

# **Adquisición de la Operante Libre bajo Condiciones de Reforzamiento Demorado: Una Revisión**

*(Acquisition of the free-operant under delayed reinforcement conditions: A review)*

**Marco Antonio Pulido Rull<sup>1</sup>, Rodrigo Sosa Sánchez y Leonor Valadez Téllez**

Universidad Intercontinental

En el estudio del condicionamiento operante Skinner centró su interés en el mantenimiento de patrones de respuesta repetitivos a través del tiempo, así pues el estudio de los fenómenos de adquisición nunca se abordó de manera sistemática (Sidman, 1960). El abordaje de Skinner difirió considerablemente del de otros científicos interesados en fenómenos de aprendizaje. Hull (1952) por ejemplo, construyó una parte importante de su sistema teórico basado en estudios en los cuales la adquisición de la respuesta era la principal variable dependiente de interés y la principal herramienta de recolección de datos el laberinto. Así pues, hoy en día adquisición es un fenómeno ampliamente documentado en procedimientos de ensayo discreto pero escasamente explorado utilizando los procedimientos desarrollados por el análisis experimental de la conducta.

Además del interés de Skinner por el mantenimiento de la tasa de respuesta como variable dependiente, el estudio de la adquisición nunca se emprendió sistemáticamente dentro del análisis experimental de la conducta debido al desarrollo de la técnica de moldeamiento por aproximaciones sucesivas. El moldeamiento le permitió a Skinner desarrollar rápidamente el comportamiento sobre el cual incidía con sus procedimientos de cambio de contingencias y por lo tanto la adquisición se convirtió en una actividad deliberada del investigador (y por lo tanto interesante como actividad propia del científico y no del sujeto).

<sup>1</sup>Dirigir toda correspondencia relacionada con este artículo a: Marco Antonio Pulido Rull. En Av. Universidad No. 1330 Edificio A, Depto. 1102. Colonia del Carmen Coyoacán. C.P. 04100, México, DF. Correo electrónico mpulido@uic.edu.mx. Los autores desean agradecer a los revisores por sus valiosos comentarios acerca del trabajo.

En términos generales el moldeamiento por aproximaciones consiste en seleccionar componentes de la conducta terminal de interés y utilizar las operaciones de reforzamiento y extinción para modificarlos gradualmente y establecer el comportamiento meta (Skinner, 1951; Savage, 2001). Schwart y Gamzu (1977) han sugerido que el moldeamiento es más un arte inefable que una técnica, sin embargo de acuerdo con el propio Skinner (1953) existen algunas reglas básicas que lo rigen. Por ejemplo de acuerdo con Skinner el moldeamiento se facilita considerablemente si la entrega del reforzador es inmediata (y se dificulta de manera notable si se entrega de forma demorada).

El papel secundario que Skinner atribuyó a los procesos de adquisición dentro del análisis experimental de la conducta, así como la adopción generalizada de los procedimientos de moldeamiento llevaron al desarrollo de una conceptualización particular acerca de los procesos de adquisición de la operante libre. Según dicha conceptualización, la adquisición “espontánea” del comportamiento puede ocurrir solamente si el reforzador es inmediato (y aun en estos casos la aparición de la conducta resulta lenta y sus resultados difíciles de predecir). Adicionalmente la idea de que el comportamiento solamente se desarrolla de manera gradual (derivada del uso generalizado de las técnicas de moldeamiento) sugería que la aparición completa de la conducta final buscada por el investigador resultaba improbable.

Las ideas preconcebidas que existían dentro del análisis experimental de la conducta con respecto a los procesos de adquisición de la operante libre fueron cuestionadas en 1990 por un estudio conducido por Lattal y Gleeson. Estos autores entrenaron palomas a recoger alimento de un contenedor de grano colocado en la cámara experimental. Sin otro entrenamiento que el descrito, expusieron a los sujetos a programas tándem RF1, TF 30-s. Sorprendentemente la tasa de respuesta de los animales sometidos a este programa incrementó consistente y gradualmente conforme transcurrieron las sesiones experimentales. Los resultados obtenidos con palomas fueron replicados con ratas. Adicionalmente Lattal y Gleeson obtuvieron clara evidencia de adquisición utilizando programas tándem intermitentes con demora fija (IV 30-s, RDO 10-s), y separando el *manipulandum* del contenedor de comida.

Los resultados de Lattal y Gleeson sugirieron que la adquisición de la operante libre era posible aun bajo condiciones de reforzamiento demorado y sin moldeamiento explícito. De hecho la complejidad de las condiciones experimentales utilizadas por los autores y el hecho de que en todas ellas las tasas de respuesta se hayan elevado consistentemente, sugería que la adquisición era más un fenómeno “inevitable” que improbable. La naturaleza contraintuitiva de los resultados de Lattal y Gleeson generó interés en el fenómeno por parte de la comunidad de analistas de la conducta. Dicho interés se materializó en una gran cantidad de investigaciones que pueden dividirse,

según su temática en varias áreas: 1) Algunos investigadores se interesaron por evaluar la generalidad del fenómeno, es decir en demostrar que puede ocurrir en condiciones diferentes a las evaluadas por Gleeson y Lattal. 2) Dado que existe una amplia variedad de procedimientos mediante los cuales se puede programar la demora de un reforzador, también se han llevado a cabo estudios en los cuales se ha comparado la adquisición de la operante libre utilizando diferentes programas de reforzamiento. 3) Algunos investigadores se han centrado en identificar los límites máximos bajo los cuales puede ocurrir el fenómeno (por ejemplo algunos científicos se han interesado en determinar los valores máximos de demora bajo los cuales puede ocurrir la adquisición). 4) Dado que en la mayor parte de los estudios sobre demora de reforzamiento, esta variable se encuentra confundida con la duración del intervalo entre reforzadores, algunos investigadores han llevado a cabo estudios para evaluar de manera independiente el efecto de demora y tasa de reforzamiento sobre la adquisición de la operante libre. 5) Por último, otro grupo de científicos se ha interesado en identificar variables que pueden facilitar la adquisición de la respuesta. En la presente revisión se abordan todas estas áreas de investigación. Al finalizar la presentación de las mismas se sintetizan sus principales alcances y se plantea una agenda de investigación basada en las limitaciones empíricas y teóricas existentes.

## ACERCA DE LA GENERALIDAD DEL FENÓMENO

Cuando un investigador se topa con un hallazgo que contradice un conjunto de ideas dominantes en su área de interés (como el reportado por Lattal y Gleeson) con frecuencia la agenda de investigación sugiere que el primer paso consiste en evaluar la generalidad del mismo. En un intento por determinar si la adquisición con reforzamiento demorado podía también ocurrir con operantes diferentes a la presión de la palanca y el picoteo a la tecla, Critchfield y Lattal (1993) llevaron a cabo el siguiente estudio. Sometieron a ratas ingenuas, privadas de alimento a programas tándem RF1, RDO 30-s en los cuales la respuesta en cuestión consistía en interrumpir un haz de luz producido por una celda fotoeléctrica. Este tipo de respuesta resultó de interés para los investigadores dado que no producía los característicos ruidos electromecánicos asociados a la utilización de *manipulanda* más convencionales. Los investigadores dividieron a los animales en dos grupos experimentales. En uno de los grupos la interrupción del haz de luz producía un breve tono auditivo; en el segundo grupo la interrupción del haz no produjo el tono. En ambos grupos se observó un incremento gradual y consistente en la tasa de respuesta; sin embargo los sujetos del grupo que recibió el tono, mostró mayor eficiencia (medida en términos de la cantidad de respuestas emitidas por reforzador) que el grupo sin tono. El estudio evidenció que el fenómeno de adquisición con demora puede también ocurrir

con respuestas sustancialmente diferentes a las empleadas por Lattal y Gleeson. También sugirió que los sonidos que produce la operación de los *manipulanda* tradicionales pueden tener efectos facilitadores sobre la adquisición de un nuevo repertorio conductual (idea que ya había sido sugerida por Bolles en 1988).

Schlinger y Blakely (1994) decidieron replicar el experimento de Critchfield y Lattal pero utilizando demoras de distinto valor. Los investigadores sometieron a ratas ingenuas, privadas de alimento, a programas tándem RF1 RDO 4-s o 10-s. Al igual que en el estudio de Critchfield y Lattal los autores dividieron a los animales en dos grupos. Los sujetos del grupo experimental recibieron un tono de .25-s cuando se producía la interrupción del haz de luz; los sujetos del grupo control no recibieron el tono. Los investigadores reportaron que la tasa de respuesta fue más alta y aumentó de manera más rápida en los grupos con tono (independientemente de la duración de la demora). Tanto los grupos con tono como los grupos sin tono mostraron tasas de respuesta más altas que las producidas por una condición de extinción (es habitual utilizar un grupo sometido a una condición de extinción como grupo control en experimentos de adquisición, sin embargo como se verá más adelante no es la única forma de evidenciar la adquisición de la respuesta). En general el estudio confirmó los resultados de Critchfield y Lattal en el sentido de que operantes diferentes al picoteo de la tecla y la presión de la palanca pueden establecerse con operaciones de reforzamiento demorado.

Con la finalidad de determinar si la adquisición de la respuesta con reforzamiento demorado ocurría en especies diferentes a la rata y la paloma, Lattal y Metzger (1994) sometieron a peces Siameses (*Betta splendens*) a la siguiente situación. Sometieron a los sujetos a un programa tándem RF1 RDO (0-s, 10-s o 25-s). La operante en cuestión consistió en atravesar un aro sumergido en la pecera, se empleó como reforzador la presentación de un espejo durante 15-s. Los resultados mostraron un incremento gradual y consistente de la tasa de respuesta conforme transcurrieron las sesiones experimentales en todos los sujetos expuestos a un programa contingente. La tasa de respuesta de animales de control, expuestos a un programa no contingente no se elevó. Los investigadores encontraron tasas de respuesta más altas en la condición de reforzamiento inmediato y progresivamente más bajas conforme se incrementó el valor de la demora. El estudio en cuestión resulta relevante dado que se mostró que la adquisición con reforzamiento demorado ocurre en especies distintas a las empleadas en los estudios previos. También mostró que el fenómeno puede ocurrir con reforzadores distintos a los empleados comúnmente dentro del análisis experimental de la conducta (principalmente apetitivos como la comida o el agua). El estudio también confirmó el hallazgo de Critchfield y Lattal en el sentido de que respuestas diferentes al picoteo de la tecla y la presión de la palanca se pueden establecer con reforzamiento demorado.

La generalidad del fenómeno de adquisición con reforzamiento demorado también

se ha demostrado utilizando procedimientos con dos palancas. Stuphin, Byrne y Poling (1998) sometieron a ratas ingenuas privadas de agua a programas tándem RF1, RDO (0-s, 8-s, 16-s, 32-s o 64-s). Los animales fueron asignados a programas que difirieron en cuanto a la duración de la demora y al funcionamiento de dos palancas que estaban presentes dentro de la caja de condicionamiento. En una de las condiciones experimentales una de las palancas producía agua después del intervalo de demora y las presiones en la otra palanca no tenían consecuencias programadas; en la segunda condición una de las palancas producía agua y las presiones a la otra cancelaban la entrega del líquido. Con las demoras breves (0-s y 8s) los sujetos emitieron una cantidad considerablemente mayor de respuestas en la palanca operativa que en la palanca no operativa (o en la palanca de cancelación). En las condiciones de demoras intermedias (16-s y 32-s) la cantidad de respuestas en la palanca operativa fue más alta solamente cuando la segunda palanca producía la cancelación del agua. Con la demora más larga las tasas de respuesta fueron sumamente bajas y no se encontraron diferencias entre las frecuencias de respuesta de las dos palancas. Así pues el estudio de Stuphin et al. sugiere que las ratas pueden aprender dos respuestas diferentes (responder, no responder) si las demoras no exceden de 32-s.

## **ADQUISICIÓN DE LA OPERANTE LIBRE BAJO DISTINTAS CONTINGENCIAS DE REFORZAMIENTO**

Una forma prototípica de evaluar la generalidad de un hallazgo dentro del análisis experimental de la conducta, consiste en evaluarlo bajo diferentes programas de reforzamiento. Dentro de la literatura relacionada con el reforzamiento demorado se pueden encontrar diversas formas de programar la separación temporal entre la respuesta procuradora y la entrega del reforzador (véase Lattal, 1987 para una revisión). De hecho los experimentos que más cercanamente siguieron al estudio germinal de Lattal y Gleeson consistieron precisamente de estudios que evaluaron adquisición variando la forma de programar el intervalo de demora. Por ejemplo, Wilkenfield, Nikel, Blakely y Poling (1992) evaluaron la adquisición de la respuesta de palanqueo en ratas ingenuas utilizando diferentes duraciones de demora (4-s, 8-s, 16-s o 32-s) y diferentes tipos de demora (demora fija, variable y reciclante). En los programas de demora fija, el intervalo de demora se programa de acuerdo a una contingencia tipo RDO. Dicha contingencia plantea una situación en la cual la ocurrencia de un respuesta durante el intervalo de demora produce el reinicio del mismo (de ahí el nombre de demora fija). Por otro lado en los procedimientos de demora variable, el intervalo de tiempo entre la respuesta y el reforzador se programa de acuerdo con un programa de TF. Esta forma de programar la demora permite la ocurrencia de respuestas durante todo el intervalo entre la respuesta

procuradora y la entrega del reforzador (de ahí el nombre de demora variable). En la condición de demora reciclante (*stacked delay*) todas las respuestas producidas durante el intervalo de demora entregan reforzamiento una vez que ha transcurrido el periodo de demora programado.

En todos los grupos experimentales Wilkenfield et al. expusieron a ratas ingenuas privadas de alimento a programas *tándem* RF1, con intervalo de demora de diferente tipo y duración. Los resultados mostraron tasas de respuesta sustancialmente mayores en todos los grupos experimentales en relación con las producidas por sujetos expuestos a una condición de control en la cual estuvo vigente un programa de extinción. La condición que produjo las tasas más bajas fue la demora fija; en contraste, las tasas producidas por las demoras reciclantes y variables fueron muy similares. Los gradientes de demora producidos por las tres condiciones experimentales fueron muy diferentes entre sí. En tanto en los sujetos expuestos a demora fija se observó una abrupta caída de la tasa de respuesta conforme se incrementó la duración de la demora, los gradientes producidos por los sujetos expuestos a demora reciclante fueron menos inclinados. Por su parte, los gradientes producidos por los sujetos expuestos a demora variable resultaron invertidos (es decir al incrementar la demora aumentó la tasa de respuesta).

Otro investigador que exploró la generalidad de los resultados de Lattal y Gleeson utilizando programas diferentes a los autores en cuestión fue van Haaren (1992). Este científico expuso a ratas privadas de alimento a programas de entrega de comida no contingente hasta que ocurría una respuesta en la palanca. Una vez que el sujeto emitía la primera respuesta, uno de cinco diferentes programas de reforzamiento entraba en vigencia. Los sujetos podían recibir un programa de reforzamiento continuo o programas *tándem* de demora fija RF1, RDO (10-s o 30-s). Los sujetos también podían recibir programas *tándem* a los que van Haaren denominó de "demora fija variable." En estos últimos programas la duración de la contingencia RDO podía tomar distintos valores que en promedio producían una de dos duraciones (10-s o 30-s). Los resultados del investigador mostraron que las tasas de respuesta más altas ocurrieron en la condición de reforzamiento inmediato seguidas por las tasas producidas en la condición de demora breve. En general las tasas de respuesta producidas por los programas de demora fija variable fueron un poco más altas que las encontradas con demora fija. Van Haaren no solo extendió los hallazgos de Lattal y Gleeson a procedimientos de demora fija variable, el autor también comparó la ejecución de ratas de diferente sexo en todos los programas descritos encontrando que la adquisición de la respuesta con reforzamiento demorado ocurría independientemente de si los sujetos eran machos o hembras.

Otro estudio contemporáneo a los de Wilkenfield et al y van Haaren fue conducido por Dickinson, Watt y Griffiths (1992). Estos investigadores evaluaron la adquisición de la respuesta de palanqueo por ratas ingenuas privadas de alimento utilizando un

arreglo similar a un programa *tándem* con demora reciclante RF1 (duración de la demora 0-s, 2-s, 4-s o 16-s). En general los investigadores encontraron que la duración de la demora fue una función inversa de la tasa de reforzamiento obtenida por los sujetos experimentales.

En los tres estudios citados previamente se utilizó como primer componente del programa de reforzamiento un programa RF1. En contraste, Bruner, Avila y Gallardo (1994) evaluaron la adquisición de la respuesta en ratas ingenuas, privadas de alimento utilizando programas *tándem* con un primer componente intermitente IA 30-s, TF (6-s, 12-s o 24-s). Los investigadores encontraron evidencia de adquisición en todas las condiciones experimentales, adicionalmente reportaron que la tasa de respuesta fue una función inversa de la duración de la demora. El resultado de Bruner et al. contrasta con el de Wilkenfield y colaboradores en el cual la condición de demora variable mostró un gradiente inverso. Bruner, Avila y Gallardo interpretaron las diferencias entre los resultados de los dos estudios señalando que las tasas de respuesta bajas obtenidas con las demoras cortas utilizadas por Wilkenfield et al. probablemente se deben al hecho de que bajo programas de reforzamiento continuo las tasas de respuesta son muy bajas (pues el animal alterna entre presiones de la palanca y visitas al comedero). Así pues cuando Wilkenfield y colaboradores aumentaron la duración de la demora en realidad lo que hicieron fue aproximarse más a una situación de reforzamiento intermitente (en las cuales se ha documentado que las tasas de respuesta son más altas) y por lo tanto no resulta extraño que hayan producido un gradiente inverso.

## ACERCA DE LOS LÍMITES DEL FENÓMENO

En el estudio germinal de Lattal y Gleeson se puede leer entre líneas un interés por plantear a los sujetos experimentales condiciones cada vez más adversas para la adquisición de la respuesta (y un asombro ante la "inevitable" aparición de tasas de respuesta bajas pero consistentes en la mayor parte de los sujetos). De hecho es difícil leer el estudio de Lattal y Gleeson sin quedarse con la idea de que los autores están trabajando para ver cuando "tira la toalla" el sujeto experimental. En el estudio de Van Haaren citado previamente pueden advertirse condiciones bajo las cuales los animales parecen "estar tirando la toalla," específicamente la condición de demora fija de 30-s muestra tasas de respuesta sumamente bajas.

En un estudio diseñado para evaluar los límites del fenómeno de adquisición con reforzamiento demorado, Avila y Bruner (1995) expusieron a ratas privadas de alimento a programas *tándem* y encadenado IA 30-s, TF 60-s, 120-s o 180-s. Los resultados mostraron poca evidencia de adquisición en cualquiera de las condiciones experimentales.

En otro estudio similar, Snyckerski, Laraway, Byrne y Poling (1999) expusieron a ratas privadas de agua a programas RF1, RDO 60-s en los cuales una de las palancas disponibles producía agua de acuerdo con el programa descrito (y una segunda palanca cancelaba la entrega de agua si se accionaba durante el intervalo de demora). Los investigadores encontraron que algunos de los sujetos experimentales presionaron más la palanca operativa que la palanca de cancelación, sin embargo la frecuencia del palanqueo en los dos *manipulanda* fue disminuyendo conforme progresaron las sesiones (las diferencias en la tasa de palanqueo a las dos palancas también se fue haciendo menos notoria conforme transcurrieron las sesiones experimentales).

Los estudios citados sugieren que una forma en que se pueden buscar los límites del fenómeno de adquisición con reforzamiento demorado es incrementar la duración del intervalo de demora. Sin embargo también es posible evaluar sus límites "complicando" la discriminación de las consecuencias de la conducta. En un estudio realizado por Bruner, Ávila y Gallardo (1996) se sometió a ratas privadas de alimento a programas de reforzamiento en los cuales se varió la proporción de reforzadores contingentes y no contingentes. Los investigadores utilizaron un diseño factorial de 3 x 3 en donde una de las variables independientes fue la duración del intervalo de demora de un programa IA 30-s TF 6s, 12-s o 24-s). La segunda variable independiente consistió del porcentaje de reforzadores que se entregaban de manera contingente (75%, 50% o 25%). Los investigadores encontraron que las tasas de respuesta resultantes fueron una función tanto de la demora como del porcentaje de reforzadores contingentes. En general las tasas fueron más bajas mientras más larga fue la demora (lo cual coincide con los hallazgos de Dickinson et al.) y menor fue el porcentaje de reforzadores entregado de manera contingente.

### **ADQUISICIÓN COMO UNA FUNCIÓN DE LA DURACIÓN DE LA DEMORA Y LA TASA DE REFORZAMIENTO**

A la fecha se han publicado al menos cuatro revisiones relacionadas con los efectos de la demora del reforzador sobre la conducta operante (Renner, 1964; Tarpy y Sawabini, 1974; Lattal, 1987 y Pulido, Lanzagorta, Morán, Reyes y Rubí, 2004). En todas estas revisiones puede apreciarse que la forma más usual de manipular la duración de la demora del reforzador consiste en alargar la duración del TF (en programas de demora variable) o del RDO (en programas de demora fija). Manipular la duración de la demora de esta forma produce cambios concomitantes en la duración del intervalo entre reforzadores. Por en ejemplo un programa tándem IF 60-s TF 10-s es posible producir un reforzador cada 70-s, sin embargo si se incrementa la demora a 30-s el reforzador puede producirse cada 90-s. Así pues al variar la duración de la demora se varía



simultáneamente el intervalo entre reforzadores y por lo tanto la tasa de respuesta programada. La prevalencia de este problema metodológico en la literatura sobre demora de reforzamiento ha llevado a algunos investigadores a sugerir que los efectos decrementales de la demora sobre la tasa de respuesta pueden en realidad atribuirse a disminuciones concomitantes en tasa de reforzamiento programada (Baum, 1995). Esta situación ha llevado a algunos científicos a conducir experimentos en los cuales se programan las contingencias de reforzamiento de forma tal que se puede variar la duración de la demora sin modificar simultáneamente la tasa de reforzamiento programada.

En un intento por evaluar de manera independiente el efecto de la demora sobre la adquisición de la operante libre, Bruner, Pulido y Escobar (1999) llevaron a cabo el siguiente estudio. Utilizando programas definidos temporalmente de 64-s (véase Schoenfeld y Cole, 1972) variaron la posición de una ventana de tiempo ( $t^d$ ) durante la cual la primera respuesta emitida producía reforzamiento al finalizar el ciclo. Para algunos sujetos  $t^d$  se colocó al inicio del ciclo (simulando una situación de reforzamiento demorado); para otros sujetos  $t^d$  se colocó al finalizar al ciclo (simulando una situación de reforzamiento más o menos inmediato). El procedimiento en cuestión permitía separar los efectos de la demora y la tasa de reforzamiento programada debido a que la separación temporal entre la respuesta procuradora y el reforzador se podía manipular cambiando la posición de  $t^d$  dentro del ciclo (sin necesidad de cambiar la duración del ciclo de reforzamiento añadiendo un componente de TF o RDO). Los resultados mostraron tasas de respuesta consistentemente más altas en las condiciones en las cuales  $t^d$  se colocó al finalizar el ciclo de reforzamiento. Los resultados mostraron que aun cuando la tasa de reforzamiento programada y la demora no covariaron, esta última variable continuó teniendo efectos decrementales sobre la tasa de respuesta.

En otro estudio similar, Bruner, Pulido y Escobar (2000) sometieron a ratas ingenuas a programas definidos temporalmente que variaron en duración (32-s, 64-s y 128-s). Los investigadores variaron la duración de la demora colocando  $t^d$  en diferentes porciones del intervalo entre reforzadores (no únicamente al inicio o al final como en el estudio descrito previamente). Los resultados mostraron abruptos gradientes de demora en las condiciones de 64-s y 128-s; las tasas de respuesta de todas las condiciones experimentales en los ciclos de 32-s fueron homogéneamente altas. Los resultados mostraron que, al menos en ciclos de reforzamiento superiores a 32-s, la demora del reforzador tiene un efecto independiente de la tasa de reforzamiento programada.

A la fecha el estudio con mayor cantidad de variantes paramétricas, orientado a evaluar el efecto de la tasa de reforzamiento y la duración de la demora sobre la adquisición de la operante libre fue conducido por Bruner, Avila, Acuña y Gallardo (1998). En dicho estudio se sometió a ratas ingenuas, privadas de alimento, a programas

tándem IA, TF en los cuales se podía dejar fijo el primer componente y variar el segundo o por el contrario se podía variar el primer componente y dejar fijo el segundo (es decir se manipularon de manera independiente el IA y la demora del reforzador). En un primer conjunto de grupos experimentales grupos de ratas ingenuas fueron expuestas a programas RF1, IA 15-s, IA 30-s, IA 60-s e IA 120-s. En general la tasa de respuesta fue baja en RF1 y relativamente más alta en IA 15-s e IA 30-s. Las tasas de respuesta volvieron a bajar con valores de IA más altos. En un segundo conjunto de grupos experimentales se evaluó la adquisición de la respuesta de palanqueo con los mismos valores descritos previamente (RF1, IA, 15-s, 30-s, 60-s y 120-s) pero se adicionó a todos los programas un TF de 12-s. Los resultados mostraron que la tasa de respuesta disminuyó gradualmente conforme incrementó la duración del IA. En un tercer conjunto de grupos experimentales el primer componente fue siempre un programa de RF1, sin embargo la demora se varió a través de valores de 0-s, 1-s, 3-s, 6-s, 12-s y 24-s. Las tasas de respuesta producidas por las diferentes condiciones experimentales fueron muy similares, sin embargo la variabilidad de la tasa incrementó conforme aumentó la duración de la demora. En un cuarto y último conjunto de grupos experimentales el IA se mantuvo fijo en 60-s y la duración de la demora se varió a través de los mismos valores utilizando en el conjunto experimental tres. Los resultados mostraron que la tasa de respuesta disminuyó gradualmente conforme aumentó la duración de la demora. En general los resultados de Bruner et al. muestran que tanto la demora de reforzamiento como la duración del primer componente tienen efectos importantes sobre la adquisición de la operante libre. La conclusión general del estudio parecería ser la de que incrementos moderados en cualquiera de los dos componentes se asocian a tasas de respuesta altas, sin embargo si los incrementos rebasan ciertos parámetros las tasas de respuesta disminuyen.

## **VARIABLES QUE FACILITAN LA ADQUISICIÓN DE LA OPERANTE LIBRE**

Como ya se mencionó una buena parte de la investigación en fenómenos de adquisición de la operante libre se ha centrado en evaluar las condiciones, programas, tipos de respuesta y especies bajo las cuales puede ocurrir el fenómeno. Complementariamente, otros experimentos se han centrado en identificar variables que pueden facilitar (o dificultar el proceso de interés). Por ejemplo, como ya se mencionó Critchfield y Lattal (1993) y Schlinger y Blakely (1994) encontraron que los breves sonidos electromecánicos que producen los *manipulanda* convencionales facilitan la adquisición de la operante libre.

Pulido, Lanzagorta y López (2005) también evaluaron los efectos facilitadores de

las señales pero utilizando programas definidos temporalmente similares a los utilizados por Bruner, Pulido y Escobar (1999). Los investigadores expusieron a ratas ingenuas, privadas de alimento, a ciclos repetitivos de reforzamiento de 32-s en los cuales  $t^d$  se colocó al inicio o al final del intervalo entre reforzadores. En algunos grupos experimentales una señal auditiva no contingente ocurría automáticamente durante  $t^d$ ; en otros grupos la señal ocurría si el sujeto presionaba la palanca. Los resultados coincidieron con los producidos por Bruner, Pulido y Escobar, en el sentido de que las tasas de respuesta alcanzaron niveles más altos en los grupos en los cuales  $t^d$  se presentó al final del ciclo. Por otro lado, tanto las señales no contingentes como las contingentes tuvieron efectos facilitadores sobre la adquisición del palanqueo. Las señales no contingentes favorecieron tasas de respuesta bajas pero se asociaron a altos niveles de eficiencia (los sujetos respondieron poco pero casi todas sus respuestas produjeron reforzamiento). Por otro lado las señales contingentes produjeron tasas de respuesta sumamente elevadas pero una baja eficiencia respuesta-reforzador.

En los estudios reseñados hasta ahora, la duración de las señales ha sido muy breve (como en los estudios de Critchfield y Lattal y Schlinger y Blakely) o se ha variado a través de solo unos cuantos valores (Pulido, Lanzagorta y López). Con la finalidad de evaluar sistemáticamente los efectos de la duración de la señal sobre la adquisición de la operante libre, Pulido, Backer y Rubí (2003) expusieron a ratas ingenuas a programas definidos temporalmente en los cuales la primera respuesta emitida durante  $t^d$  podía producir señales de diferente duración. Los investigadores mantuvieron fija la posición de  $t^d$  al inicio del ciclo y evaluaron dos distintos valores de  $t^d$  (4-s y 8-s). En las condiciones de  $t^d$  breve la duración de la señal fue de 0-s, .5-s, 4-s, 8-s o 32-s; en las condiciones de  $t^d$  largo la duración de la señal fue de 0-s, .5-s, 8-s, 16-s y 32-s. En general los resultados mostraron una relación directa entre la duración de la señal y la tasa de respuesta. El análisis de la tasa local mostró que con señales de 8-s (o más) aparece un patrón de respuesta distintivo a través del intervalo entre reforzadores. El patrón en cuestión consiste de un incremento gradual de la tasa de respuesta que llega a su punto más alto justo antes de accionar la señal exteroceptiva; una vez accionada la señal se observa un notorio decremento en la tasa de respuesta que permanece en niveles bajos hasta que ocurre la entrega del reforzador. El patrón sugiere que las señales largas podrían promover un proceso de discriminación temporal durante el intervalo entre reforzadores (mismo que podría facilitar considerablemente una alta eficiencia respuesta-reforzador).

Además de la presencia o ausencia de señales exteroceptivas durante el intervalo de demora, diversos estudios han identificado otras variables que facilitan el proceso de adquisición de la operante libre. Por ejemplo Lattal y Williams (1997) expusieron a tres grupos de ratas ingenuas que diferían en nivel de privación (90%, 80% y 70% del

peso *ad libitum*) a un programa tándem IV 15-s RDO 30-s. Los resultados mostraron una relación inversa entre el peso experimental y la tasa de respuesta. El resultado se mantuvo aun con un diseño intrasujeto en el cual ratas fueron expuestas al mismo programa en diferentes condiciones de privación.

Una variable que ha sido identificada recientemente como facilitadora en la adquisición de la operante libre es la historia previa a la primera sesión experimental. Snycerski, Laraway, Bradley, Huitema y Poling (2004) sometieron a ratas ingenuas, privadas de agua, a diferentes tipos de historias antes de someterlas a programas RF1 o RF1, RDO 15-s. Algunas ratas fueron expuestas a los programas sin haber sido colocadas antes en el cubículo experimental; otros sujetos fueron colocados durante una sesión en el cubículo experimental pero no recibieron agua; por último otros sujetos recibieron diferente cantidad de sesiones de agua no contingente en el cubículo experimental (1 sesión o 5 sesiones) de acuerdo con un programa TV 60-s. En todas las condiciones en las cuales los sujetos experimentales fueron expuestos a la cámara de condicionamiento los investigadores compararon condiciones en las cuales la palanca estuvo presente o ausente. Los resultados mostraron que la tasa de respuesta fue baja en aquellos grupos en los cuales los sujetos no fueron expuestos a la cámara de condicionamiento antes de la primera sesión experimental. Las tasas fueron un poco mayores en los grupos que recibieron al menos una sesión (sin TV) en el cubículo experimental y llegaron a los niveles más altos en los sujetos que recibieron una o más sesiones de TV en el cubículo experimental. Los efectos fueron consistentes tanto en la condición de RF1 como en el tándem. No se encontraron efectos sistemáticos de dejar o quitar la palanca durante las sesiones de exposición a la cámara de condicionamiento. El estudio de Snycerski et al. es relevante dado que en la mayor parte de los experimentos reseñados en este trabajo, las condiciones preparatorias a las que son sometidos los animales antes de iniciar el estudio formal, son sumamente diferentes entre sí. Los datos de Snycerski y colaboradores sugieren que los procedimientos empleados no son intercambiables y podrían explicar diferencias en la magnitud de los efectos encontrados entre estudios que manipulan variables similares.

En un estudio reciente se evaluó el efecto de fármacos sobre la adquisición de la operante libre. Le Sage, Byrne y Poling (1996) evaluaron la adquisición de la respuesta de palanqueo en una palanca operativa (y la abstención de la conducta de palanqueo en una palanca no operativa) en función de la dosis de d-anfetamina que recibieron diferentes grupos de ratas ingenuas privadas de agua. Los investigadores utilizaron programas tándem en los cuales el primer componente fue siempre un programa de RF1 y el segundo componente podía ser un programa de TF o de RDO de 0-s, 8-s o 16-s. Los resultados de los investigadores mostraron que una dosis de 1 mg/kg de d-anfetamina produjo efectos facilitadores en algunos sujetos independientemente del tipo de demora

programada. Bajo dosis de 5.6 y 10 mg/kg la adquisición de la respuesta en la palanca operativa y en la palanca no operativa fue menos notoria. Los autores discutieron sus resultados en términos de la similitud de sus hallazgos con otros estudios en los cuales se utilizaron variables dependientes diferentes. Adicionalmente discutieron la posibilidad de utilizar la preparación general de adquisición con demora para evaluar los efectos de otros fármacos sobre la conducta y el aprendizaje.

## CONCLUSIÓN

Las investigaciones realizadas a la fecha sugieren que la adquisición de la operante libre con reforzamiento demorado es un fenómeno general a través de diferentes especies animales, procedimientos experimentales y tipos de respuestas. Así pues la idea preconcebida de que la adquisición de la respuesta solamente puede ocurrir utilizando la técnica de moldeamiento por aproximaciones sucesivas y reforzamiento inmediato no parece tener sustento empírico. Las investigaciones realizadas a la fecha, sin embargo, sugieren que sí existen límites en las condiciones bajo las cuales es posible la adquisición de la operante libre con reforzamiento demorado. Por lo menos dos de los estudios reseñados en este trabajo sugieren que los límites podrían ser cercanos a un minuto (Avila y Bruner, 1995; Snyckerski, Laraway, Byrne y Poling, 1999). Por supuesto este límite debe tomarse con reservas pues estudios clásicos han mostrado que los gradientes de demora generados en cajas de condicionamiento son considerablemente más abruptos que los generados en laberintos (Perin, 1943a, 1943b). Así es posible que en preparaciones diferentes a la caja de condicionamiento tradicional sea posible evidenciar la adquisición de la respuesta con demoras de reforzamiento superiores a un minuto.

Para los presentes autores, la idea de que la aparición de la respuesta solamente puede ocurrir con reforzamiento inmediato resulta intuitivamente menos atractiva que la idea opuesta. Los organismos enfrentan cotidianamente situaciones de las que deben aprender para poder sobrevivir y en las cuales el reforzamiento inmediato es más la excepción que la regla (Baum, 1995).

Los estudios reseñados en este estudio sugieren que a pesar de que demora de reforzamiento y tasa de reforzamiento programada se encuentran confundidas en la mayor parte de los estudios, ambas variables tienen efectos independientes importantes sobre la adquisición de la operante libre. El efecto reportado de manera más habitual de la demora del reforzador consiste en una relación inversa entre duración de la demora y tasa de respuesta. Por otro lado los efectos de la tasa de reforzamiento sobre la tasa de respuesta en estudios sobre adquisición parecen describirse mediante una función de "U" invertida en la cual las tasas son bajas con intervalos entre reforzadores "demasiado breves" o "demasiado largos."

En lo concerniente a las variables que modulan la adquisición de la operante libre los resultados reportados a la fecha sugieren que la presentación de una señal durante el intervalo de demora tiene importantes efectos facilitadores sobre el aprendizaje. Al parecer los efectos en cuestión podrían deberse tanto al papel de reforzador condicionado que puede desempeñar las señales como a su capacidad para constituirse en estímulos discriminativos que permiten maximizar la eficiencia respuesta-reforzador. El efecto facilitador de algunas variables sobre los fenómenos de adquisición no resulta sorprendente (tal es el caso de la relación entre nivel de privación y tasa de respuesta reportado por Lattal y Williams). Sin embargo los hallazgos de Snycerski et al. relativos a los efectos que pueden tener las condiciones preparatorias sobre la adquisición de la operante libre sí parecen novedosos (dado que es casi imposible encontrar dos estudios sobre adquisición en los cuales las condiciones preparatorias sean comparables). La heterogeneidad en las condiciones preparatorias que se encuentra en la literatura del área sugiere que los autores de los estudios ignoran la importancia de las mismas para la adquisición de la operante libre.

Por último, el hecho de que se haya descubierto que bajo ciertas dosis algunos fármacos pueden tener efectos facilitadores (o inhibidores) sobre la adquisición de la operante libre con reforzamiento demorado, abre la puerta a la experimentación de diferentes sustancias utilizando la variable dependiente de interés.

Es probable que el lector se haya percatado a través de la revisión que los criterios seleccionados para determinar que ocurrió (o no) la adquisición se manejan de una forma muy laxa. La responsabilidad del manejo laxo del concepto puede en parte atribuirse a la falta de consenso por parte de los investigadores revisados con respecto al desarrollo de criterios basados empírica y estadísticamente que permitan definir operacionalmente el fenómeno. Así uno de los aspectos que deben resolverse de manera prioritaria en la agenda de investigación del área es la de llegar a consensos aceptados que permitan determinar cuándo ha ocurrido la adquisición de la operante libre (cabe mencionar que Snycerski et al (2004) han realizado algunas propuestas interesantes para definir operacionalmente el concepto, sin embargo difícilmente se puede hablar de un consenso dado que estos autores son los únicos que utilizan los criterios de adquisición desarrollados por ellos mismos).

Otro pendiente importante en la agenda de investigación tiene que ver con lo que algunos investigadores han dado en llamar el nivel operante (véase Premack y Schaeffer, 1962;1963). En general por nivel operante se entiende aquel comportamiento que surge "espontáneamente" y que no puede atribuirse a las contingencias de reforzamiento manipuladas por el investigador. El nivel operante parece crucial para los procesos de adquisición pues los sujetos experimentales deben presentar respuestas "espontáneas" para poder exponerse a las contingencias programadas por el investigador. Así pues un

segundo pendiente importante en la agenda de investigación del área consiste en identificar las variables que modulan la frecuencia de las respuestas “espontáneas.”

Un tercer pendiente para los investigadores interesados en adquisición con demora es la difusión de sus hallazgos. Una cantidad importante de estudios de corte farmacológico se llevan a cabo todavía con procedimientos de ensayo discreto, principalmente porque para muchos investigadores de orientación farmacológica la variable dependiente de interés es la adquisición y no el mantenimiento (Antón, 2005 comunicación personal). Si los procedimientos de adquisición de la operante libre se difunden fuera del análisis experimental de la conducta sería posible ofrecer a otros científicos una alternativa a los procedimientos de ensayo discreto cuyos problemas ya han sido señalados ampliamente por Skinner (1956).

## REFERENCIAS

- Avila, R. y Bruner, C. (1995). Adquisición de la respuesta bajo demoras largas de reforzamiento señalado y no señalado. *Revista Mexicana de Análisis de la Conducta*, 21, 117-127.
- Baum, W.M. (1995) Introduction to molar behavior analysis. *Revista Mexicana de Análisis de la Conducta*, 21, 17-35.
- Bolles, R.C. (1988). The bathwater and everything. *Behavioral and Brain Sciences*, 11, 449-450.
- Bruner, C., Avila, R. y Gallardo, L. (1994). Acquisition of lever pressing in rats under an intermittent Schedule of delayed reinforcement. *Revista Mexicana de Análisis de la Conducta*, 20, 119-129.
- Bruner, C., Avila, R. y Gallardo, L. (1996) Acquisition with delayed reinforcement under combinations of response dependent and independent reinforcement. *Revista Mexicana de Análisis de la Conducta*, 22, 29-39.
- Bruner, C., Ávila, R., Acuña, L., y Gallardo, L., (1998). Effects of reinforcement rate and delay on the acquisition of lever pressing by rats. *Journal of the Experimental Analysis of Behavior*, 69, 59-75.
- Bruner, C., Pulido, M. y Escobar, R. (1999). Response acquisition and maintenance with a temporally defined schedule of delayed reinforcement. *Revista Mexicana de Análisis de la Conducta*, 25, 379-391.
- Bruner, C., Pulido, M. y Escobar, R. (2000). La adquisición del palanqueo con programa temporales de reforzamiento demorado. *Revista Mexicana de Análisis de la Conducta*, 26, 91-103.
- Critchfield, T.S. y Lattal, K.A. (1993). Acquisition of a spatially defined operant with delayed reinforcement. *Journal of the Experimental Analysis of Behavior*, 59, 373-387.
- Dickinson, A., Watt, A. y Griffiths, W.J.H. (1992). Free-operant acquisition with delayed reinforcement. *The Quarterly Journal of Experimental Psychology*, 45B 241-258.
- Hull, C.L. (1952). *A behavior system*. New Haven: Yale University Press.
- Lattal, K.A. (1987). The effect of delay and of intervening events on reinforcement value. En M.L. Commons, J.E. Mazur, J.A. Nevin y H. Rachlin (Eds.): *Quantitative Analysis of Behavior* (vol 5). New Jersey: Lawrence, Erlbaum Associates Publisher.
- Lattal, K.A. y Gleeson, S. (1990). Response acquisition with delayed reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes*, 16, 27-39.

- Lattal, K.A. y Metzger, B. (1994). Response acquisition by Siamese fighting fish. *Journal of the Experimental Analysis of Behavior*, 61, 35-44.
- Lattal, K.A. y Williams, A.M. (1997). Body weight and response acquisition with delayed reinforcement. *Journal of the Experimental Analysis of Behavior*, 67, 131-144.
- Le Sage, M.G., Byrne, T. y Poling, A. (1996). Effects of d-amphetamine on response acquisition with immediate and delayed reinforcement. *Journal of the Experimental Analysis of Behavior*, 66, 349-367.
- Perin, C.T. (1943a) The effect of delayed reinforcement upon differentiation of bar responses. *Journal of Experimental Psychology*, 32, 95-109.
- Perin, C.T. (1943b) A quantitative investigation of the delay of reinforcement gradient. *Journal of Experimental Psychology*, 32, 110-123.
- Premack, D. y Schaeffer, R.W. (1962). Distributional properties of operant-level locomotion in the rat. *Journal of the Experimental Analysis of Behavior*, 5, 89-95.
- Premack, D. y Schaeffer, R.W. (1963). Some parameters affecting the distributional properties of operant-level running in rats. *Journal of the Experimental Analysis of Behavior*, 6, 473-475..
- Pulido, M., Backer, C. y Rubi, M. (2003). *Response acquisition with signaled delay of reinforcement: Effect of signal duration*. Trabajo presentado en la XXIX Convención Anual de la Association for Behavior Analysis.
- Pulido, M., Lanzagorta, N., Morán, E., Reyes, A. y Rubí, M. (2004) El efecto de las señales en programas de reforzamiento demorado: Una revisión contemporánea. *Revista del Consejo Nacional para la Enseñanza e Investigación en Psicología*, 9, 321-339.
- Pulido, M., Lanzagorta, N., y López, L., (2005) Effects of contingent and non-contingent signals during delay interval on response acquisition by rats. Enviado para dictamen a la *Revista Mexicana de Análisis de la Conducta*.
- Renner, K.E. (1964). Delay of reinforcement: A historical review, *Psychological Bulletin*, 61, 341-361.
- Savage, T. (2001) Shaping: A multiple contingencies analysis and its relevance to behaviour-based robotics. *Connection Science*, 13, 199-234.
- Schlinger, H.D. y Blakely, E. (1994). The effects of delayed reinforcement and a response-produced auditory stimulus on the acquisition of operant behavior in rats. *Psychological Record*, 44, 391-409.
- Schoenfeld, W.N., y Cole, B.K. (1972). *Stimulus schedules: The t-T systems*. New York, Harper and Row.
- Schwartz, B and Gamzu, E. (1977). Pavlovian control of operant behavior. En W.K. Honig y J.E.R. Staddon (Eds.) *Handbook of operant behavior*. Prentice-Hall, Englewood Cliffs, NJ, 53-97.
- Sidman, M. (1960). *Tactics of scientific research*. New York: Basic Books.
- Skinner, B.F. (1951). How to teach animals. *Scientific American*, 185, 26-29.
- Skinner, B.F. (1953). *Science and human behavior*. New York: McMillan.
- Skinner, B.F. (1956). A case history in scientific method. *American Psychologist*, 2, 221-233.
- Snycerski, S., Laraway, S., Byrne., T. y Poling, A. (1999). Acquisition of lever-press responding with delayed consequences in rats: Is a minute to long? *Revista Mexicana de Análisis de la Conducta*, 25, 341-350.
- Snycerski, S., Laraway, S., Bradley, E., Huitema., y Poling, A. (2004). The effects of behavioral history on response acquisition with immediate and delayed reinforcement. *Journal of the Experimental Analysis of Behavior*, 81, 51-64.



- Stuphin, G., Byrne, T., y Poling, A. (1998). Response acquisition with delayed reinforcement: A comparison of two-lever procedures. *Journal of the Experimental Analysis of Behavior*, 69, 17-28.
- Tarpy, R.W. y Sawabini, F.L. (1974) Reinforcement delay: A selective review of the last decade. *Psychological Bulletin*, 81, 984-997.
- van Haaren, F. (1992). Response acquisition with fixed and variable resetting delays of reinforcement in male and female Wistar rats. *Physiology and Behavior*, 55, 769-772.
- Wilkenfield, J., Nickel, M., Blakely, E. y Poling, A. (1992). Acquisition of lever-press responding in rats with delayed reinforcement: A comparison of three procedures. *Journal of the Experimental Analysis of Behavior*, 58, 431-443.

## RESUMEN

Estudios recientes han mostrado que la adquisición de la operante libre bajo condiciones de reforzamiento demorado y en ausencia de moldeamiento explícito es posible. Esta revisión presenta la investigación generada dentro del área. Se presentan estudios que han evaluado la generalidad del fenómeno, sus límites y las variables que lo modulan. La revisión sugiere que el fenómeno es general y robusto. La revisión también sugiere que demoras cercanas a un minuto y mezclas de reforzamiento contingente y no contingente dificultan considerablemente el fenómeno. La presencia de señales exteroceptivas durante el intervalo de demora así como la entrega de alimento no contingente antes de iniciar la primera sesión experimental facilitan el aprendizaje. Se sugiere llevar a cabo la evaluación empírica de diferentes definiciones operacionales del concepto; también se sugiere analizar sistemáticamente los efectos que las diferencias en los distintos procedimientos experimentales tienen sobre los fenómenos de adquisición.

Descriptores: Adquisición, demora de reforzamiento, operante libre, animales, revisión.

## ABSTRACT

Recent studies have shown that the acquisition of free operant responding may occur under conditions of delayed reinforcement and without explicit shaping. The purpose of the present review is to coherently present the studies generated on the subject. The studies conducted so far suggest the phenomenon is general across a number of different animal species and procedures. The studies also suggest the phenomenon is less apparent with delay durations in the vicinity of one minute and with combinations of contingent and non contingent reinforcers. Exteroceptive stimuli presented during the delay interval and non-contingent food delivered before the exposure to the schedule considerably enhance response acquisition. As a future research agenda, this review suggests different operational definitions of response acquisition be empirically tested; additionally procedural differences between the studies conducted so far should be systematically studied.

Key words: Acquisition, delayed reinforcement, free operant responding, animals, review.