

THE ROLE OF INHIBITION IN THE SUBOPTIMAL CHOICE TASK

VALERIA V. GONZÁLEZ

AARON P. BLAISDELL

2021

En elección subóptima las palomas prefieren la alternativa con señales predictivas a pesar de tener una menor densidad de reforzamiento. El procedimiento tiene su base en el paradigma de respuestas de observación, en el cual los animales están dispuestos a responder para obtener estímulos que informan sobre el programa vigente, a pesar de no poder cambiarlo.

Una interpretación viene de la teoría de información, según la cual la información positiva y negativa deberían ser igualmente valiosas. La información positiva es directamente valiosa, pero la negativa lo es porque permite redirigir los recursos a otro lugar (Vasconcelos). Pero experimentos más recientes han mostrado que las buenas y malas noticias no son iguales, en tanto que los animales prefieren las buenas que las malas noticias. Pero los experimentos no indican si los animales aprenden que las malas noticias son señales de ausencia de reforzamiento, o si aprenden a ignorar la señal. Evidencia indica que los humanos prefieren las malas noticias que la ausencia de noticias, y las ratas prefieren un choque inescapable señalado que uno no señalado.

Se ha investigado poco la *preferencia* por la información separada de las respuestas de observación. Un procedimiento en el que se hizo usó un laberinto en E en el que ambos brazos proveían la misma probabilidad de comida, pero uno informaba anticipadamente mediante el color de las paredes y otro no. Se encontró que las ratas prefieren la información anticipada a pesar de que esta no cambie el resultado final. Más tarde, el grupo de Zentall encontró que esta preferencia por la información sucede aun si implica menor acceso a comida.

Varias explicaciones presumen que el S^- es ignorado y no tiene efecto en la elección (*e.g.*, el reinforcement rate model), lo que contrasta con la evidencia que indica que los animales prefieren un mal resultado señalado que uno sin señalar. Del mismo modo, la explicación de la información temporal de Cunningham y Shahan supone que el S^- , al no señalar ninguna relación temporal con un reforzador, es ignorado. Por otro lado, la hipótesis $\Delta - \Sigma$ supone que los animales deben prestar atención a todas las probabilidades, incluida la de S^- . En resumen, no es claro si S^- afecta la elección; y si lo hace, no es claro si se debe a un proceso perceptual, atencional o de aprendizaje.

Se propone que el S^- sí contribuye a la elección subóptima al volverse un inhibidor

condicionado. Laude, utilizando el procedimiento de elección subóptima de magnitudes, encontró que se desarrolla inhibición condicionada, pero esta se pierde con entrenamiento extendido. No se provee una explicación para este fenómeno, pero una posibilidad es que el S^- inicialmente redujera las respuestas debido a inhibición externa, la cual menguó mientras el S^- se hizo más familiar.

El uso del procedimiento de magnitudes implica que todas las teclas son informativas. Es incierto si el mismo mecanismo causa la elección subóptima en ambos procedimientos. Por otro lado, S^- era una línea vertical, mientras que los demás estímulos eran colores. Se ha mostrado que los colores son más salientes que las formas, por lo que usar al estímulo menos saliente como S^- puede haber contribuido a que este estímulo perdiera control sobre la conducta. Además, solo se utilizó una prueba de S^+S^- para determinar la inhibición. Se requieren de más pruebas (por ejemplo, con estímulos novedosos) para descartar la inhibición externa.

Fortes incrementó la probabilidad y longitud del S^- , manipulaciones que deberían incrementar la inhibición condicionada, pero no hubo cambios en la preferencia. Trujano, por otro lado, encontró que la inhibición en ratas incrementa con el entrenamiento, pero no encontró elección subóptima y concluyó que la diferencia en resultados está relacionada con una diferencia en el impacto de los inhibidores condicionados.

Estos experimentos pretendían determinar si S^- adquiría propiedades inhibitorias en un procedimiento más convencional, sin las desviaciones de Laude; y probar la relación entre el desarrollo de la inhibición a S^- y la fuerza de la preferencia subóptima. Trujano dijo que las ratas no son subóptimas debido a la inhibición, pero evidencia reciente muestra que sí pueden ser subóptimas (Cunningham, Ojeda).

En el experimento 1 se entrenaron palomas en automoldeamiento solo con los estímulos de TL ($S^+, S^-, S3, S4$), señalando 1, 0, .5 y .5 de probabilidad de reforzamiento. Se presentaron ensayos de sondeo con compuestos de S^+S^- , $S3S4$, $S3S^-$, y S^+S4 para medir las propiedades inhibitorias de S^- . Si es inhibidor, se esperaría una reducción en S^+S^- comparada con S^+ o $S3S4$. También se probaron los compuestos $S3S^-$ y S^+S4 para comparar las propiedades excitatorias de un estímulo parcialmente reforzado con las de uno continuamente reforzado.

El experimento 2 midió la inhibición de S^- después de entrenamiento en elección subóptima. Después de establecerse una preferencia confiable, se probó la inhibición pavloviana con una prueba de sumación. Se presentaron ensayos de sondeo con compuestos de S^+S^- , pero también S^+ en compuesto con un estímulo novedoso para descartar la inhibición externa. Si S^- se vuelve inhibidor, debería evocar una tasa de respuestas menor que la del compuesto con estímulo novedoso.

Para explorar la relación entre la inhibición y la elección subóptima en el experimento 3 las palomas fueron entrenadas en elección subóptima con estímulos tanto en eslabones iniciales como terminales. Se midió el desarrollo de inhibición ante S^- usando pruebas de sumación con compuestos como los de los experimentos 1 y 2, y el desarrollo de la preferencia por el estímulo de eslabón inicial subóptimo en ensayos de elección durante el entrenamiento. Si la preferencia subóptima está relacionada con la inhibición condicionada, entonces se predice que la fuerza de la preferencia subóptima estará correlacionada con la inhibición condicionada. Una ausencia de correlación implicaría una ausencia de

relación entre las variables. Y si la elección subóptima depende de ignorar el S^- como sugiere Laude, entonces la fuerza de la preferencia subóptima debería correlacionar negativamente con la inhibición. Se usarán pruebas de sumación negativa, en las cuales el S^+ y el supuesto estímulo inhibitorio S^- se presentan en compuesto y las respuestas a ellos se comparan con las respuestas a presentaciones no reforzadas de S^+ .

1. Experiment 1

Primero se quiso determinar si los estímulos de los eslabones terminales entrenados por sí solos adquirirían propiedades excitatorias o inhibitorias. Se realizó un procedimiento de automoldeamiento pavloviano con los cuatro estímulos y las probabilidades de reforzamiento dichas, y se midió la tasa de respuesta a cada estímulo. Las tasas de respuesta a los compuestos se compararon con las tasas de respuesta a los estímulos individuales y otros compuestos. Se predice una reducción en la tasa de respuesta en los compuestos S^+S^- y $S3S^-$.

1.1. Method

1.1.1. Subjects

Cinco palomas con experiencia en múltiples procedimientos, pero ingenuas con respecto a este procedimiento y estímulos.

1.1.2. Materials

Los estímulos se presentaron en una pantalla LCD mediante PsychoPy. Las picadas se detectaron mediante una pantalla táctil infrarroja.

Los estímulos estaban compuestos por dos círculos (rojo, verde, amarillo o azul) alineados vertical u horizontalmente.

1.1.3. Procedure

Se entrenó a las palomas a picar cada estímulo en FR1 y FR10.

En sesiones diarias se presentó el procedimiento de automoldeamiento, que consistía en 80 ensayos. S^+ se presentaba 8 veces seguido de comida, S^- se presentaba 32 veces seguido de omisión, $S3$ se presentaba 8 veces seguido en 4 de comida y 4 de omisión, y $S4$ se presentaba 32 veces seguido en 16 por comida y 16 por omisión.

Tras 15 sesiones de entrenamiento se hicieron pruebas en bloques de 5 sesiones intercaladas con bloques de 5 sesiones de solo entrenamiento. Cada sesión de prueba de 56 ensayos comenzaba con diez presentaciones de los estímulos de entrenamiento. Después, nueve ensayos no reforzados se presentaban aleatoriamente entre ensayos de entrenamiento por lo restante de la sesión. Los ensayos de prueba no eran reforzados y duraban 30 segundos.

1.1.4. Data analysis

Se calculó la tasa de respuesta a cada estímulo y se promedió en bloques de cinco sesiones.

Un análisis de varianza factorial de medidas repetidas se implementó para analizar los bloques de entrenamiento y prueba.

1.1.5. Results

La tasa de respuesta a S^- fue menor que la de todos los demás estímulos en solitario. Hubo una menor tasa de reforzamiento a S^+S^- y $S3S^-$ que a los demás compuestos. Esto es importante ya que indica que presentar dos estímulos juntos en un compuesto novedoso no produce inhibición externa ni decremento de generalización.

2. Experiment 2

El experimento 2 pretendía replicar los resultados de la prueba de sumación después de entrenamiento en elección subóptima, y controlar la inhibición externa con la inclusión de un estímulo novedoso en las pruebas. Si la presentación de S^+ junto con un estímulo no entrenado no produce sumación negativa en la prueba, entonces se puede descartar la inhibición externa como explicación para la inhibición observada en S^+S^- , apoyando la verdadera inhibición condicionada de S^- .

2.1. Method

2.1.1. Subjects

Ocho palomas nuevas con experiencia en otros procedimientos.

2.1.2. Materials

Mismo aparato que antes. Se usaron siete estímulos. Dos eran círculos de color con patrones geométricos usados como estímulos de eslabón inicial. De los cinco restantes, se usaron cuatro al azar para cada paloma como estímulos de eslabón terminal. Los estímulos eran dos círculos de color (rojo, verde, amarillo, azul o naranja) alineados vertical u horizontalmente.

2.1.3. Procedure

Ensayos forzados y de elección. Al picar un estímulo de eslabón inicial, ambos desaparecían y el elegido era reemplazado por su estímulo de eslabón terminal durante 30 segundos. Las probabilidades de reforzamiento eran 20 % para la alternativa discriminativa y 50 % para la no discriminativa. Durante el ITI la pantalla se oscurecía por 10 segundos. Se usó un solo estímulo para la alternativa discriminativa debido a que se ha encontrado que esto no impacta la preferencia (nota, esto solo es cierto en palomas).

La localización de las alternativas en la pantalla (izquierda/derecha) fue aleatorizada en cada ensayo.

Al inicio del entrenamiento el eslabón terminal duraba 10 segundos. Después se extendió a 30.

La primera sesión del entrenamiento también incluyó dos presentaciones de 10 segundos de la clave E, que después se usó para medir la inhibición externa.

Hubo dos sesiones en las que se midió la inhibición de S^- con prueba de sumación. La prueba consistía en 20 ensayos, 10 libres y 10 de prueba. Los ensayos de prueba consistían en 4 presentaciones de los compuestos S^+S^- y S^+E^- y dos presentaciones no reforzadas de S^+ . Cada estímulo era presentado durante 30 segundos.

2.1.4. Data analysis

Se calculó la preferencia por la alternativa informativa por sesión y la tasa de respuesta a cada estímulo de eslabón terminal.

2.1.5. Results

Todas las palomas mostraron preferencia subóptima. Se encontraron menores tasas de respuesta a S^+S^- que a S^+E y S^+ , y la tasa de S^+E fue menor que la de S^+ . Así, aunque hubo evidencia de una pequeña cantidad de inhibición externa por E , hubo un efecto mucho mayor de inhibición condicionada producido por S^- .

3. Experiment 3

Habiendo establecido que S^- es inhibitorio, se puede probar la hipótesis de que el desarrollo de la preferencia subóptima se relaciona con la adquisición de inhibición por S^- . Para ello, en elección subóptima se introdujeron periódicamente en el entrenamiento en las cuales se presentaban los compuestos S^+S4 , $S3S4$, S^+S^- , y $S3S^-$. Se esperaba el desarrollo de preferencia subóptima al decrementar la tasa de respuestas a S^- , y una correlación entre la fuerza de la inhibición y la fuerza de la preferencia subóptima.

3.1. Method

3.1.1. Subjects

Los mismos del experimento 1.

3.1.2. Materials

Mismo aparato, pero estímulos nuevos. Los mismos estímulos de eslabón inicial se repitieron, y se introdujeron cuatro pares de figuras geométricas con patrones blancos y negros como estímulos de eslabón terminal, todo organizado vertical u horizontalmente.

3.2. Procedure

Todos los estímulos fueron presentados individualmente como en los experimentos 1 y 2. Después se entrenó elección subóptima.

Para la prueba se presentaron cuatro bloques de cinco sesiones de prueba. Una sesión de prueba tenía 96 ensayos: 80 de entrenamiento y 16 de prueba no reforzados. Los primeros 10 ensayos de cada sesión de prueba siempre eran forzados. Después, 50 forzados, 20 libres y 16 de prueba se mezclaban aleatoriamente. Los ensayos de prueba consistían en la presentación de uno de cuatro compuestos presentado durante 30 segundos.

3.3. Data analysis

Se calculó la preferencia y tasa de respuesta. Los datos se normalizaron debido a la variabilidad entre palomas para permitir hacer comparaciones.

3.4. Results

La tasa de desarrollo de la preferencia subóptima varió entre palomas.

Una tasa de respuestas similar se mantuvo entre bloques para S^+ , $S3$ y $S4$, mientras que la tasa de S^- disminuyó con los bloques hasta llegar casi a 0.

En general la tasa de respuesta ante compuestos fue menor que ante estímulos solitarios, lo que sugiere *generalization decrement* o que los animales aprendían que los compuestos nunca eran reforzados.

La tasa ante S^+ era mayor que ante S^+S^- y S^+S4 , lo que sugiere que las pruebas repetidas a través de varios ciclos de entrenamiento/prueba pueden haber resultado en que las palomas aprendieran que los ensayos compuestos no eran reforzados. Dado que se observó *generalization decrement* para todos los compuestos, los análisis usan la tasa de respuesta a S^+ y S^- como medidas de excitación e inhibición.

Se hipotetizó que aprender la inhibición de S^- correlacionaría con la preferencia subóptima, pero otra posibilidad era que aprender sobre S^+ también siguiera a la preferencia. Dado que las palomas desarrollaron preferencia a tasas distintas se hicieron correlaciones entre la tasa ante S^+ y la preferencia subóptima por cada ave. Se encontró una correlación positiva fuerte entre S^+ y la preferencia para algunas palomas, pero pobre o ninguna en otras.

La correlación entre S^- y la preferencia fue fuertemente negativa para algunas y débilmente negativa para otras. Una prueba r de Pearson mostró que la correlación promedio entre la tasa de S^- y la preferencia era menor a cero, y la correlación entre la tasa de S^+ y la preferencia era mayor a cero. Esto apoya la hipótesis de que la inhibición de S^- se relaciona con el desarrollo de la preferencia subóptima.

Dado que la cantidad de entrenamiento puede ofuscar los resultados, se hizo un análisis por cada paloma en asíntota. La correlación entre niveles asíntóticos de preferencia y S^+ fue leve pero positiva, y entre preferencia y S^- fue leve y negativa. Aunque esos resultados no fueron significativos, sugieren una relación entre el aprendizaje de S^+ y S^- y la preferencia subóptima.

4. General discussion

Se probó la hipótesis de que S^- desarrolla propiedades inhibitorias, y ese desarrollo se relaciona con el desarrollo de preferencia subóptima.

En el primer experimento se encontró que S^- desarrolla inhibición condicionada. Más aun, se encontró que la inhibición tenía un efecto supresor más fuerte en estímulos parcialmente reforzados que en estímulos continuamente reforzados.

En el experimento 2 se replicó el patrón de respuesta a los estímulos compuestos, pero después de entrenamiento en elección subóptima. Se encontró evidencia de inhibición condicionada independiente de la inhibición externa. Es decir, las tasas de respuesta a S^+S^- eran menores que a S^+SE .

En el experimento 3 se replicó elección subóptima y se encontró una correlación negativa entre la fuerza de la preferencia subóptima y la tasa de respuesta a S^- , lo que indica que según cada ave adquirió la preferencia subóptima, el S^- se hizo más inhibitorio.

Se observó correlación positiva entre las tasas a S^+ y la fuerza de la preferencia, y negativa entre las tasas de S^- y la fuerza de la preferencia al final del entrenamiento, lo que indica que cuanto mayor es la inhibición de S^- , más fuerte es el nivel asintótico de preferencia subóptima. Aunque a este análisis le faltó poder, y los datos son solo correlacionales.

Una posible explicación es que el desarrollo de la preferencia por el estímulo inicial de la alternativa subóptima depende del desarrollo de inhibición al estímulo S^- de eslabón terminal. Aunque ambas podrían depender de un tercer proceso. Una tercera alternativa es que el aprendizaje de las propiedades de S^+ y S^- actúa en conjunto para el desarrollo de la preferencia.

Estos resultados contradicen estudios previos. Laude no encontró evidencia de relación entre la inhibición de S^- y la elección subóptima, y de hecho encontró que la inhibición se pierde durante el entrenamiento. Sus resultados podrían indicar una de-correlación entre inhibición y preferencia.

Se identificaron limitaciones en su estudio que impiden apoyar fuertemente sus conclusiones. En este estudio se usó la versión tradicional de probabilidad de la tarea. También, el estímulo S^- en este experimento era de la misma dimensión que los otros estímulos terminales, evitando así posibles efectos de confusión. Además, los estímulos de eslabones terminales se contrabalancearon entre palomas. Más aun, se midió la inhibición mediante pruebas de sumación incluyendo un estímulo no entrenado para controlar inhibición externa. Todavía más, se midió la inhibición en varios puntos del entrenamiento para rastrear su desarrollo. Y finalmente, se analizaron datos individuales para medir diferencias en aprendizaje y ejecución.

Trujano reporta encontrar inhibición en elección subóptima en ratas, pero no preferencia subóptima. Atribuyen la diferencia en preferencia a la inhibición, y sugieren que las palomas y las ratas no codifican la tarea de la misma manera. Estos resultados contrastan, y más aun, otros investigadores han reportado suboptimalidad en ratas cuando se disminuye la diferencia en probabilidades de reforzamiento entre alternativas (Ojeda) o se aumenta la demora de reforzamiento (Cunningham), o cuando se usan palancas (Chow). Sería interesante evaluar la inhibición en una tarea en la cual las ratas desarrollen prefe-

rencia por la alternativa subóptima. Quizá los resultados de Trujano se explican por una diferencia entre parámetros y no por una diferencia intrínseca entre especies.

Estos resultados pueden aclarar el papel de S^- en modelos de elección subóptima. El modelo ecológico y la hipótesis de información temporal suponen que el S^- es ignorado en el sentido de que no juega un papel en la elección. Estos resultados desafían al menos los mecanismos presumidos por esos modelos: si un animal aprende de un estímulo, esto debería contribuir a la elección (no estoy seguro de esto).

El modelo $\Delta - \Sigma$, en contraste, trata a S^- solo como otro valor que se contrasta para asignar valor a una alternativa. Es posible que la inhibición de S^- incremente el atractivo de S^+ , lo que resultaría en una sobre-ponderación.

En términos de información, la evidencia de que los animales aprenden que S^- se vuelve inhibitorio, y que la fuerza de la inhibición siga a la fuerza de la preferencia subóptima, sugiere que los animales prefieren la información a la no-información, incluso cuando la información son malas noticias. Esto contrasta con supuestos de la literatura de observación, que indican que solo la información positiva es preferida.

Se propone que al adquirir propiedades de inhibición, el S^- predice la ausencia explícita de comida. Se sugiere que el que S^- se vuelva inhibidor puede estar causalmente relacionado con la preferencia subóptima, aunque esto debe probarse. Esto se suma a modelos que indican que el efecto combinado de S^+ y S^- produce la preferencia subóptima. Puede ser que aprender de S^- incremente el valor de S^+ , sesgando la preferencia. Esto está en línea con experimentos que no encuentran un gran impacto de la tasa global de reforzamiento cuando ambas alternativas son informativas. Se debe investigar más la relación entre la inhibición condicionada de S^- y el desarrollo de la preferencia subóptima.