

# Capítulo 1

## El efecto de la respuesta recién reforzada: ¿una cuestión de señal o fuerza?

*Sarah Cowie,<sup>1</sup> Michael Davison y Douglas Elliffe*

SCHOOL OF PSYCHOLOGY  
UNIVERSITY OF AUCKLAND

### Resumen

Los efectos locales de los reforzadores sobre la conducta han jugado un rol central incentivando el debate sobre cómo los reforzadores afectan la conducta. Cambios en la elección en una escala temporal pequeña tienden a señalar los cambios en la probabilidad de obtener el siguiente reforzador para determinada respuesta, independientemente de qué respuesta ha sido reforzada recientemente, sugiriendo un control ejercido por la estructura del ambiente y lo que un reforzador predice acerca de subsecuentes relaciones conducta-ambiente. Literatura en desarrollo muestra que esta aproximación prospectiva del control del reforzador ofrece una alternativa viable a la interpretación tradicional, la del fortalecimiento. Sin embargo, aunque los análisis a nivel local muestran un efecto principal fuerte que es consistente con la idea de señalización, estos también muestran un efecto considerablemente menor que parece, a primera vista, más consistente con el fortalecimiento. El control del reforzador no puede ser entendido

---

1 Correspondencia: Dirigirla a Sarah Cowie, correo: <sarah.cowie@auckland.ac.nz>.

apropiadamente sin el conocimiento de los orígenes de este efecto de la respuesta recién reforzada. En este capítulo usamos ejemplos de documentos publicados para explorar el efecto de la respuesta recién reforzada y discutir si el efecto es consistente con la función de los reforzadores de fortalecer la respuesta.

*Palabras clave:* reforzador, ambiente, perspectiva prospectiva, perspectiva del fortalecimiento, respuesta-recién-reforzada

## Abstract

The fine-grained, local-level effects of reinforcers on behavior have played a key role in driving debate about how reinforcers affect behavior. Changes in choice on a small timescale tend to track changes in the likelihood of obtaining the next reinforcer for a given response, irrespective of which response has most recently been reinforced, suggesting control by the structure of the environment and what one reinforcer predicts about subsequent behavior-environment relations. A developing literature shows this prospective account of reinforcer control offers a viable alternative to the traditional, strengthening interpretation. Nevertheless, although local-level analyses show a strong main effect that is consistent with a signaling account, they also display a much smaller affect that seems, at first sight, more consistent with strengthening. Reinforcer control cannot be properly understood without knowledge of the origins of this just-reinforced-response effect. In this chapter we use examples from published literature to explore the just-reinforced-response effect, and discuss whether the effect is consistent with a response-strengthening function of reinforcers.

*Keywords:* reinforcer, environment, prospective account, strengthening account, just-reinforced-response.

## Introducción

Las consecuencias de nuestro comportamiento son fundamentales para determinar si es probable que una conducta<sup>2</sup> se repita o no en el futuro. La elección

---

2 El comportamiento se puede medir de diferentes maneras; la tasa de respuesta, la asignación de tiempo o respuestas entre alternativas (elección) y la latencia de respuesta después de un evento son medidas comunes. Si bien todas estas medidas parecen ser sensibles al tiempo y ubicación probables de los refor-

entre conductas favorece la opción con mayor probabilidad de producir un reforzador, ya sea que esa opción sea consistentemente más probable (e.g., Baum, 1974; Herrnstein, 1961), o simplemente más probable en el momento (e.g., Davison & Baum, 2002; Cowie *et al.*, 2013). Aunque tradicionalmente se pensaba que los reforzadores<sup>3</sup> aumentaban la probabilidad de que ocurriera una conducta porque se fortalecía (Skinner, 1938), perspectivas más recientes aseguran que es el control por la relación entre los eventos en el ambiente de un organismo (Baum, 2012; Cowie & Davison, 2016; Shahan, 2017). Los eventos que normalmente denominamos reforzadores (alimento para un organismo hambriento, agua para uno sediento) son eventos filogenéticamente importantes que forman parte de la estructura del ambiente (Baum, 2012; Davison & Baum, 2006). El efecto de un reforzador sobre la conducta no se deriva de una función única (es decir, no porque altere la fuerza de la conducta o las conexiones), sino de su capacidad para señalar las condiciones ambientales que pueden seguir, de la misma manera que cualquier otro componente del ambiente (incluidos los estímulos y las conductas) con los que pueden estar asociados, y, por lo tanto, llegar a señalar, o en el lenguaje de Shahan (2010), proporcionar señales para futuras condiciones particulares. Cualquier evento –estímulo, respuesta o reforzador– que normalmente es seguido por más reforzadores para una conducta en particular, aumentará la ocurrencia de dicha conducta. Cualquier evento que normalmente es seguido por la ausencia de reforzadores para cierta conducta resultará en una disminución de ese comportamiento. Este marco más reciente, basado en la estructura para

---

zadores, el efecto de un reforzador en la conducta que le sigue está determinado en gran medida por sus propiedades en relación con las de los reforzadores que siguen a otras conductas (e.g., Herrnstein, 1961; 1970); es decir, la decisión de responder depende del valor relativo, más que absoluto, de la recompensa por responder. Como señaló Herrnstein (1970), toda conducta es una elección –incluso cuando se dispone de una sola respuesta medida, los animales pueden no hacer nada, e incluso no hacer nada es en sí mismo un comportamiento, que es precisamente lo que hacen entre la emisión de las respuestas definidas por el experimentador–. La medición de la tasa de respuesta (o la latencia para responder) de forma aislada no capta la importancia de los beneficios de responder en contexto, y, como resultado, puede ser insensible a las variables que afectan la elección (e.g., Catania, 1963; Neuringer, 1967). En consecuencia, tomamos la elección como la medida fundamental de la conducta.

- 3 En este manuscrito, usamos el término “reforzador” para denotar una cosa (un evento, probablemente de importancia filogenética para el organismo) más que un proceso. Aunque el término “reforzador” es problemático en sus connotaciones, su significado como sustantivo, en referencia a comida para una paloma hambrienta, agua para una rata sedienta, un juguete favorito para un niño aburrido, generalmente se entiende bien.

comprender la conducta, ha sido discutido en términos de *control de estímulos* (Cowie & Davison, 2022), *señalización* (Cowie *et al.*, 2021), *inducción*<sup>4</sup> y una *vía ambiental* (Tonneau, 2001).

Este cambio de paradigma del fortalecimiento retrospectivo del reforzador a un control más bien prospectivo mediante la señalización ha sido impulsado en gran medida por la observación de los efectos a nivel local de reforzadores primarios (e.g., Cowie *et al.*, 2011; véase Cowie & Davison, 2016) y condicionales (Boutros *et al.*, 2009; Davison & Baum, 2006; 2010; ver Shahan, 2010, para una revisión). En modelos animales (e.g., Cowie *et al.*, 2011) y –al menos para los reforzadores primarios– en experimentos con niños con un desarrollo típico (Cowie *et al.*, 2021a, b), los cambios de elección en una escala temporal pequeña tienden a seguir los cambios en la probabilidad de obtener el *próximo* reforzador para una determinada respuesta. Para que un reforzador produzca un aumento en la probabilidad de la conducta que sigue, ese reforzador debe haber sido correlacionado con la disponibilidad futura –en lugar de la no disponibilidad– de reforzadores para la misma respuesta (Cowie *et al.*, 2011; Davison & Baum, 2006; 2010; Krägeloh *et al.*, 2005) en la propia historia de aprendizaje del organismo de forma no redundante y discriminable (Boutros *et al.*, 2009), o bien, debe haber sido correlacionada con reforzadores posteriores para la misma respuesta en la historia filogenética y evolutiva del organismo (Baum, 2012). Un elemento crítico de este control es cómo se traduce la experiencia actual con una estructura ambiental en la estructura *percibida*; cuanto más cercana sea la estructura percibida del ambiente a su estructura actual, mayor será el control de la contingencia (ver Cowie & Davison, 2016; Cowie 2018; Cowie & Davison, 2022 para una discusión).

Literatura actualmente en desarrollo muestra que estas explicaciones más recientes, basadas en la estructura, ofrecen una alternativa viable a la interpretación tradicional de que la entrega de un reforzador tiene como efecto principal fortalecer la conducta. Sin embargo, puede haber más efectos menores que se explican de manera menos obvia de esta manera. Estos pueden revelarse mediante análisis a grano fino y a nivel local que muestran

---

4 Por supuesto, los marcos específicos de inducción y señalización/control de estímulo difieren, al menos en algún grado. Su terreno en común es que predicen el control sobre la base de la estructura ambiental –la relación entre eventos en el tiempo, el lugar y otras dimensiones– y generan reforzadores como eventos que son distintos de los demás porque son filogenéticamente importantes, usualmente teniendo un valor hedónico, más que porque tengan una función única en el cambio de comportamiento.

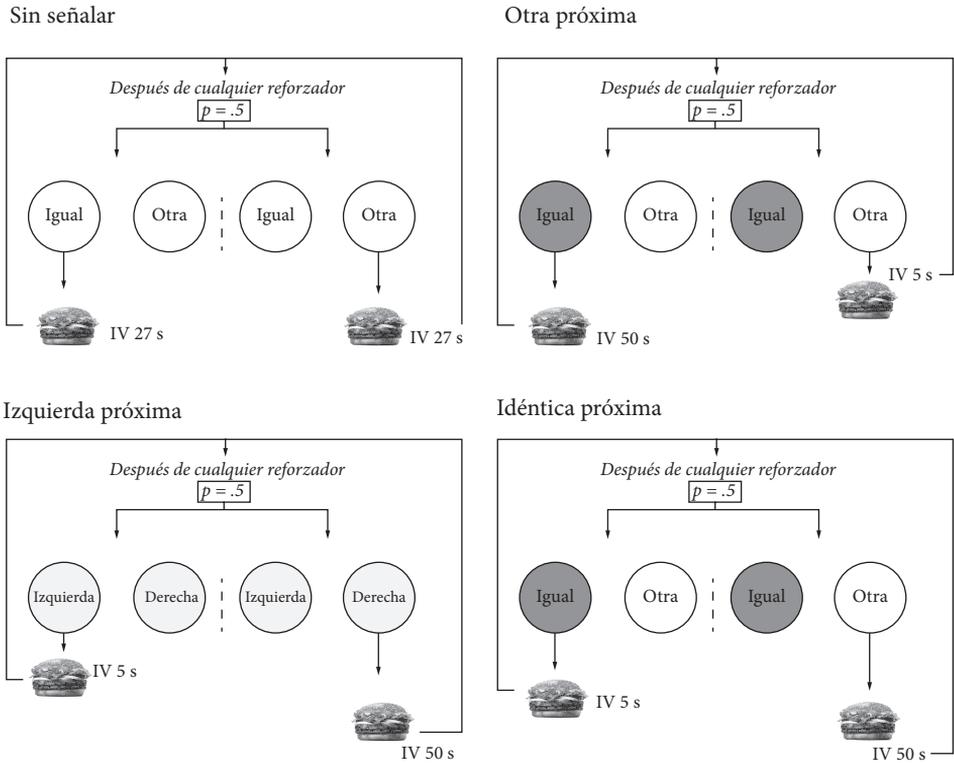
cambios en la preferencia a lo largo del tiempo desde la entrega del reforzador. En el ejemplo que discutimos en este capítulo, tales análisis muestran un fuerte efecto principal que es consistente con una explicación de señalización, pero también un efecto mucho más pequeño que parece, a primera vista, más consistente con el fortalecimiento. Por ejemplo, Cowie *et al.* (2011) implementaron, utilizando palomas como sujetos y comida como reforzador, un procedimiento de elección concurrente en el que la probabilidad de reforzar las respuestas a las teclas izquierda y derecha cambiaba en función del tiempo transcurrido desde el reforzador anterior, independientemente de qué tecla había producido ese reforzador anterior. En una condición, llamada izquierda-pronto, derecha-después, un reforzador de la tecla izquierda era mucho más probable durante los primeros 30 s después de recibir un reforzador en cualquier tecla, pero transcurridos los 30 s esas probabilidades se revirtieron. En otra condición, llamada derecha-pronto, izquierda-después, los reforzadores de la tecla derecha eran mucho más probables durante los primeros 30 s, y los de la tecla izquierda eran más probables después de los 30 s. El patrón de cambio de la elección después del reforzamiento mostró un fuerte efecto de los cambios en la probabilidad de reforzamiento en las dos teclas, lo que sugiere que el efecto principal se explica con una interpretación basada en la estructura. No obstante, la función de elección se desplazó un poco hacia la respuesta que produjo el reforzador más reciente. Este efecto de *última-respuesta-reforzada* es consistente con una función de fortalecimiento de la respuesta, y sus orígenes deben ser entendidos para que, a su vez, el control del reforzador se entienda correctamente.

En este capítulo, exploramos la naturaleza del efecto de respuesta-recién-reforzada, utilizando como ejemplo la elección en condiciones seleccionadas a partir de un conjunto de datos publicados, para explorar el impacto de los reforzadores individuales y las secuencias de reforzadores en la elección. A continuación, discutimos qué revelan estos efectos sobre *cómo* los reforzadores controlan la conducta. El conjunto de datos que hemos utilizado para ilustrar el efecto de respuesta-recién-reforzada proviene de un experimento con palomas reportado por Cowie *et al.* (2011). En este experimento, cada ensayo dentro de una serie de ensayos terminó en un solo reforzador dispuesto en un programa de Intervalo-Variable (IV). El programa con el que se entregaba el reforzador y la respuesta que lo produciría, diferían según la contingencia vigente. Una vez que se obtuvo un reforzador, el ensayo terminó;

el siguiente ensayo comenzaba inmediatamente después de ese reforzador. La mitad de todos los reforzadores se entregaron para respuestas dadas en la tecla izquierda y la otra mitad para respuestas en la tecla derecha. En todas las condiciones, excepto en la condición Sin Señal, la respuesta con mayor probabilidad de producir el siguiente reforzador fue discriminable sobre la base del tiempo transcurrido desde el comienzo del ensayo –es decir, el tiempo transcurrido desde el reforzador anterior–, y en algunas condiciones, desde la respuesta que había producido el reforzador en el ensayo anterior. No se programó ninguna demora de cambio. Las condiciones seleccionadas representan una amplia gama de contingencias (Figura 1; ver también Cowie *et al.*, 2011, para mayor detalle); una condición de Sin Señal en la que los reforzadores fueron igualmente probables para las respuestas a las teclas izquierda y derecha, una condición Izquierda-Pronto en la que los reforzadores de la tecla izquierda se proporcionaban en un programa IV de 5-s y las respuestas a la tecla derecha en un programa IV de 50-s; una condición de Otra-Pronto en la que un reforzador fue entregado según un programa de IV 5-s si fue producido por la respuesta que *no* fue reforzada en el ensayo anterior y con un programa de IV 50-s *si fue* producido por la respuesta que fue reforzada en el ensayo anterior, y una condición La Misma-Pronto en la que se dispuso un reforzador en un programa IV de 5 s si fue producido por la misma respuesta que fue reforzada en el ensayo anterior, y en un programa IV de 50 s si fue producido por la respuesta que *no* se reforzó en el ensayo previo. En las condiciones Misma-Pronto y Otra-Pronto, la respuesta que proporcionaba el último reforzador se señalaba con una tecla de luz roja. Elegimos estas condiciones sobre aquellas en las que la respuesta no fue señalada porque estas condiciones proporcionaron la mejor demostración de control discriminativo y, por lo tanto, fueron las mejores para ilustrar la interacción entre los efectos discriminativos y de respuesta-recién-reforzada. Nos enfocamos en la conducta estable, después de una exposición prolongada a la misma estructura ambiental, usando datos de las últimas 55 de 85 sesiones de cada condición. Seleccionamos estas condiciones porque los datos demuestran efectos que se han tomado para respaldar una aproximación de control estructural, así como un patrón de efectos de respuesta-recién-reforzada, que refleja aquellos observados en otros conjuntos de datos donde el control estructural y de fortalecimiento es más difícil de separar.

**Figura 1**

Diagrama de contingencias de las condiciones seleccionadas por Cowie et al. (2011)

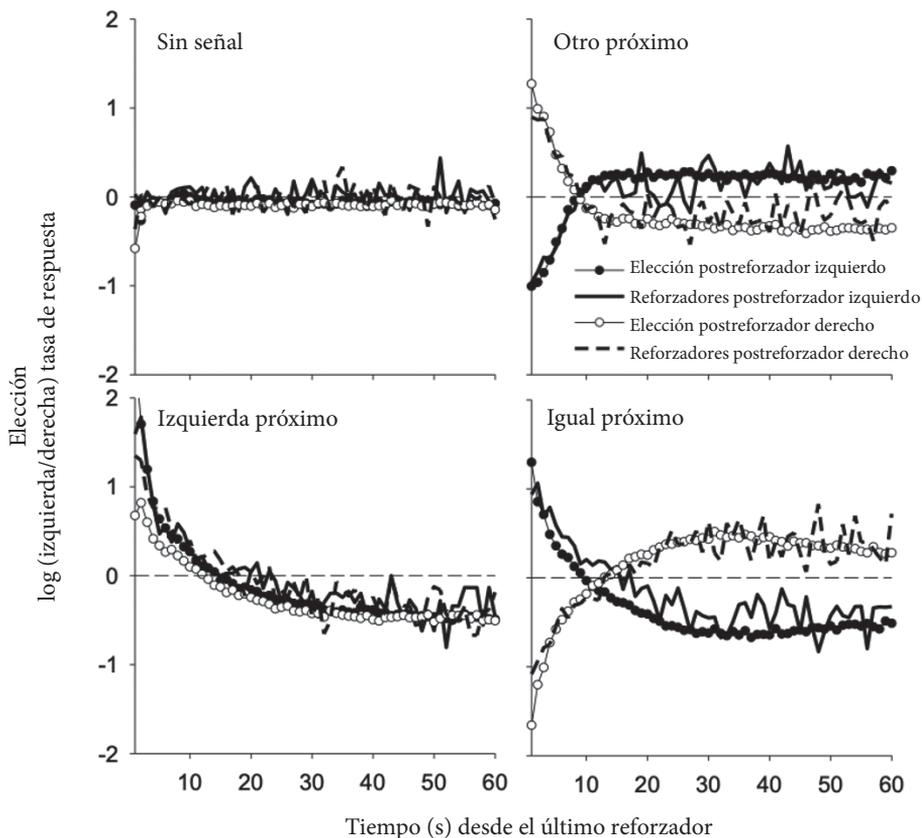


## Efectos del reforzador individual

Los análisis de *elección-local* (Davison & Baum, 2002) muestran cómo la entrega de un solo reforzador afecta la asignación de conducta entre dos alternativas. Estos análisis representan una medida de elección (agregada a través de muchos casos de entregas de reforzadores) como una función del tiempo desde el reforzador más reciente, de forma separada para respuestas que ocurren después de cada tipo de reforzador. La razón de reforzadores locales –calculada a partir de los reforzadores obtenidos en cada momento desde el reforzador más reciente, por separado, para cada tipo de reforzador– ilustra la relación entre el

tiempo desde un reforzador, la elección y los reforzadores. Efectivamente, esta razón cambiante de reforzadores locales representa la estructura actual del ambiente, tanto en términos de tiempo como de lugar de la respuesta, que resultó de los programas de refuerzo descritos anteriormente. Es decir, documentan qué respuesta era más probable que produjera un reforzador en un momento dado. Cuanto más cerca esté la función de elección de la función de reforzador, mejor será el control por parte de la estructura del ambiente.

**Figura 2**  
*Elección-local y funciones de reforzamiento*



La Figura 2 muestra la elección local –círculos llenos y vacíos– y las funciones del reforzador –líneas sólidas y discontinuas; la razón logarítmica de la respuesta izquierda-derecha y las tasas de entrega del reforzador– como funciones de tiempo desde un reforzador de la tecla izquierda o derecha, de las condiciones seleccionadas de Cowie *et al.*, (2011). En los cuatro paneles, la elección –círculos– siguió el logaritmo de la razón de reforzadores –líneas– a lo largo del intervalo entre reforzadores –es decir, el patrón de cambio en la elección a lo largo del tiempo en la Figura 2 fue muy similar al patrón de cambio en las razones de reforzadores a lo largo del tiempo–, indicando un fuerte control por parte de la estructura del ambiente. La conducta de elección del sujeto rastreando la disponibilidad cambiante de los reforzadores a través del tiempo y las ubicaciones es un hallazgo típico en este nivel de análisis local. La excepción es cuando la disponibilidad de reforzadores cambia con el tiempo de una manera compleja y no monótonica, que requiere una discriminación precisa del tiempo transcurrido (e.g., Miranda-Dukoski *et al.*, 2014), o cuando la dirección del cambio depende de la ubicación del último reforzador, lo cual requiere de memoria de un evento pasado que pudo ser olvidado (Cowie *et al.*, 2011; 2017). Cuando el control a nivel local requiere una discriminación o memoria demasiado precisa, la elección permanece constante a lo largo del tiempo, en un nivel consistente con los reforzadores agregados en contextos de estímulo separados –e.g., agregados en una escala de tiempo más larga o en intervalos iniciados por un reforzador, ya sea en el lado izquierdo o derecho–.

La estrecha adherencia de la elección a las razones cambiantes de reforzamiento, significa que el patrón de elección local difiere según el último reforzador solamente cuando el patrón de disponibilidad del reforzador –log de las razones del reforzador– difiere según el último reforzador. Este efecto se ilustra en la Figura 2; las funciones del reforzador –líneas discontinuas y sólidas– son diferenciales con respecto al reforzador precedente solo en las condiciones Misma-Pronto y Otra-Pronto y, en consecuencia, solo en estas condiciones la elección sigue un patrón diferente dependiendo de si el reforzador precedente fue producido por una respuesta izquierda o derecha. Como resultado de los diferentes patrones de elección, el efecto de respuesta-recién-reforzada es difícil de discernir en estas condiciones; en cambio, se ilustra claramente en las condiciones Izquierda-Pronto y –en menor medida– Sin-Señal, donde la disponibilidad de reforzadores no difería según el reforzador precedente. En las condiciones Izquierda-Pronto y Sin-Señal, el *patrón* de cambio en la elec-

ción local es exactamente el mismo después de los reforzadores izquierdo y derecho, pero el *nivel* de elección –qué tan extremo es el logaritmo de razón de respuesta– difiere, con funciones de elección ligeramente desplazadas hacia la respuesta que produjo el reforzador precedente. Es decir, la elección después de una entrega de comida a la izquierda se desplaza hacia respuestas a la izquierda, y la elección después de una entrega de comida a la derecha hacia respuestas a la derecha<sup>5</sup>. Este efecto de respuesta-recién-reforzada es más evidente en tiempos más cortos desde un reforzador, y más prominente en la condición de Izquierda-Antes que en la condición No-Señal. Si bien el desvanecimiento del efecto a lo largo del tiempo podría atribuirse a un efecto de fortalecimiento transitorio producido por el reforzador que precedió a la prueba, la diferencia en la magnitud de este efecto entre las dos condiciones obviamente no es comprensible en tales términos.

## Secuencias de reforzador

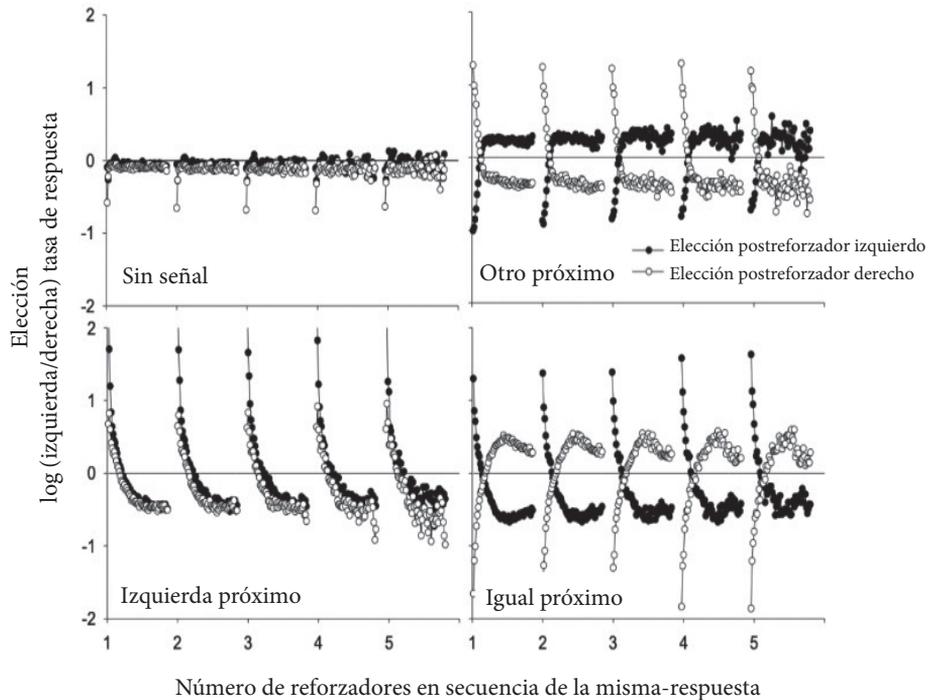
El efecto de respuesta-recién-reforzada es acumulativo entre entregas sucesivas de alimento para la misma respuesta (e.g., Davison & Baum, 2000; Landon, Davison & Elliffe, 2002). Para ilustrar, la Figura 3 muestra la misma medida de elección que en la Figura 2 –el logaritmo de la razón de respuesta graficada como una función del tiempo desde un reforzador–, pero siguiendo secuencias de uno a cuatro reforzadores producidos por la misma respuesta. La Figura 3 muestra que el patrón de cambio en la elección a lo largo del intervalo entre comidas, después de secuencias de dos a cinco entregas de comida con la misma tecla, fue similar a la que siguió a una sola entrega de comida (Figura 2), pero la elección se desplaza cada vez más hacia la respuesta-recién-reforzada con cada reforzador sucesivo de la misma tecla (Figura 3). No obstante, incluso después de secuencias relativamente largas de reforzadores para la misma respuesta, el efecto principal es que la elección local continúa siguiendo los cambios en la razón local de comida.

---

5 Por brevedad, no hemos presentado datos para la condición paralela Derecha-Pronto. El mismo patrón de elección cambiante a lo largo del tiempo también fue evidente en esa condición –la elección inmediatamente después de la entrega del alimento se movió fuertemente hacia la tecla derecha, independientemente de qué respuesta había producido esa entrega del alimento, se movió hacia la tecla izquierda con el paso del tiempo, pero en todos los puntos en el tiempo la elección se desplazó levemente hacia la alternativa que produjo la última entrega de alimentos–.

**Figura 3**

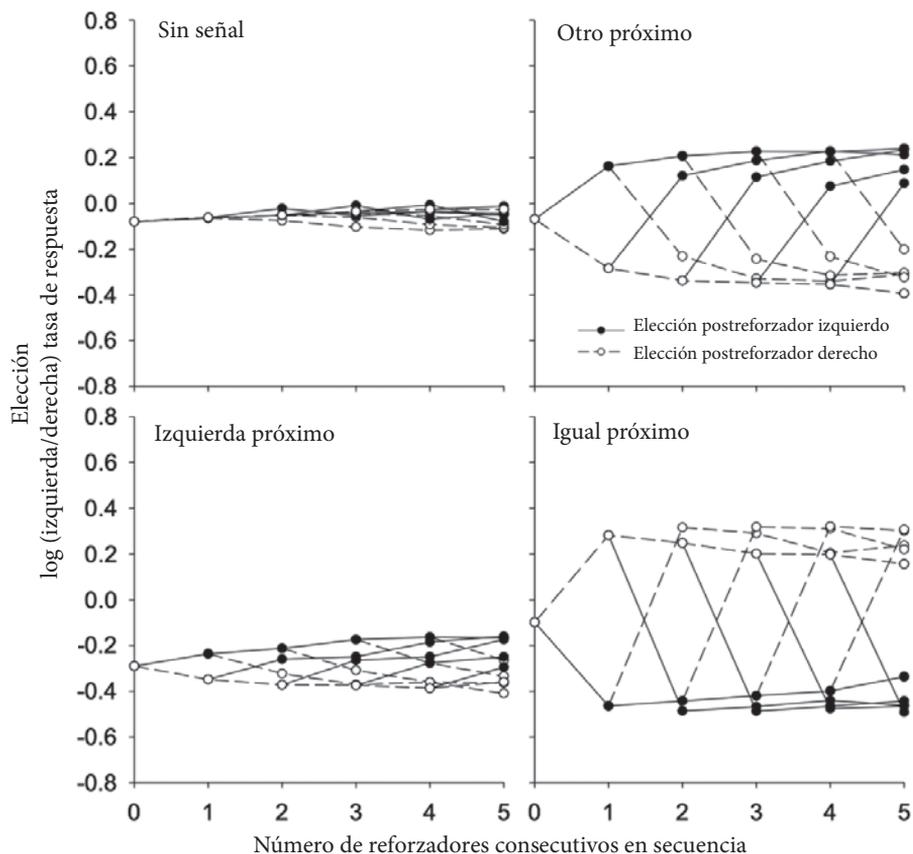
*Funciones de elección local a lo largo de secuencias de reforzadores a la misma respuesta*



Un análisis particularmente apropiado para identificar la presencia y el alcance del efecto de respuesta-recién-reforzada es un *análisis de árbol* (Davison & Baum, 2000), que muestra una sola medida de elección de todas las respuestas emitidas durante el tiempo entre reforzadores, como una función del número de reforzadores en una secuencia. La Figura 4 muestra una selección de datos de Cowie *et al.*, (2011), en la que se calcula en cada punto secuencias que comprenden en su totalidad reforzadores a la misma respuesta, así como secuencias que comprenden algunos reforzadores de la misma respuesta más una *descontinuación* de esta secuencia con un reforzador obtenido en la otra respuesta. El primer dato de cada gráfica muestra el logaritmo de la razón de respuesta obtenida después de cualquier secuencia de cinco reforzadores, independientemente de la ubicación de la respuesta. A partir de entonces, los puntos de datos más externos (las “ramas” más externas del árbol) muestran

la elección después de una secuencia de cinco reforzadores que terminan en una, dos, tres o cuatro entregas de alimento a la misma tecla –*continuación* de entregas de alimentos–. Los puntos de datos internos de la estructura de árbol muestran discontinuaciones seleccionadas de las secuencias de continuación –secuencias de hasta cuatro reforzadores de la misma tecla seguidas de una secuencia de reforzadores de la otra tecla–.

**Figura 4**  
*Análisis de árbol*



La Figura 4 muestra varios aspectos que son característicos de los análisis de árbol en general: independientemente de la contingencia en efecto, el desplazamiento en la elección hacia la respuesta-recién-reforzada es simétrico entre las secuencias de reforzadores de tecla-izquierda y derecha, y disminuye marginalmente con la longitud de la secuencia previa de la misma-respuesta. Por otro lado, el desplazamiento total en la elección generalmente incrementa con el aumento de la longitud de la secuencia (lo que refleja un efecto acumulativo entre los reforzadores), la extensión del desplazamiento disminuye marginalmente con cada reforzador de la misma respuesta en una secuencia, y la discontinuación de los reforzadores va seguida de cambios mucho mayores en la elección que los de continuación (ver también Davison & Baum, 2000; 2022; Landon *et al.*, 2002).

El conjunto de datos de Cowie *et al.*, (Figura 4) muestra que estos efectos a nivel de árbol dependen en cierta medida de las contingencias. Primero, el desplazamiento causado por la discontinuación de los reforzadores en el experimento de Cowie *et al.*, fue mucho más evidente en condiciones en las que una discontinuación del reforzador señaló un patrón diferente de disponibilidad de reforzador que uno de continuación (condiciones Misma-Pronto y Otra-Pronto) que en aquellos en los que todos los reforzadores estaban asociados a la misma disponibilidad subsecuente (condiciones Sin-Señal e Izquierda-Pronto). La diferencia en la magnitud del efecto surge aquí no porque el efecto de ubicación del último reforzador sea mayor en las condiciones Misma-Pronto y Otra-Pronto (véanse las Figuras 2 y 3), sino porque una discontinuación del reforzador en estas condiciones difiere de una continuación en términos de qué señala respecto a los próximos reforzadores, y en términos de su ubicación, mientras que una discontinuación del reforzador en las condiciones Sin-Señal e Izquierda-Pronto difiere solo en términos de su ubicación, y no en términos de lo que señala sobre los próximos reforzadores. Al menos en ambientes que cambian rápidamente, las ratas muestran cambios similares en la elección a través de secuencias de continuación y discontinuación de reforzadores (Aparicio & Baum, 2009) a los que se muestran en la Figura 4. En segundo lugar, en la condición Sin-Señal de Cowie *et al.*, (Figura 3), en la que los reforzadores eran igualmente probables para cualquier respuesta en cualquier momento, los reforzadores no desplazaron consistentemente la elección hacia la respuesta-recién-reforzada. Este hallazgo no es consistente con los patrones de desplazamiento en las otras condiciones de Cowie *et al.*, y tam-

poco es consistente con los hallazgos en programas IV-IV concurrentes más estándar (e.g., Davison & Baum, 2000; 2002; Landon & Davison, 2002). Por lo tanto, al menos en el nivel local de análisis, el efecto de respuesta-recién-reforzada depende en gran medida del contexto.

### ¿Qué causa el efecto de respuesta-recién-reforzada?

Las Figuras 2 a 4 demuestran que un efecto de respuesta-recién-reforzada ocurre incluso cuando el control con el que un reforzador señala la disponibilidad de un siguiente reforzador es fuerte (condiciones Izquierda-Pronto, Misma-Pronto, Otra-Pronto). Si bien un cambio en la elección hacia la respuesta-recién-reforzada es generalmente consistente con el fortalecimiento de la respuesta (es decir, una mayor tendencia a repetir la respuesta más recientemente reforzada), tal explicación resulta embarazosa a la luz del análisis más detallado del efecto anterior. Específicamente, las Figuras 2 a 4 demuestran que el mismo reforzador, entregado con la misma probabilidad y frecuencia global, produce una magnitud diferente del efecto de respuesta-recién-reforzada en diversas situaciones. Primero, la magnitud del efecto varía con la contingencia. La comparación de la magnitud del cambio en la elección a través de reforzadores consecutivos en las condiciones Sin Señal e Izquierda-Pronto (Figura 4) sugiere que el efecto del último reforzador es mayor cuando los reforzadores están asociados con algún cambio a nivel local en la disponibilidad del próximo reforzador (condición Izquierda-Pronto) que cuando no lo están (condición Sin Señal). En segundo lugar, el efecto de respuesta-recién-reforzada es mayor para los reforzadores que descontinúan una secuencia de reforzadores de la misma respuesta y que siguen secuencias más cortas que largas de continuación de reforzadores (Figura 4). Si estos efectos fueran el resultado del fortalecimiento, indicarían que el mismo reforzador fortalecería más –es decir, es más efectivo para controlar la conducta– cuando ese reforzador tiene fuertes propiedades de señalización y cuando otra respuesta ha sido reforzada más recientemente (con el mismo reforzador).

¿Pueden entenderse estos efectos en términos de control por parte de la estructura del medio ambiente? En los datos explorados aquí, los efectos de respuesta recién reforzada fueron evidentes cuando los reforzadores se asociaron con exactamente el mismo cambio en la disponibilidad de reforzadores subsecuentes y, por lo tanto, la estructura del ambiente permaneció sin

cambios. Pero, por supuesto, los ambientes tienen una estructura a lo largo de múltiples escalas de tiempo. En el nivel de estructura más local, la disponibilidad del reforzador puede cambiar de manera predecible a lo largo del tiempo –por ejemplo, segundos desde un evento marcador–; a un nivel menos local, la disponibilidad del reforzador también puede diferir según el número de eventos o el orden en una secuencia de eventos. Las secuencias de entrega de comida a la misma tecla son instancias de parcelas ricas en la sesión, independientemente de la contingencia a nivel local. Pero todas las cosas buenas –incluidas las secuencias de continuación– eventualmente llegan a su fin. Cuando la probabilidad de recibir un reforzador de la misma respuesta es mayor, las secuencias de continuación son en promedio más largas (ver Krägeloh *et al.*, 2005), y esta experiencia puede causar secuencias más largas de entregas de comida a la misma tecla para indicar una *menor probabilidad* de obtener la siguiente entrega de comida en la tecla productiva más reciente, similar a cómo la falacia del jugador (Laplace, 1796; Jarvik, 1951) precipita una apuesta a cruz después de una racha de caras, una reducción en la cantidad apostada en un número de lotería recientemente ganador (Clotfelter & Cook, 1993), o la venta prematura de inversiones a las que les ha ido (y les sigue yendo) bien (Odean, 1998, Weber & Camerer, 1998, Shapira & Venezia, 2001, Rabin, 2002, Chen *et al.*, 2007). El efecto de respuesta-recién-reforzada como control discriminativo (aunque erróneo) por secuencias de reforzadores sería consistente con la observación de que el efecto generalmente disminuye a medida que aumenta la extensión de las continuaciones, y en tanto las secuencias son en general más cortas y menos probables (Krägeloh *et al.*, 2005).

Ciertamente, en los datos revisados aquí (Figura 4), el efecto de la última-respuesta-reforzada fue más pequeño y menos consistente en la condición Sin Señal (Figura 4). Podría decirse que, en la condición Sin-Señal de Cowie *et al.*, los efectos de fortalecimiento del reforzador deberían ser *más* prominentes porque no hay efectos de señalización de reforzadores que compitan por el control. Por lo tanto, si el desplazamiento hacia la respuesta recién reforzada refleja un efecto de fortalecimiento de la respuesta, debería ser *mayor* en esta condición. En cambio, la variación en el efecto de respuesta recién reforzada en las condiciones de Cowie *et al.*, sugiere que este efecto refleja el control de eventos subsecuentes. De hecho, Krägeloh *et al.*, (2005) notaron que el efecto de respuesta recién reforzada era mayor cuando las secuencias de continuación ocurren con más frecuencia y, por lo tanto, más probable que sean más

largas. Aunque Boutros *et al.*, (2011) observaron que se producía un efecto de respuesta-recién-reforzada cuando los reforzadores rigurosamente se alternaban, su procedimiento empleó un requisito de cambio, el cual provoca un breve período inmediatamente después de cada reforzador durante el cual no se pueden obtener reforzadores para la respuesta que no es la recién reforzada, independientemente de la contingencia establecida por el experimentador (ver Cowie & Davison, 2016; Davison *et al.*, 2013; Gomes-Ng *et al.*, 2018 para una discusión).

Por supuesto, las discontinuaciones también pueden constituir eventos sorprendidos, al menos en el contexto de que no hayan ocurrido en un pasado muy reciente. Eventos con efectos desproporcionadamente grandes que podrían considerarse «sorprendidos» en el contexto del patrón de eventos recientes son evidentes no solo en las mediciones de comportamiento y memoria (por ejemplo, Cowie *et al.*, 2021; Maki, 1979; Rescorla & Holland, 1982; William & Wagner, 1975), sino también en la actividad cerebral (por ejemplo, Cavanagh *et al.*, 2010; Roesch *et al.*, 2012), a través de una variedad de paradigmas experimentales. Los eventos que no son consistentes con la estructura percibida de un ambiente pueden sugerir que la estructura está cambiando o que se ha percibido erróneamente. De hecho, sería una mala adaptación comportarse como si los cambios de contingencia nunca fueran a ocurrir y los eventos que están fuera de lugar en el contexto de los eventos recientes podrían ser señales (potencialmente poco confiables) de un cambio inminente. Sin embargo, mientras que un mecanismo exploratorio, preparado para el cambio, podría explicar una pequeña parte del efecto de ubicación-del-último-reforzador, no puede explicar el conjunto completo de sus características.

Aunque la conducta operante es en gran medida el producto de la ontogenia –la propia experiencia del organismo– es imposible escapar a la influencia de la filogenia (ver, por ejemplo, Baum, 2012). De hecho, la interacción entre filogenia y ontogenia está bien documentada tanto en el condicionamiento respondiente como operante. Lo que Breland y Breland (1961) denominaron «mala conducta» es la intrusión de la filogenia en la experiencia, que se produce cuando la estructura ambiental operante induce conductas incompatibles (véase también Boakes *et al.*, 1978), como podrían ocurrir con muchos reforzadores alimenticios cuando esos alimentos en ocasiones son almacenados o requieren procesamiento antes de ser consumidos. Se podría esperar que el trigo induzca el picoteo en las palomas, ya que desde un punto de vista

filogenético existe una covarianza entre el picoteo y el trigo. Ciertamente, la ocurrencia regular y predecible de trigo en ausencia de cualquier requerimiento de picotear suele ser suficiente para inducir el picoteo en las palomas (Staddon & Simmelhag, 1971; Palya & Zacny, 1980), quizás de la misma manera que otros comportamientos son inducidos por reforzadores en otras especies (Faulk, 1966). Por lo tanto –como Jenkins y Moore (1973) señalaron hace cincuenta años–, incluso cuando el experimentador se enfoque en la ontogenia, el picoteo en la cámara operante estará impulsado tanto por la historia filogenética como ontogenética de una paloma. En consecuencia, el efecto de ubicación del último reforzador puede ser (al menos en parte) un simple caso de «mala conducta» –es decir, una intrusión filogenética–.

Una segunda influencia filogenética en el efecto de respuesta-recién-reforzada podría provenir del tipo de reforzador utilizado. Está bien documentado que la topografía de la conducta operante depende de las características específicas del reforzador que mantiene la conducta. Las palomas picotean con el pico más abierto cuando los picoteos se mantienen con chícharos en comparación con granos más pequeños como el milo, y con sus picos casi cerrados cuando los picoteos se mantienen con agua (LaMon & Zeigler, 1984), y los picoteos son más enérgicos y de menor duración cuando preceden a los reforzadores alimentarios que a los de agua (Spetch *et al.*, 1981; véase también Wolin, 1968). Las ratas presionan la palanca por más tiempo cuando responden por comida en lugar de agua con sacarosa y este efecto no disminuye con la exposición a la contingencia (Hull, 1977). Spetch (1981) mostró que el impacto del reforzador en la topografía de la respuesta era anticipatorio, dependiendo críticamente del próximo reforzador más que del anterior. Se han observado efectos de tipo reforzador similares en el automoldeamiento (e.g., Jenkins & Moore, 1973) y, por supuesto, en el condicionamiento respondiente, donde la naturaleza del estímulo incondicionado –también un evento filogenéticamente importante– determina (en parte) la forma de la respuesta condicionada. De hecho, bajo un marco basado en la estructura, el condicionamiento operante y respondiente son el resultado del control combinado de las relaciones ontogenéticas y filogenéticas; son distintos por la forma en que se ordenan las correlaciones ontogenéticas, no porque sean procesos de aprendizaje diferentes.

El impacto del reforzador en el comportamiento no se limita a su forma; cómo ese reforzador se distribuye típicamente en el mundo natural también

afecta el control por relaciones ontogenéticas. Los recursos dependen unos de otros hasta cierto punto (Taylor, 1961; Taylor, Woiwod & Perry, 1978) y rara vez se distribuyen al azar. Cuando los reforzadores se agrupan en parcelas –como suele ser el caso de los granos de trigo en el mundo natural, la aparición de uno predice más de lo mismo para el mismo comportamiento. Por lo tanto, la tendencia a repetir la respuesta-recién-reforzada hasta cierto punto -conducta de permanecer-ganar- puede reflejar el control a partir de una relación entre reforzadores construidos a lo largo de una escala de tiempo mucho más larga: la historia evolutiva del organismo. Esta hipótesis es consistente con una función de señalización de los reforzadores (Boutros, Davison y Elliffe, 2011), simplemente representa el control por parte de las señales que se derivan de la historia filogenética de un organismo.

Ciertamente, los animales que normalmente buscan recursos que están agrupados en parcelas son generalmente más lentos para aprender las contingencias de cambiar-ganar que las contingencias de permanecer-ganar (e.g., Randall & Zentall, 1997). Lo contrario es cierto para los animales que no buscan alimento agrupados en parcelas (e.g., Cole, Hainsworth, Kamil, Mercier & Wolf, 1982). Estos efectos son evidentes incluso en poblaciones criadas en cautiverio sin experiencia (Burke & Fulham, 2003; pero ver Sulikowski & Burke, 2011); es decir, no se aprenden a través de la experiencia del organismo, sino que surgen por la historia filogenética de la especie. Tanto la especie como el propio reforzador afectan la probabilidad de que un organismo adopte un comportamiento de permanecer-ganar; los reforzadores que normalmente se agrupan en parches promoverán una mayor conducta de permanecer-ganar que los reforzadores no distribuidos de esta manera (e.g., Sulikowski y Burke, 2010). Incluso los humanos son propensos al comportamiento de permanecer-ganar, particularmente cuando la tarea involucra recursos que ocurren naturalmente y pueden ser buscados en un sentido más primitivo (Wilke & Barrett, 2009). Estos sesgos filogenéticos para repetir o cesar una conducta después de un reforzador aún pueden considerarse el resultado de lo que un reforzador predice sobre los reforzadores futuros, pero esta predicción se hace sobre la base de la evolución del organismo, en lugar de su historia de aprendizaje (Boutros *et al.*, 2000).

La fricción entre la filogenia y la ontogenia significa que los reforzadores pueden convertirse en señales de segundo orden o estímulos discriminativos (o lo que Baum, 2012, podría denominar inductores condicionales). A menu-

do, los estímulos exteroceptivos ejercen un control conductual más fuerte que los estímulos reforzadores que señalan la contingencia equivalente (Cowie *et al.*, 2011; Davison & Smith, 1986; Davison *et al.*, 1980; Gomes-Ng *et al.*, 2023). El reforzador no solo es un estímulo breve que ya no está presente y debe recordarse a medida que ocurren eventos posteriores (Cowie *et al.*, 2017); los reforzadores también llevan un bagaje filogenético que ocasiona conductas que no conducen necesariamente al aprendizaje de relaciones ontogenéticas (e.g., véase Faulk, 1970). Las asociaciones filogenéticas preexistentes entre los reforzadores y los eventos subsiguientes pueden ocasionar un comportamiento que es totalmente incompatible con el control ontogenético y, en estas situaciones, las relaciones filogenéticas suelen triunfar sobre las ontogenéticas (Breland & Breland, 1961; Boakes *et al.*, 1978). Cuando los reforzadores no tienen una historia ontogenética -como podría ocurrir durante el moldeamiento- es su historia filogenética la que determina sus efectos.

## Conclusión

Los efectos de ubicación-del-último-reforzador son menores en comparación con los efectos discriminativos a nivel local (Figura 2), pero siguen siendo una característica del control ambiental que debe entenderse. Los efectos de ubicación-del-último-reforzador revisados aquí sugieren que el efecto se entiende mejor como un producto del control por parte de la estructura del ambiente, en lugar de un fortalecimiento de la respuesta. El patrón de los efectos de ubicación-del-último-reforzador en diferentes condiciones y diferentes secuencias sugiere fuertemente un control dinámico tanto por la ontogenia (en varias escalas de tiempo) como por la filogenia. Este entrelazamiento de la experiencia específica de la especie y del organismo permite que el organismo se comporte de manera adaptativa, de una manera apropiada para eventos futuros -el objetivo final, como señala Baum (2012), es la supervivencia del animal y su especie. Un siguiente paso fundamental para comprender el control del reforzador es explorar cómo estos diferentes niveles de experiencia se combinan, compiten e interactúan para determinar el control del reforzador, examinando una variedad de especies, comportamientos, reforzadores y estímulos.

## Referencias

- Bain, A. (1868). *The senses and the intellect* (3rd ed.). Longmans, Green and Co. <https://doi.org/10.1037/12273-000>
- Baum, W. M. (1974). Choice in a changing environment: Changes in reinforcing efficacy. *Journal of the Experimental Analysis of Behavior*, 22(2), 189–197.
- Baum, W. M. (2012). Rethinking reinforcement: Allocation, induction, and contingency. *Journal of the Experimental Analysis of Behavior*, 97(1), 101–124. <https://doi.org/10.1901/jeab.2012.97-101>
- Boakes, R. A., Poli, M., Lockwood, M. J., & Goodall, G. (1978). A study of misbehavior: Token reinforcement in the rat. *Journal of the Experimental Analysis of Behavior*, 29(1), 115–134. <https://doi.org/10.1901/jeab.1978.29-115>
- Bolles, R. C. (1975). *Theory of motivation*. 2nd ed. New York (N.Y.): Harper and Row.
- Boutros, N., Davison, M., & Elliffe, D. (2009). Behavioural sources of and functions of “last reinforcer location” effects. *Journal of the Experimental Analysis of Behavior*, 91(2), 155–166.
- Breland, K., & Breland, M. (1961). The misbehavior of organisms. *American Psychologist*, 16(11), 681–684. <https://doi.org/10.1037/h0040090>
- Burke, D., & Fulham, B. J. (2003). An evolved spatial memory bias in a nectar-feeding bird? *Animal Behaviour*, 66(4), 695–701. <https://doi.org/10.1006/anbe.2003.2246>
- Cavanagh, J. F., Frank, M. J., Klein, T. J., & Allen, J. J. (2010). Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *NeuroImage*, 49(4), 3198–3209. <https://doi.org/10.1016/j.neuroimage.2009.11.080>
- Chen, G., Kim, K. A., Nofsinger, J. R., & Rui, O. M. (2007). Trading performance, disposition effect, overconfidence, representativeness bias, and experience of emerging market investors. *Journal of Behavioral Decision Making*, 20(4), 425–451. <https://doi.org/10.1002/bdm.561>
- Clotfelter, C. T., & Cook, P. J. (1993). The “gambler’s fallacy” in lottery play. *Management Science*, 39(12), 1521–1525.
- Cowie, S., & Davison, M. (2016). Behavioural indicators of control by structural properties. *Behavioural Processes*, 122, 2–13.

- Cowie, S., Davison, M., Elliffe, D., & Kulesza, A. (2013). Temporal control by structural relations in behavioral systems. *Journal of the Experimental Analysis of Behavior*, 99(2), 124–145.
- Davison, M., & Baum, W. M. (2002). Choice in a variable environment: Effects of blackout duration and extinction between components. *Journal of the Experimental Analysis of Behavior*, 77, 65–89. doi: 10.1901/jeab.2002.77-65
- Davison, M., & Baum, W. M. (2006). The interaction of phylogeny and ontogeny in behaviour. *Animal Behaviour*, 72(3), 581–589.
- Davison, M., & Baum, W. M. (2010). Stimulus effects on local preference: Stimulus–Response contingencies, stimulus–food pairing, and stimulus–food correlation. *Journal of the Experimental Analysis of Behavior*, 93, 45–59. <https://doi.org/10.1901/jeab.2010.93-45>.
- Davison, M., & Baum, W. M. (2020). The importance of reinforcer distribution in the evolution of behavior. *Journal of the Experimental Analysis of Behavior*, 113(3), 427–441.
- Gilovich, Thomas (1993). *How We Know What Isn't So*. Free Press.
- Herrnstein R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the experimental analysis of behavior*, 4(3), 267–272. <https://doi.org/10.1901/jeab.1961.4-267>
- Jarvik, M. E. (1951). Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *Journal of Experimental Psychology*, 41(4), 291–297. <https://doi.org/10.1037/h0056878>
- Laplace, P. S. (2012). *A philosophical essay on probabilities*. Courier Corporation.
- Maák, I., Lórincki, G., Le Quinquis, P., Módra, G., Bovet, D., Call, J., & d’Ettorre, P. (2017). Tool selection during foraging in two species of funnel ants. *Animal Behaviour*, 123, 207–216. <https://doi.org/10.1016/j.anbehav.2016.11.005>
- Maki, W. S. (1979). Pigeons’ short-term memories for surprising vs. expected reinforcement and nonreinforcement. *Animal Learning & Behavior*, 7(1), 31–37. <https://doi.org/10.3758/BF03209653>
- Odean, T. (1998). Are investors reluctant to realize their losses?. *The Journal of finance*, 53(5), 1775–1798.
- Palya, W. L., & Zacny, J. P. (1980). Stereotyped adjunctive pecking by caged pigeons. *Animal Learning & Behavior*, 8(2), 293–303.
- Rabin, M. (2002). Inference by believers in the law of small numbers. *The Quarterly Journal of Economics*, 117(3), 775–816.

- Rescorla, R. A., & Holland, P. C. (1982). Behavioral studies of associative learning in animals. *Annual review of psychology*, 33(1), 265-308.
- Roesch, M. R., Esber, G. R., Li, J., Daw, N. D., & Schoenbaum, G. (2012). Surprise! Neural correlates of Pearce–Hall and Rescorla–Wagner coexist within the brain. *European Journal of Neuroscience*, 35(7), 1190-1200.
- Shahan T. A. (2010). Conditioned reinforcement and response strength. *Journal of the experimental analysis of behavior*, 93(2), 269–289. <https://doi.org/10.1901/jeab.2010.93-269>
- Shapira, Z., & Venezia, I. (2001). Patterns of behavior of professionally managed and independent investors. *Journal of Banking & Finance*, 25(8), 1573-1587.
- Taylor, L. R. (1961). Aggregation, variance and the mean. *Nature*, 189, 732–735. [10.1038/189732a0](https://doi.org/10.1038/189732a0)
- Taylor, L. R., Woiwod, I. P., & Perry, J. N. (1978). The density- dependence of spatial behaviour and the rarity of randomness. *Journal of Animal Ecology*, 47, 383–406. [doi:10.2307/3790](https://doi.org/10.2307/3790)
- Terry, W. S., & Wagner, A. R. (1975). Short-term memory for “surprising” versus “expected” unconditioned stimuli in Pavlovian conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, 1(2), 122–133. <https://doi.org/10.1037/0097-7403.1.2.122>
- Tonneau, F. (2001). Equivalence relations: A critical analysis. *European Journal of Behaviour Analysis*, 2, 1-128. <https://doi.org/10.1080/15021149.2001.1434165>
- Weber, M., & Camerer, C. F. (1998). The disposition effect in securities trading: An experimental analysis. *Journal of Economic Behavior & Organization*, 33(2), 167-184.
- Wilke, A., & Barrett, H. C. (2009). The hot hand phenomenon as a cognitive adaptation to clumped resources. *Evolution and Human Behavior*, 30, 161–169. [doi:10.1016/j.evolhumbehav.2008.11.004](https://doi.org/10.1016/j.evolhumbehav.2008.11.004)
- Wolin, B. R. (1968). Difference in manner of pecking a key between pigeons reinforced with food and with water. *Contemporary research in operant behavior*, 286.