

RATS MAINTAIN OPTIMAL CHOICE WHEN FACING LONG TERMINAL LINKS IN A “SUBOPTIMAL CHOICE” TASK

RODRIGO ALBA RODRIGO GONZÁLEZ-TORRES
VLADIMIR ORDUÑA

2021

Aunque la diferencia entre ratas y palomas en elección subóptima podría implicar una diferencia en los procesos de decisión involucrados en la tarea, es más probable que haya diferencias procedimentales. Es importante encontrar generalidad dado que se ha propuesto a la ejecución de las palomas como un modelo animal de *gambling*.

La primera explicación para las diferencias se basó en la saliencia incentiva de los estímulos, pero se encontraron resultados conflictivos.

Dos modelos de los muchos que hay han abordado directamente la ejecución diferencial entre especies: el de *associability decay*, que asume que el valor de la alternativa se actualiza de acuerdo con un parámetro de asociabilidad relacionado inversamente con la certidumbre con la cual los estímulos predicen sus consecuencias. La asociabilidad de la alternativa discriminativa rápidamente llega a cero y su valor deja de actualizarse, pero la asociabilidad de la alternativa no-discriminativa prevalece. La dinámica de la pérdida de la asociabilidad y el valor más alto del estímulo pareado con certeza de reforzamiento resultan en preferencia por la alternativa discriminativa. El modelo se ajusta a ambas especies considerando que en las ratas no decae la asociabilidad. Esta distinción podría relacionarse con la sensibilidad de cada especie a los estímulos asociados con no-reforzamiento: las palomas dejan de ser influidas por él, pero las ratas no.

El *temporal information-theoretic model* asume que el valor de las alternativas es determinado por la información temporal que provee sobre la entrega de comida y por la tasa de reforzamiento primario. La diferencia entre especies se explica con una distinta ponderación para cada componente: las palomas dan más peso a la información temporal; las ratas, al reforzamiento primario. El peso es modulado por la sensibilidad a la demora media a la comida y a los estímulos discriminativos en el momento de la elección y por un parámetro de sesgo en contra de la utilización de la información temporal. Incrementos en la longitud de los TL deberían incrementar el peso dado a la información temporal e incrementar la elección subóptima. Es posible que las ratas necesiten TL más largos que las palomas para mostrar elección subóptima. Esto fue probado presentando a las

ratas TLs de 10 a 50 segundos. En TL más largos la preferencia de las ratas pasaba a la suboptimalidad.

Bajo suposiciones distintas, los modelos de utilidad anticipatoria de Iigaya y SiGN de McDevitt predicen elección subóptima también según incrementan los TL, predicción que fue confirmada en palomas y en humanos.

Este estudio pretende analizar si estos resultados son generalizables a oro procedimiento de elección subóptima en el que los estímulos tienen saliencia incentiva. Esto puede proveer información sobre la posible interacción entre la saliencia incentiva y la demora a la comida.

EXPERIMENT 1

METHOD

Se usaron nueve ratas en cajas experimentales con dos paneles idénticos. Cada panel tenía dos palancas, comedero y *nosepoke*.

Los sujetos fueron entrenados a responder en *nosepoke*.

Hubo tres grupos separados con base en el orden en que experimentaron las condiciones: 10-30-50, 30-10-50, 50-30-10. La diferencia entre las condiciones era el tiempo en segundos durante el cual las palancas permanecían extendidas. Las sesiones tenían 60 ensayos: 20 forzados a cada alternativa y 20 libres. El procedimiento utilizado fue el de elección subóptima con probabilidades de .2 y .5, con una alternativa en cada panel. Hubo un intervalo entre ensayos de 10 segundos. Tras 40 sesiones en una condición se pasaba a la siguiente a la vez que se realizaba una reversión de las posiciones de las alternativas.

RESULTS

En todas las condiciones se encontró una preferencia consistente por la alternativa óptima. Dado que no se encontró un efecto del orden, los datos de cada longitud de TL se colapsan para el análisis. Se encontró un efecto principal de sesiones solamente, pero no de condición ni de interacción.

No se encontró diferencia en la proporción de elección entre las condiciones, pero todas fueron distintas de la indiferencia.

Las respuestas a las palancas mostraron que los animales discriminaban bien en la alternativa discriminativa y no lo hacían en la no discriminativa.

Se exploró si el nivel de sign-tracking asociado con el reforzamiento se asocia con la elección subóptima al correlacionar la tasa de respuestas en los últimos 10s en presencia de la palanca durante ensayos forzados con la proporción de elección. No se encontró ninguna correlación.

Se evaluó el efecto de los estímulos (S^+ , S^-) y las condiciones en la tasa de respuestas a comedero y se encontró un efecto significativo de estímulo y condición, pero sin interacción. Las tasas de respuesta en presencia de S^+ eran mayores cuando el TL era de 10s que cuando era de 50s. No hubo diferencias en las entradas a comedero en la alternativa no-discriminativa.

DISCUSSION

Las ratas mantienen su preferencia sin importar la longitud del TL a pesar de que se sostiene la discriminación. Esto contradice el estudio de Cunningham y Shahan (2019) en el que se reportó una gran influencia del TL en la preferencia subóptima. Dado que se usó la misma longitud de TL, es razonable suponer que la diferencia está en el procedimiento, por

lo que se realizó un segundo experimento que utilizaba la mayor parte de las características de Cunningham y Shahan, incluyendo el TL 50s, y se contrastó la ejecución de dos grupos que difirieron en el estímulo del TL: para un grupo se usó una combinación de tono y luz (como Cunningham y Shahan); para el otro se empleó solamente luz. Además, en la alternativa discriminativa había dos posibles estímulos: una combinación intermitente de tono y luz como estímulo positivo, y un blackout como negativo. La alternativa no discriminativa tenía solo un estímulo (presentación continua de tono y luz).

EXPERIMENT 2

METHOD

Se utilizaron 16 ratas en cajas operantes con un solo panel. El panel tenía dos palancas, cada una con triple estimulador, un comedero, una luz ambiental y un altavoz.

Los animales fueron entrenados a responder en las palancas y luego divididas en dos grupos: “tono + luz” y “luz”. Los grupos solo diferían en los estímulos de TL.

Cada ensayo comenzaba con el encendido de la luz ambiental y la presentación de una o ambas palancas con la luz blanca asociada con ella. Las probabilidades de reforzamiento eran de .2 y .5. Para el grupo “tono + luz” el estímulo positivo fue la presentación intermitente de una luz y un tono; para el grupo “tono” se presentó intermitentemente solo la luz. El estímulo permaneció por 50s seguido de la entrega de un reforzador y el ITI. El estímulo negativo fue un blackout en ambos casos. En la alternativa no-discriminativa el único estímulo fue la presentación constante del compuesto tono + luz o solo la luz.

RESULTS

El grupo “tono + luz” mostró preferencia por la alternativa subóptima; el grupo “luz”, no. Un ANOVA con las últimas 5 sesiones mostró que el efecto de *grupo* fue significativo, pero no los de *sesión* ni la interacción.

Las respuestas a comedero mostraron buena discriminación en la alternativa discriminativa, y no hubo diferencias entre las respuestas a S^+ y las respuestas al estímulo no-discriminativo.

DISCUSSION

El grupo “tono + luz” mostró conducta subóptima en niveles similares a los reportados por Cunningham y Shahan. El resultado principal es que los TL largos no son suficientes para generar conducta subóptima en rats, y que la duración del TL interactúa con las características de los estímulos. Cunningham y Shahan sugirieron que su estímulo altamente saliente podría adquirir un mayor valor como reforzador condicionado e incrementar el valor de la alternativa. Estos resultados son consistentes con esa interpretación. Quizá también la saliencia del inhibidor condicionado podría relacionarse con su fuerza como inhibidor. Esta inhibición contrarrestaría el reforzamiento condicionado derivado del S^+ , bajando el valor neto de la alternativa discriminativa. Eso explicaría por qué un *blackout* (con poca saliencia) preserva la elección subóptima.

Quizá la saliencia podría relacionarse con la naturaleza intermitente de los estímulos, pero eso no fue evaluado aquí.

Hay evidencia que indica que estímulos altamente salientes incrementan el valor de las alternativas y promueven la elección maladaptativa. Esto sugiere la necesidad de evaluar los valores excitatorio e inhibitorio de los estímulos usados en la tarea de elección subóptima, igual que su interacción con la longitud del TL.

GENERAL DISCUSSION

En el experimento 1 se obtuvo preferencia por la alternativa óptima a pesar del uso de TL largos y estímulos con alta saliencia incentiva.

No se encontró relación entre la elección subóptima y el nivel de conducta de sign-tracking.

Se apoya la idea de una nula influencia de las palancas en la promoción de suboptimalidad, y se muestra la primera evidencia de no-interacción entre longitudes de TL y el uso de palancas como estímulos discriminativos.

Quizá los resultados de Cunningham y Shahan se debieron a una particularidad de su procedimiento: la alta saliencia de la combinación tono + luz sumada a la baja saliencia del *blackout*.

Se ha mostrado que incrementar la duración de los TL promueve la elección subóptima en palomas. Se ha investigado el efecto de variar la longitud de los IL, pero esos resultados no son tan claros, aunque indican que la suboptimalidad de las palomas disminuye con IL más largos, pero las ratas no parecen ser afectadas por esta manipulación. Esto apoya la noción de que los procesos que determinan la elección subóptima en ratas y palomas son distintos. El *temporal information-theoretic model* sugiere que la diferencia está en las variables que contribuyen al mecanismo de ponderación del modelo: las palomas toman en cuenta los IL, pero las ratas no. Este experimento indica que los TL tampoco contribuyen a la ponderación de las ratas, o que lo hacen solo en ciertas condiciones. Otra posibilidad es que las diferencias sean generadas por la saliencia incentiva de los estímulos (quizá las palancas no tienen para las ratas la misma saliencia que las teclas para las palomas, como sugieren procedimientos de automoldeamiento).

En un estudio participantes humanos eligieron entre alternativas discriminativas y no-discriminativas en condiciones definidas por la duración del TL (2.5 a 40s). Los resultados indicaron que los participantes prefieren la alternativa no-discriminativa en duraciones cortas, pero cambian en duraciones mayores a 20s. El modelo de utilidad anticipatoria explica las diferencias asumiendo que elegir la alternativa discriminativa dispara un error de predicción de recompensa que amplifica una señal anticipatoria de acuerdo con el tipo de estímulo encontrado: *savouring* y *dreading*. El modelo asume que la influencia de las señales anticipatorias para determinar el valor crece con la duración del estímulo y la valencia de las recompensas primarias. Pero la anticipación negativa es menos duradera que la positiva. Esto permite predecir mayor suboptimalidad para TL más largos (hasta cierta duración) y para mayores diferencias entre *savouring* y *dreading*. En palomas se asume que la influencia del *dreading* es nula, lo que lleva a más suboptimalidad. En humanos, el *dreading* es menor que el *savouring*, así que la preferencia fluctúa de óptima a subóptima cuando se alargan las duraciones. Los resultados actuales indican que las ratas podrían no tener *savouring* o que la influencia del *dreading* es equivalente a la del *savouring* y ambos efectos se cancelan. La segunda posibilidad es consistente con el mayor efecto de inhibición condicionada reportado en ratas comparadas con palomas.

La mayoría de los modelos de elección subóptima en palomas dicen explícitamente que el inhibidor es irrelevante, pero los resultados de ratas dicen lo contrario. Para explicar la presencia cada vez más demostrada de optimalidad es necesario asumir que el estímulo negativo tiene un impacto en la determinación del valor de la alternativa discriminativa.

Esta idea es apoyada por un análisis que resalta que en la mayoría de los estudios que reportan suboptimalidad en ratas la saliencia del predictor de reforzamiento es mayor que la del *blackout* que predice no-reforzamiento. Aunque cabe resaltar que la mayoría de los estudios que reportan optimalidad han usado ratas Wistar, mientras que los que reportan suboptimalidad usan otras cepas. Se requiere ahondar en el efecto de la saliencia relativa de los estímulos y la cepa.

Cuando los estímulos son luces o palancas, las preferencias no son afectadas por la duración de los TL. La excepción a ese patrón ocurre cuando los estímulos son combinaciones salientes de tono y luz, lo que sugiere una interacción entre las características de los estímulos y la duración de los TL. Se requiere mayor investigación para desarrollar un modelo general de elección subóptima. Este modelo deberá tener implicaciones directas para el entendimiento de la conducta maladaptativa en humanos.