

BEING THERE ON TIME: REINFORCER EFFECTS ON TIMING AND LOCATING

SARAH COWIE

MICHAEL DAVISON

2020

La correlación entre eventos ambientales ejerce control sobre la conducta suponiendo que las condiciones futuras señaladas por un *evento marcador* sean relevantes en términos de los *affordances*, disposiciones, e historia filogenética del organismo. Eso se conoce como *control de estímulo*.

Las correlaciones entre eventos ocurren en una escala temporal dada, y su efectividad para ejercer control de estímulos depende de la habilidad de los organismos para recordar los eventos en la escala temporal relevante: no se puede aprender una correlación si un evento se olvida antes de que el siguiente tenga lugar.

El recuerdo impreciso—generalización—en tiempo o espacio distorsionará la relación discriminada entre eventos y conducta.

El control por correlaciones depende también de los reforzadores que forman parte de ellas, por ejemplo, cuando la disponibilidad de reforzadores cambia entre localizaciones en un tiempo dado, las diferencias en el porcentaje de reforzadores obtenidos antes y después del tiempo causan diferencias confiables en el patrón de elección a través del tiempo a pesar de que la correlación entre localización y reforzador permanezca igual.

Un mecanismo posible por el que los reforzadores afectan la discriminación de cambios basados en el tiempo es un efecto en el *arousal*: el *arousal* alteraría la velocidad del “marcapasos” que mueve a los animales a través de estados, y eso cambia la velocidad del tiempo subjetivo. Diferencias en *arousal* determinarían cuán precisamente la elección sigue a cambios en la localización de reforzadores a través del tiempo. Las aproximaciones de *timing* asumen que los reforzadores afectan la conducta bajo el control de correlaciones basadas en el tiempo porque afectan al *timing*, y porque afectan la fuerza de las cadenas asociativas.

Otra aproximación supone que los reforzadores no tienen ninguna función especial que les permita influir en la discriminación, sino que son una parte más del entorno. La elección seguiría así a la disponibilidad *percibida* de reforzadores, que puede diferir de la real debido a la generalización en tiempo y espacio. Esta aproximación asume que los reforzadores más frecuentes se generalizan en mayor medida que los menos frecuentes, y por lo tanto tendrán un mayor impacto en la relación *discriminada* entre reforzadores y otros aspectos del entorno (estímulos, respuestas).

La aproximación de *Learning to Time* de Machado es similar, pero se distingue en que (1) aquella sí supone que los reforzadores tienen una función especial para alterar la discriminación del tiempo (mediante su impacto en el *arousal*); y (2) una aproximación de *timing* supone que el control imperfecto de las contingencias viene solo del error en discriminación del tiempo, mientras que la aproximación de la discriminación supone que las fuentes de error son errores en tiempo y espacio.

Este estudio pretende determinar cómo la distribución de reforzadores antes y después de un cambio en una contingencia temporal afecta al control ejercido por ese cambio. Se utilizó un programa en el que una alternativa tenía nueve veces más probabilidad de producir reforzamiento que otra, y su posición se revirtió en puntos determinados por el tiempo pasado desde el último evento marcador. Se midió la velocidad del cambio en preferencia, el tiempo en que la preferencia se revirtió, y el grado del cambio en la preferencia a través de condiciones. Se buscó explicar los datos con dos versiones del modelo de discriminación. Se examinó cómo la frecuencia de reforzamiento contribuye a discriminar las relaciones entre espacio, tiempo, y disponibilidad de reforzamiento.

1. Method

Se evaluaron cuatro palomas en cajas con tres teclas y un comedero.

1.1. Procedure

Se entregaban 60 reforzadores en sesiones de un máximo de 60 minutos.

En cada ensayo dos teclas estaban disponibles, y las respuestas a una tenían siempre nueve veces más probabilidad de producir reforzamiento. La tecla con la mayor probabilidad se revertía tras pasar 19 segundos desde el inicio del ensayo. Las teclas fueron etiquetadas como *higher-to-lower* y *lower-to-higher*. La posición de las teclas era alterna-da entre condiciones.

Terminar el ensayo con la entrega del reforzador aseguraba que el único estímulo discriminativo para la localización de los reforzadores era el tiempo pasado desde la última entrega.

Se determinó que todo reforzador programado en una tecla pero no obtenido antes de la reversión permanecía disponible pero era reasignado a la tecla con la misma probabilidad después de la reversión. Esto impedía que la elección desplazara los reforzadores a través del tiempo y potencialmente creara un pico local en la tasa de reforzadores después de la reversión.

Los reforzadores se entregaron con base en un programa de intervalo variable para asegurar que el número de reforzadores programados permanecía constante a través del tiempo después de los reforzadores.

Al inicio de cada sesión y tras cualquier reforzador, el programa determinaba al azar si el siguiente reforzador ocurriría antes o después de la reversión de acuerdo con una probabilidad que variaba a través de condiciones. Después, los tiempos de reforzamiento

se escogían al azar sin reemplazo de dos listas de tiempos de 1 segundo para reforzadores antes o después de la reversión.

La lista pre-reversión tenía tiempos de 1 a 19 segundos en pasos de 1s; la post-reversión tenía tiempos de 10 a 60 en pasos de 1s. Tras elegir el reforzador, éste se asignaba a la tecla de mayor probabilidad con 90 % de probabilidad. Así, siempre había mayor probabilidad de reforzamiento en esa tecla, pero la localización (izquierda o derecha) dependía de si el reforzador estaba programado antes o después de la reversión de probabilidades.

La probabilidad de reforzamiento permaneció en .9 y .1 todo el experimento, pero se manipuló la proporción de reforzadores obtenidos antes de la reversión de 5 % a 95 % en una tecla con relación a la otra.

Se calculó la proporción de respuestas a las teclas y los reforzadores obtenidos por ellas en cada bin de 1s. Esto describe el cambio en la elección en función del tiempo desde un evento marcador.

Se ajustó a los datos una función Gaussiana dado que se ha visto que tienen buen ajuste con cambios basados en el tiempo en la disponibilidad de reforzadores. La función está descrita por la ecuación

$$\frac{B_{L \rightarrow H}}{B_{L \rightarrow H} + B_{H \rightarrow L}} = Y_0 + \frac{a}{1 + e^{-\left(\frac{t - X_0}{\beta}\right)}}$$

donde B son las respuestas, Y_0 es el mínimo de la función y $Y_0 + a$ es el máximo; X_0 es el punto en que la ojiva está en el punto medio del mínimo y máximo, y β es el inverso de la pendiente en el punto X_0 y determina la velocidad del cambio en la ojiva a través de unidades de t . La ecuación provee medidas separables de características de la función de elección mediadas por la localización y por el tiempo. Ajustarla a los datos permite cuantificar el impacto de la tasa relativa de reforzamiento antes y después de la reversión del programa en el sesgo, la precisión y exactitud de la discriminación del momento del cambio, y en el grado del control ejercido por el cambio en la localización de los reforzadores probables antes y después de la reversión.

Los ajustes de los datos con la ecuación y la obtención de Factores de Bayes mostraron que el porcentaje de reforzadores obtenidos antes de la reversión tuvieron un impacto claro en la forma en que la elección seguía a la *localización* probable de los reforzadores en cualquier lado de la reversión, pero tenían prácticamente ningún impacto en el grado en que la elección era controlada por el tiempo del cambio en la localización probable.

Los cambios en la elección entre localizaciones fueron más extremos cuando los porcentajes de reforzadores obtenidos antes y después de la reversión eran más similares. Se hipotetiza que la razón está en la generalización de los reforzadores entregados a través del tiempo y localización. Se ajustó un modelo de discriminación que supone que la elección iguala al diferencial de reforzadores discriminado, que se obtiene desplazando los reforzadores obtenidos a través de ambas dimensiones. Un modelo inicial en el cual se asumía que la verosimilitud de la generalización entre localizaciones y tiempos no es afectada por el tiempo en que se obtiene un reforzador (y por lo tanto asume un cambio lineal en estos parámetros) probó un ajuste incorrecto con los datos—se encontraron desviaciones sistemáticas entre los datos y las predicciones. Un modelo no-lineal en el cual los

parámetros incrementaron de forma ojival fue ajustado en su lugar y mostró un mejor ajuste—desviaciones pequeñas y no-sistemáticas.

1.2. Discussion

Cambiar la frecuencia relativa de reforzadores antes y después de la reversión alteró el grado en que la elección era controlada por la probable *localización* de los reforzadores, pero no por los aspectos *temporales* de la contingencia. Los cambios en la elección fueron mayores cuando las proporciones de reforzamiento eran más similares.

Hubo un mayor impacto sobre la elección de los reforzadores antes de la reversión, lo que indica que el control por las correlaciones basadas en el tiempo es más influenciado por el número de reforzadores en tiempos posteriores que en tiempos anteriores.

Este procedimiento se distingue de otro similar, el *Free Operant Psychophysical Procedure*, en que aquél entrega más de un reforzador por ensayo, lo que podría funcionar como un estímulo adicional para o mellar la discriminación temporal. El procedimiento actual entrega un único reforzador por ensayo, por lo que lo único que permite discriminar el tiempo de cada reforzador es discriminable solo con base en el paso del tiempo desde el evento marcador.

La discriminación a través de espacio y de tiempo son necesarias en conjunto. Cuando cualquiera es imperfecta habrá generalización a través del tiempo y localización de los reforzadores. Un parámetro que pretenda explicar la traducción de inputs del entorno a conducta debe reflejar ambos tipos de generalización. Interpretar un parámetro que caracterice la discriminación como producto de una sola fuente de generalización probablemente dará un estimado de generalización inflado o enteramente erróneo.

Los ajustes del modelo de discriminación con los datos indican que las diferencias en la proporción de reforzadores obtenidas antes de la reversión tuvieron virtualmente ningún efecto sobre la precisión de la generalización temporal. El modelo sugiere que los cambios en la elección se debieron a la generalización de *localización*. Así, cuando la duración que debe estimarse no cambia, la discriminación *temporal* no cambia. Esto solo sorprende si uno se basa en los modelos de generalización basados en *timing* debido a que en ellos no hay cabida para ningún error más que los errores de *timing*.

Estos resultados sugieren que los reforzadores no alteran la discriminación del tiempo transcurrido, sino que forman parte del tejido del entorno. Es la forma en que ese tejido es discriminado—la estructura temporal aparente—lo que guía a los organismos al sitio correcto en el momento correcto.

Esto añade al cuerpo de evidencia que indica que los reforzadores no son especiales por su función para controlar la conducta, sino con base en su valor para el organismo. La conducta depende de las correlaciones discriminadas entre eventos del entorno.