las inconsistencias en el entrenamiento de las RO y que Slezak y Anderson (2014) reportaran dificultades para entrenar RO se deban a que aún se desconocen las variables determinantes del reforzamiento condicionado. Las características del programa ajustable utilizado en el presente trabajo y los datos del Experimento 1 y el Experimento 2 sugieren que la proximidad temporal entre el E+ y el reforzador primario son suficientes para entrenar RO. En ese caso, utilizar el programa ajustable del presente trabajo podría contribuir a que continúe la investigación de reforzamiento condicionado en los programas encadenados. No obstante, existen otras preocupaciones en la literatura respecto a la conceptualización del reforzamiento condicionado que han dirigido la investigación hacia otras direcciones.

Una definición comúnmente aceptada de un reforzador condicionado es un estímulo inicialmente neutro que adquiere la capacidad de fortalecer las respuestas que los producen debido a su asociación con un reforzador establecido. En ese sentido, los reforzadores condicionados funcionan como tales en tanto se mantengan emparejados con el reforzador primario. Sin embargo, algunos autores han formulado otras interpretaciones del reforzamiento condicionado (véase Shahan, 2010 para una breve revisión). Las otras interpretaciones proponen nombres diferentes para los reforzadores condicionados, pero coinciden en que los estímulos que incrementan las respuestas que les preceden lo hacen porque señalan de manera consistente la

localización espacial o temporal de más reforzadores, no porque exista un proceso de fortalecimiento de las respuestas.

Por ejemplo, Davison y Baum (2006) elaboraron un par de experimentos con palomas como sujetos usando programas concurrentes en los que una luz en el comedero siempre acompañaba la entrega de comida. La frecuencia relativa de reforzamiento de las dos opciones del programa concurrente cambiaba a través de distintos componentes que no estaban señalados, de tal forma que la preferencia hacia una opción u otra cambiaba cuando cambiaba la frecuencia relativa de reforzamiento. En algunas condiciones, una proporción de la entrega de reforzadores consistía únicamente en la luz del comedero sin entrega de comida. De esta forma, la entrega de comida siempre estaba acompañada de la luz del comedero, pero la luz del comedero no siempre precedía la entrega de comida. En otras condiciones, una luz verde en la tecla correlacionaba (correlación de +1, 0 o -1 según la condición) con la opción de mayor frecuencia de reforzamiento sin que se emparejara con el reforzador primario. Los resultados de los experimentos mostraron que la preferencia se dirigía hacia una opción si las respuestas en esa opción producían un estímulo que predijera una mayor frecuencia relativa de reforzamiento, incluso si este estímulo no se emparejaba nunca con el reforzador primario. Davison y Baum argumentaron que interpretar a los reforzadores condicionados como señales predictoras de reforzadores es más parsimonioso, pues el emparejamiento entre los estímulos y el reforzador primario no parece ser relevante para que los organismos respondan si esos estímulos no son buenos predictores de más reforzadores.

Por otro lado, Shahan (2010) coincide con Davison y Baum y considera principalmente dos experimentos para proponer su propia conceptualización sobre los reforzadores condicionados. El primer experimento es el reportado por Shuster (1969), en el que palomas se

encontraron en un programa de reforzamiento concurrente encadenado de dos opciones. Los eslabones de las dos opciones del programa eran idénticos en términos de los programas de reforzamiento vigentes. La entrega del reforzador en ambas opciones era precedida de un estímulo. La única diferencia entre las opciones del programa era que en una de ellas había presentaciones adicionales del estímulo que precedía al reforzador, pero sin la entrega del reforzador. Según Shahan, el efecto esperado según la conceptualización tradicional de reforzador condicionado era que las palomas eligieran la opción con las presentaciones adicionales del estímulo emparejado al reforzador. No obstante, las palomas elegían la opción donde el estímulo sólo se presentaba previo a la entrega del reforzador. El segundo experimento se realizó por Bolles (1961) con ratas como sujetos que se encontraban en un programa concurrente de dos opciones con una palanca para cada opción. En este experimento se programaron entregas de comida asociadas con un click de tal forma que la entrega de comida del lado de una palanca precedía más comida en el lado de la otra palanca. Cuando se colocó a los sujetos en una condición de extinción, el sonido del click de una palanca hacía que las ratas respondieran en la palanca opuesta. Bolles concluyó con los datos de su experimento que el click no funcionó como un reforzador condicionado, sino como una señal para la obtención del reforzador.

La propuesta de Shahan basada en los experimentos como los de Bolles, Shuster y Davison y Baum, consiste en reconsiderar la función de los reforzadores condicionados en términos de señales que guían a los organismos en la obtención de reforzadores. Shahan cuestionó la conceptualización tradicional de los reforzadores condicionados por dos razones principales: 1) por la noción implícita de que los reforzadores condicionados fortalecen las respuestas que los producen y que implica que haya una transferencia de valor del reforzador

primario hacia el reforzador condicionado y 2) por considerar que la contigüidad temporal es la variable determinante para que un estímulo adquiriera un efecto sobre la conducta (Shahan, 2010; Shahan, 2017). En cambio, Sahahan argumenta que los estímulos adquieren efectos sobre la conducta dado que predicen o tienen una correlación con el reforzador primario y, por lo tanto, se vuelven necesarios para tener contacto con él, es decir, un “medio para llegar a un fin” (*means-to-an-end*). En ese sentido, Shahan retoma una analogía de Davison y Baum que propone reemplazar la metáfora de fortalecimiento de las respuestas que plantea el reforzamiento condicionado por una metáfora de guía por parte de los estímulos para los organismos. La conceptualización de Shahan nombra a los reforzadores condicionados como *signposts*.

Shahan plantea que debido a la historia evolutiva, los organismos son capaces de detectar correlaciones entre los eventos del ambiente y los reforzadores. La detección de las regularidades en el ambiente se volvería fundamental para que los organismos encontraran fuentes de reforzamiento. Así, una vez que los organismos detectan que un estímulo está correlacionado con la presencia de reforzadores, los organismos tienden a seguir ese estímulo (*signpost*) o emitir conductas que les pongan en contacto con él. Shahan (2013) sugiere que la función de los *signpost* es similar a la de los estímulos discriminativos, pues señalan la disponibilidad espacial y temporal de un reforzador, así como la conducta asociada a su obtención. Si es así, entonces los reforzadores condicionados y los estímulos discriminativos afectarían la conducta por la misma razón: señalar el reforzador y guiar al organismo hacia él, haciendo innecesaria la noción del fortalecimiento de las respuestas por parte de los reforzadores condicionados.

El experimento de Shuster (1969) daría soporte a la importancia de la correlación entre estímulo y reforzador para que el estímulo funcione como *signpost* y tenga un efecto en la conducta. En el experimento de Shuster, al programar presentaciones adicionales del estímulo

emparejado con el reforzador en una opción, se reduce la predictibilidad o correlación entre el estímulo y el reforzador, por lo que las palomas prefirieron la opción sin presentaciones adicionales. Por otro lado, el experimento de Bolles (1961) ilustraría que los signposts señalan la ubicación espacial de los reforzadores, pues los clicks del lado de una palanca predecían comida en la palanca del lado opuesto y las respuestas de los sujetos se ajustaban a esa predicción.

En procedimientos de observación, Shahan y Podlesnik (2005, 2008) han reportado que la tasa de reforzamiento condicionado y el valor del reforzador condicionado (E+) afecta a la tasa de RO, lo que apoya a la conceptualización de reforzador condicionado. Sin embargo, Shahan y Podlesnik también encontraron que en condiciones de interrupción, la frecuencia de E+ y el valor del E+ no tiene efectos sobre la resistencia al cambio de las RO, contrario a lo esperado desde la conceptualización de reforzador condicionado. Shahan (2010, 2013) explicó que el hecho de que no se observen cambios en la resistencia al cambio de RO brinda soporte a la conceptualización de los signposts. Cabe mencionar que se ha señalado que los procedimientos de Shahan y Podlesnik son complejos y los resultados son difíciles de interpretar (Escobar, 2011; Shahan, 2010).

La función de los estímulos en los experimentos del presente trabajo puede interpretarse de manera congruente con las afirmaciones de autores como Davison y Baum (2006) y Shahan (2010, 2013, 2017). Desde el punto de vista de los signposts, las RO son respuestas que se mantienen por producir signposts que señalan el momento en el que un reforzador está disponible y si las respuestas deben distribuirse hacia el operando de reforzamiento o hacia el operando de observación. Como se mencionó anteriormente, existe un patrón de respuestas reportado en procedimientos de observación que resulta en tasas de RO mayores durante el componente de extinción respecto al componente de reforzamiento (Escobar & Bruner, 2009).

Los resultados del Experimento 2 mostraron esta diferencia de RO entre los dos componentes en tres de los cuatro sujetos. Nociones como la de los signposts podría explicar estos resultados, pues es posible que el E- funcionara como una señal de que las respuestas debían continuar en la palanca de observación para que eventualmente se produjera el E+. En el momento que el E+ se presentara, éste funcionaría como una señal para que las respuestas cambiaran a la palanca de reforzamiento porque el E+ funcionaba como un predictor de disponibilidad de agua. Además, el programa ajustable para el entrenamiento de las RO hacía que el E+ fuera un buen predictor de agua desde el inicio del procedimiento.

El origen de las RO fue un estudio de control de estímulos (Wyckoff, 1952, 1969), y desde entonces se han considerado importantes dentro de esa línea de investigación (Dinsmoor, 1985). Las RO introducidas por Wyckoff, pueden considerarse respuestas de observación “artificiales”, pues naturalmente, las respuestas que ponen a los organismos en contacto con los estímulos del ambiente suelen ser diferentes (e.g. mirar, escuchar, olfatear, tocar o saborear) (Dinsmoor, 1985; Gaynor & Shull, 2002). Las RO de Wyckoff son respuestas que se utilizan en la investigación conductual con el fin de tener un registro más confiable de que los organismos tienen contacto con los estímulos discriminativos de los programas. De hecho, las RO se han utilizado como un análogo de la atención desde su origen, porque pueden considerarse un indicador de que los organismos “atienden” a los estímulos (Kelleher, 1958; Shahan, 2013). Con base en los datos de varios estudios, Dinsmoor (1985) sugirió que el establecimiento del control de estímulos depende de qué tanto ocurran las RO (naturales o artificiales). A su vez, Dinsmoor señaló que las RO se emiten en mayor o menor medida dependiendo de la correlación que existe entre los estímulos y la disponibilidad o ausencia de reforzadores. Es decir, en qué medida señalan la ocasión para que una respuesta sea reforzada (E+) o no (E-). Si la correlación es alta,

las RO ocurren consistentemente y viceversa, tal como encontró Wyckoff al invertir los estímulos de su procedimiento de observación, y como se replicó en el Experimento 2 de este trabajo. El entrenamiento propuesto en el presente trabajo puede ser útil para futuras investigaciones que evalúen si entrenar RO favorece el control de estímulos, sobre todo en estudios de discriminación condicional o de control de estímulos complejo. Los datos de estas investigaciones podrían servir como guía para el diseño de intervenciones en humanos en los que se observe una discriminación de estímulos pobre, o que presenten lo que coloquialmente se les llama problemas de atención.

## Referencias

- Békèsy, G. V. (1947). A new audiometer. *Acta Oto-Laryngologica*, 35, 411–422.  
<https://doi.org/10.3109/00016484709123756>
- Bolles, R. C. (1961). Is the “click” a token reward? *The Psychological Record*, 11, 163–168.  
<https://doi.org/10.1007/BF03393399>
- Case, D. & Fantino, E. (1981). The delay-reduction hypothesis of conditioned reinforcement and punishment: observing behavior. *Journal of the Experimental Analysis of the Behavior*, 35(1), 93–108. <https://doi.org/10.1901/jeab.1981.35-93>
- Colares, T. R., De Faria, A. L., De Almedia, L. A., De Faria, O., McIlvane, W. J. (2020). Acquisition and maintenance of delayed matching-to-sample in tufted capuchin monkeys. *Journal of the Experimental Analysis of the Behavior*, 113(3), 549–564.  
<https://doi.org/10.1002/jeab.599>
- Davison, M., & Baum, W. M. (2006). Do conditional reinforcers count? *Journal of the Experimental Analysis of Behavior*, 86(3), 269–283.  
<https://doi.org/10.1901/jeab.2006.56-05>
- DeFulio, A. & Hackenberg, T. (2008). Combinations of response-dependent and response-independent schedule correlated stimulus presentation in an observing procedure. *Journal of the Experimental Analysis of Behavior*, 89(3), 299–309.  
<https://doi.org/10.1901/jeab.2008.89-299>
- Dinsmoor, J., Mulvaney, D., & Jwaideh, A. (1981). Conditioned reinforcement as a function of duration of stimulus. *Journal of the Experimental Analysis of Behavior*, 36(1), 41–49.  
<https://doi.org/10.1901/jeab.1981.36-41>

- Dinsmoor, J. (1983). Observing and conditioned reinforcement. *The Brain and Behavioral Sciences*, 6, 693–728. <https://doi.org/10.1017/S0140525X00017969>
- Dinsmoor, J. A. (1985). The role of observing and attention in establishing stimulus control. *Journal of the Experimental Analysis of Behavior*, 43, 365–381. <https://doi.org/10.1901/jeab.1985.43-365>
- Escobar, R., & Bruner, C. A. (2009). Observing responses and serial stimuli: searching for the reinforcing properties of the S-. *Journal of the Experimental Analysis of Behavior*, 92(2), 215–231. <https://doi.org/10.1901/jeab.2009.92-215>
- Escobar, R. (2010). Travel Distance and Stimulus Duration on Observing Responses by Rats. *European Journal of Behavior Analysis*, 11(1), 79–91. <https://doi.org/10.1080/15021149.2010.11434336>
- Escobar, R. (2011). Respuestas de observación y resistencia a la extinción. *Revista Mexicana de Análisis de la Conducta*, 37(2), 53–71. <https://doi.org/10.5514/rmac.v37.i2.26139>
- Escobar, R., Gutiérrez, B., & Benavides, R. (2022). 3D-printed operant chambers for rats: Design, assembly, and innovations. *Behavioural Processes*, 199, 104647. <https://doi.org/10.1016/j.beproc.2022.104647>
- Escobar, R., & Perez-Herrera, C. (2015). Low-cost USB interface for operant research using Arduino and Visual Basic. *Journal of the Experimental Analysis of Behavior*, 103(2), 427–435. <https://doi.org/10.1002/jeab.135>
- Evans, W. O. (1963). A titration Schedule on a treadmill. *Journal of the Experimental Analysis of Behavior*, 6(2), 2219–221. <https://doi.org/10.1901/jeab.1963.6-219>

- Fernández, G., Torres, C., Flores, C. & Escobar, R. (2015). El papel de la duración de los estímulos en la adquisición de la conducta de observación. *Revista Mexicana de Análisis de la Conducta*, 41(3), 28–47. <https://doi.org/10.5514/rmac.v41.i3.63779>
- Ferster, C., & Skinner, B. F. (1957). *Schedules of reinforcement*. Englewood Cliffs, NJ; Prentice-Hall.
- García-Leal, O., Barrón, E., Camarena-Pérez, H. & Vílchez, Z. (2019). Response rate correlates with indifference points in a delay-discounting procedure. *Journal of the Experimental Analysis of Behavior*, 112(2), 167–176. <https://doi.org/10.1002/jeab.548>
- Gaynor, S. T., & Shull, R. L. (2002). The generality of selective observing. *Journal of the Experimental Analysis of Behavior*, 77, 171-187. <https://doi.org/10.1901/jeab.2002.77-171>
- Green, L., Myerson, J., Shah, A. K., Estle, S. J. & Holt, D. D. (2007). Do adjusting-amount and adjusting-delay procedures produce equivalent estimate of subjective value in pigeons? *Journal of the Experimental Analysis of Behavior*, 87(3), 337–347. <https://doi.org/10.1901/jeab.2007.37-06>
- Hendry, D., & Dillow, P. (1966). Observing behavior during interval schedules. *Journal of the Experimental Analysis of Behavior*, 9, 337–349. <https://doi.org/10.1901/jeab.1966.9-337>
- Hirota, T. T. (1972). The Wyckoff observing response-a reappraisal. *Journal of the Experimental Analysis of Behavior*, 18(2), 263–276. <https://doi.org/10.1901/jeab.1972.18-263>
- Kangas, B. D. & Branch, M. N. (2012). Relations among acute and chronic nicotine administration, short-term memory and tactics of data analysis. *Journal of the Experimental Analysis of Behavior*, 98(2), 155–167. <https://doi.org/10.1901/jeab.2012.98-155>

- Kangas, B. D., Vaidya, M. & Branch, M. N. (2010). Titrating-delay matching-to-sample in the pigeon. *Journal of the Experimental Analysis of Behavior*, 94(1), 69–81.  
<https://doi.org/10.1901/jeab.2010.94-69>
- Kelleher, R. T. (1958). Stimulus-producing responses in chimpanzees. *Journal of the Experimental Analysis of Behavior*, 5, 3–13. <https://doi.org/10.1901/jeab.1958.1-87>
- Kelleher, R.T., Gollub, L.R., 1962. A review of positive conditioned reinforcement. *Journal of the Experimental Analysis of Behavior*, 5, 543–597. <https://doi.org/10.1901/jeab.1962.5-s543>
- Kendall, S. & Gibson, D. (1965). Effects of discriminative stimulus removal on observing behavior. *The Psychological Record*, 15, 545–551. <https://doi.org/10.1007/BF03393628>
- Lieberman, D. A. (1972). Secondary reinforcement and information as determinants of observing behavior in monkeys (*Mucaca mulatta*). *Learning and Motivation*, 3, 341–358.  
[https://doi.org/10.1016/0023-9690\(72\)90030-6](https://doi.org/10.1016/0023-9690(72)90030-6)
- Lieving, G., Reilly, M. & Lattal, K. (2006). Disruption of responding maintained by conditioned reinforcement: alterations in response–conditioned-reinforcer relations. *Journal of the Experimental Analysis of Behavior*, 86(2), 197–209.  
<https://doi.org/10.1901/jeab.2006.12-05>
- McMillan, J. C. (1974). Average uncertainty as a determinant of observing behavior. *Journal of the Experimental Analysis of Behavior*, 22, 401–408.  
<https://doi.org/10.1901/jeab.1974.22-401>
- Mazur, J. E. (1998). Estimation of indifference points with an adjusting-delay procedure. *Journal of the Experimental Analysis of Behavior*, 49(1), 37–47.  
<https://doi.org/10.1901/jeab.1988.49-37>

- Nevin, J. A., & Grace, R. C. (2000). Behavioral momentum and the law of effect. *Behavioral and Brain Sciences*, 23, 73–130. <https://doi.org/10.1017/S0140525X00002405>
- Oldfield, R. C. (1949). Continuous Recording of Sensory Thresholds and Other Psycho-Physical Variables. *Nature*, 164(4170), 581–581. <https://doi.org/10.1038/164581a0>
- Preston, G. C. (1985). Observing Responses in Rats: Support for the Secondary Reinforcement Hypothesis. *The Quarterly Journal of Experimental Psychology*, 37, 23–31. <https://doi.org/10.1080/14640748508402084>
- Schoenfeld, W. N., & Cole, B. K. (1972). *Stimulus Schedules: The T- $\tau$  Systems*. Harper and Row.
- Schuster, R. H. (1969). A functional analysis of conditioned reinforcement. In D. P. Hendry (Ed.), *Conditioned reinforcement* (pp. 192–235). Homewood, IL: The Dorsey Press.
- Shahan, T. A. (2002). The Observing Response Procedure: A Novel Method to Study Drug-Associated Conditioned Reinforcement. *Experimental and Clinical Psychopharmacology*, 10(1), 3–9. <https://doi.org/10.1037/1064-1297.10.1.3>
- Shahan, T. A. (2013). Attention and conditioned reinforcement. In G. J. Madden, W. V. Dube, T. D. Hackenberg, G.P. Hanley, & K. A. Lattal (Eds), *APA handbook of behavior analysis, Vol. 1. Methods and principles* (pp. 387–410). American Psychological Association. <https://doi.org/10.1037/13937-017>
- Shahan, T. A. (2017). Moving beyond reinforcement and response strength. *The Behavior Analyst*, 40, 107-121. <https://doi.org/10.1007/s40614-017-0092-y>
- Shahan, T. A. & Cunningham, P. (2015). Conditioned reinforcement and information theory reconsidered. *Journal of the Experimental Analysis of Behavior*, 103(2), 405–418. <https://doi.org/10.1002/jeab.142>

- Shahan, T. A., & Podlesnik, C. A. (2005). Rate of conditioned reinforcement affects observing rate but not resistance to change. *Journal of the Experimental Analysis of Behavior*, 84(1), 1-17. <https://doi.org/10.1901/jeab.2005.83-04>
- Shahan, T. A., & Podlesnik, C. A. (2008). Conditioned reinforcement value and resistance to change. *Journal of the Experimental Analysis of Behavior*, 89(3), 263-298. <https://doi.org/10.1901/jeab.2008-89-263>
- Slezak J. M., & Anderson K. G. (2014). Observing of chain-schedule stimuli. *Behavioural Processes*, 105, 19-27. <https://doi.org/10.1016/j.beproc.2014.02.004>
- Villegas, T., & Bruner, C.A. (2008a). El papel del entrenamiento preliminar en la adquisición de respuestas de observación. *Acta Comportamentalia*, 16(1), 41–62.
- Villegas, T., & Bruner, C.A. (2008b). Reforzar cuatro respuestas por comida establece conducta de observación. *Acta Comportamentalia*, 16(3), 289–303.
- Williams, B. (1994). Conditioned Reinforcement: Experimental and Theoretical Issues. *The Behavior Analyst*. 17(2), 261–285. <https://doi.org/10.1007/BF03392675>
- Wyckoff, L. B., Jr. (1952). The role of observing responses in discrimination learning. *Psychological Review*, 59(6), 58–68. <https://doi.org/10.1037/h0053932>
- Wyckoff, L. B., Jr. (1969). The role of observing responses in discrimination learning. Part II. En Hendry, D. P. (Eds.), *Conditioned reinforcement*, (pp. 237–260). Homewood, III: Dorsey Press.