

# PARAMETRIC SHIFT FROM RATIONAL TO IRRATIONAL DECISIONS IN MICE

NATHAN A. SCHNEIDER, BENJAMIN BALLINTYN,  
DONALD KATZ, JOHN LISMAN, HYUN-JAE PI

2021

Los modelos económicos clásicos de toma de decisiones indican que los animales analizan los costos y beneficios de sus opciones disponibles y optimizan mediante sus decisiones. Estas predicciones no siempre se cumplen, sin embargo. Esto se debe a que, si bien las características ostensibles del ambiente son objetivas, la toma de decisiones se ve influida por factores encubiertos como el estado intrínseco del sujeto. Por ello los modelos de decisión deben incluir ambos factores.

Estudios recientes aplican lecturas y marcos de referencia conductuales para investigar estos factores encubiertos, y han revelado conducta similar a la falacia del costo hundido.

Para estudiar más en profundidad estos factores ostensibles y encubiertos, se diseñó una tarea de toma de decisiones económicas en la que el costo y el beneficio se pueden ajustar independientemente, y en la que hay una solución óptima identificable. En este estudio se define la optimalidad con respecto a el modelo de *expectation of ratios* en teoría de forrajeo; y racionalidad se refiere al estado en el cual un animal toma decisiones basado en la valoración objetiva y no en factores encubiertos.

En cada ensayo los ratones deciden presionar una palanca en un programa de razón progresiva o de razón fija (PR; FR) con distintas cantidades de agua como recompensa. El cambio en la preferencia fue calculado mediante el requisito de razón progresiva y el punto de indiferencia en el cual los valores de ambas alternativas eran equivalentes. Se evaluó si los ratones ajustaban sus decisiones de acuerdo con los parámetros de costo-beneficio y se encontró que era así.

Al comparar las decisiones con las soluciones óptimas se pudo cuantificar la desviación de la optimalidad a través de los valores de los parámetros. Se encontró que los costos tienen un mayor efecto que los beneficios en la conducta subóptima. Además, los ratones mostraron conducta similar a la de estudios de costo hundido. La contribución de esta falacia a los resultados subóptimos era específica al parámetro en el sentido de que se identificó un régimen en el espacio de parámetros en el cual el parámetro de costo y la susceptibilidad al costo hundido no tenían ningún impacto significativo.

## MÉTODO

Se utilizaron 10 ratones de la cepa C57BL/6.

La cámara experimental contenía tres nosepokes con detector de entradas y dos palancas debajo de los nosepoke laterales. Dentro del nosepoke central se entregaba agua como reforzador.

En 5 fases de preentrenamiento se entrenó a los ratones para presionar las palancas para obtener comida hasta llegar a un programa de RF10.

**Switching task.** Había dos programas vigentes: PR y FR. El de PR estaba asociado con una cantidad grande de agua y el de FR con una cantidad pequeña. El FR podía ser de 6 o 12 presiones. Al comienzo de un ensayo, los dos nosepokes se encendían ligeramente. Cuando un ratón elegía un lado y comenzaba a responder en su palanca correspondiente, el nosepoke se iluminaba con cada vez mayor intensidad hasta que el nosepoke central se encendía también indicando la entrega del reforzador. Los ratones podían cambiar de lado en cualquier momento, pero si lo hacían antes de terminar las presiones necesarias en uno de los lados se presentaba un ruido blanco como castigo y se terminaba el ensayo, que se clasificaba como incompleto. Lo mismo sucedía si el ratón comenzaba a responder pero antes de terminar se detenía durante más de 10 segundos. Las sesiones terminaban si un ratón no respondía durante más de 5 minutos. Tras tres horas, la sesión se daba por concluida y se entregaba agua adicional para mantener el peso de los animales.

**Biological vs technical replication.** Al no encontrarse diferencias entre machos y hembras, sus datos se combinaron. Se recolectaron cuatro sesiones para cada ajuste de los parámetros (160 en total: 4 sesiones por cuatro pares de parámetros por 10 ratones). La replicación biológica son los 10 ratones, y la técnica son las 4 sesiones por cada ajuste de los parámetros.

**Data analysis.** La identificación de los puntos de indiferencia se realizó ajustando los datos binarios a una función sigmoide (la función Boltzmann):

$$f(x) = \frac{1}{1 + e^{\frac{(x-x_0)}{\tau}}},$$

donde  $x_0$  y  $\tau$  son un umbral del 50 % y una pendiente. Al asignar el valor de elegir PR como 1, y el valor de elegir FR como 0, y asumir que la curva comienza en PR y termina en FR, la curva de ajuste se genera. Un número de ensayo de indiferencia se estimó en el punto en el que la curva sigmoide cruzaba el punto medio, y la cantidad de presiones de palanca requeridas en ese ensayo por la opción de PR se extrajo de los datos. Este número de presiones requeridas provee el requisito de PR en el punto de indiferencia de la sesión. La significancia estadística entre los cuatro pares de parámetros se probó con un análisis no paramétrico.

**Analysis of sensitivity to sunk cost.** El análisis comenzó tomando todos los ensayos de PR intentados en una sesión y separándolos por el número de presiones de palanca remanentes después de que el ratón emitió una presión. Después se determinó cuántos ensayos en cada grupo fueron completados. Todos los ensayos de sesiones de la misma configuración de parámetros se agregaron para cada ratón, y se promediaron entre ratones. Para evitar sesgar los datos por un solo ensayo se incluyeron solo los datos si había más de cinco ensayos completados en esa condición. La proporción de ensayos completados para cada valor de presiones remanentes se calculó dividiendo el número promedio de

ensayos completados entre el número promedio de ensayos totales. Estos valores luego fueron ajustados de forma lineal. Todos los ensayos de PR se muestrearon de nuevo después de que 5, 10, 15,..., 45 presiones habían sido emitidas y las proporciones correspondientes de ensayos recompensados fueron calculadas u ajustadas a nuevas líneas de regresión. Las regresiones solo se incluyeron en el análisis si al menos incluían 15 datos. El efecto global de la sensibilidad al costo hundido en cada contexto se calculó con un ANOVA de dos vías con la probabilidad como variable dependiente y el número de presiones remanentes x grupos de costo hundido como factores.

**Calculation of EoR optimality.** Se han propuesto dos monedas de cambio como *proxy* de la adecuación evolutiva que los animales buscan maximizar. Una es la tasa de ingesta energética a largo plazo (llamada *ratio of expectations*, o RoE), y la tasa a corto plazo (*expectation of ratios*, EoR). En términos de energía total ingerida a través del tiempo, RoE es la mejor cantidad para maximizar, sin embargo experimentos con starlings han mostrado que EoR podría ser lo que se utiliza en realidad, por lo que se computó la estrategia óptima en términos de EoR y se comparó con la conducta de los ratones. EoR para un número  $N$  de ensayos es

$$EoR(N) = \frac{1}{N} \sum_{k=1}^N \frac{r_k}{p_k},$$

donde  $N$  es el número de ensayos total,  $r_k$  es la recompensa en  $\mu$ l recibida en el ensayo  $k$ , y  $p_k$  es el número de presiones de palanca en el ensayo  $k$ . Este caso es distinto de otros estudios con EoR dado que el denominador de cada fracción está en términos de un costo energético y no temporal. Para cada tipo de sesión (2xFR6, 2xFR12, 5xFR6, 5xFR12) se calcula el número óptimo de ensayos en el programa de PR ( $N_{PR}^*$ ). Para esto se debe encontrar cuál  $N_{PR}$  da el EoR más alto para una  $N$  dada. Esto se puede encontrar dividiendo la ecuación para EoR en términos para ensayos en PR y ensayos en FR:

$$EoR(N) = \frac{1}{N} \left( r_{PR} \times \sum_{k=1}^{N_{PR}} \frac{1}{k+1} + N_{FR} \times \frac{r_{FR}}{p_{FR}} \right),$$

donde  $r_{PR}$  es la recompensa del lado de PR,  $r_{FR}$  es la recompensa del lado de FR,  $p_{FR}$  es el número de presiones requeridas para la recompensa en FR, y  $N_{FR}$  es el número de ensayos completados en el lado de FR. Se asume que un agente óptimo no aborta ningún ensayo dado que eso disminuiría EoR. Al sustituir  $N_{FR} = N - N_{PR}$  y evaluar la suma, se obtiene

$$EoR(N) = \frac{1}{N} \left( r_{PR} \times \left( \psi^{(0)}(N_{PR} + 2) + \gamma - 1 \right) + (N - N_{PR}) \times \frac{r_{FR}}{p_{FR}} \right),$$

donde  $\psi^{(0)}(x)$  es la función digamma, y  $\gamma$  es la constante Euler-Mascheroni. Para encontrar  $N_{PR}^*$  se toma la derivada de esta expresión con respecto a  $N_{PR}$

$$\frac{dEoR(N)}{dN_{PR}} = \frac{1}{N} \left( r_{PR} \times \psi^{(1)}(N_{PR} + 2) - \frac{r_{FR}}{p_{FR}} \right),$$

donde  $\psi^{(1)}(x)$  es ahora la función trigamma (derivada de digamma). Estableciendo la derivada en 0 y resolviendo para  $\psi^{(1)}(N_{PR} + 2)$  obtenemos

$$\psi^{(1)}(N_{PR}^* + 2) = \frac{r_{FR}}{r_{PR} \times p_{FR}}$$

de donde podemos obtener aproximaciones numéricas para  $N_{PR}^*$ , las cuales son 10.5, 22.5, 28.5 y 58.5 para 2xFR6, 2xFR12, 5xFR6 y 5xFR12 respectivamente. Estos números tienen una interpretación directa: por ejemplo, en el ensayo 10 de PR en una sesión 2xFR6, el ratón recibirá 6  $\mu$ l de agua por 11 presiones de palanca, lo que da una mejor tasa que el lado de FR (3  $\mu$ l por 6 presiones). En el ensayo 11 estas tasas serán iguales (6/12 y 3/6). El mismo patrón es verdad para los cuatro tipos de sesiones, por lo tanto, el 0.5 en cada uno de los  $N_{PR}^*$  anteriores refleja que es equivalente dejar de elegir el lado PR cuando las tasas son iguales o un ensayo antes. Para este trabajo se tomó como  $N_{PR}^*$  óptimo para cada tipo de sesión a 10, 22, 28 y 58. Para calcular después la fracción de optimalidad EoR de cada ratón en cada sesión, se calculó la EoR observada para cada ratón en una sesión determinada (la media de la recompensa por ensayo dividida entre el número de presiones a la palanca, incluyendo las de ensayos abortados), y eso se dividió entre el EoR óptimo dado el número total de ensayos que el ratón realizó y el tipo de sesión.

**Comparison to random behavior.** Para computar las distribuciones de optimalidad producidas por conducta aleatoria para cada conjunto de parámetros experimentales, se simularon 10,000 agentes que elegían aleatoriamente para cada número de ensayos totales de 1 a 1000. En cada ensayo simulado había una probabilidad de 50% de elegir FR o PR. El EoR observado de estas elecciones aleatorias fue comparado con el EoR óptimo para ese número de elecciones. Esto produjo, para cada número de ensayos totales, una distribución muestral (de 10,000 muestras) de optimalidades de EoR obtenidas por elección aleatoria. Después se computó la optimalidad EoR corregida descrita por la ecuación

$$\text{corrected EoR optimality} = \frac{EoR_{mouse}^{optimality}(n) - EoR_{random}^{optimality}(n)}{1 - EoR_{random}^{optimality}(n)}.$$

En esta ecuación se interpreta  $EoR_{random}^{optimality}(n)$  como la media de la distribución de optimalidades de los agentes aleatorios para el número de ensayos que el ratón completó en la sesión  $n$ . Estas tasas cuantifican la mejora en los resultados dadas las elecciones de los ratones comparadas con la elección aleatoria para un parámetro experimental particular y un dado número de ensayos.

## RESULTADOS

**A flexible economic choice behavior for mice.** Ratones bajo restricción de agua debía presionar palancas para obtener una recompensa. La palanca de FR proveía de una cantidad pequeña de agua, mientras que la palanca de PR proveía una cantidad mayor, pero su razón incrementaba tras cada entrega de reforzador. Los ratones preferían confiablemente la palanca de PR al inicio de cada sesión dado que les daba una recompensa mayor con poco esfuerzo. Mientras su requisito incrementaba, el costo (número de presiones) superaba al beneficio, y los ratones cambiaban su preferencia hacia la palanca de FR, apoyando la hipótesis de que los ratones evalúan los valores relativos de las opciones en términos tanto de costos como de beneficios.

Al fijar el parámetro de beneficio y variar el costo, su contribución se puede evaluar de forma semi-independiente. En este estudio se eligieron cuatro combinaciones de parámetros al cruzar la razón entre las recompensas (2:1 o 5:1) con el número de presiones requeridas para la recompensa de FR (6 y 12). La nomenclatura es: 2xFR12 se refiere a que el volumen de la recompensa grande es el doble de la pequeña, y que el requisito de FR es de 12 presiones. Las cuatro combinaciones son 2xFR6, 2xFR12, 5xFR6, y 5xFR12. Un análisis del tiempo entre presiones de palancas mostró que mientras hay variabilidad en los intervalos entre respuesta cuando se incrementa el requisito, no hay un incremento relevante sino hasta llegar a alrededor de 60 presiones, por lo que se puede saber que la tarea no resultaba en fatiga rápidamente.

***Mice adjust switching decisions proportional to the values of the session parameters.*** Todos los ratones podían hacer decisiones de cambio evaluando con precisión la relación de costo-beneficio de las opciones presentadas, pero el punto en que se realizó la decisión cambió entre condiciones. El cambio a FR pasa más pronto en 2xFR6 y más tarde en 5FR12 debido probablemente a los distintos valores relativos del PR. El valor de la recompensa grande es más bajo en 2xFR6 dado que la recompensa de PR es baja y el esfuerzo y tiempo de FR también son bajos. El valor de PR es más alto en 5xFR12 dado que la recompensa es alta y la elección de la otra alternativa requiere de más esfuerzo y tiempo.

Los valores de PR de 2xFR12 y 5xFR6 están en algún punto intermedio. Reflejando estos valores, el número de recompensas obtenidas en PR incrementó según incrementó su valor.

Los ensayos sin completar también fueron consistentes con estas conclusiones: los ratones fallaron en completar ensayos más a menudo cuando los valores de PR eran más altos, y el fallo era más frecuente en el lado de PR. Incrementar el costo (número de presiones de FR) contribuyó a incrementar los ensayos incompletos. El cambio en el costo también tuvo un mayor efecto en la ejecución según indica el número total de ensayos y el agua obtenida. Estos resultados indican que los ratones pueden diferenciar entre los valores relativos de los parámetros de la sesión y ajustar su conducta en función de ellos.

***Quantification of switching decisions by indifference points.*** En esta tarea el PR es inicialmente más valioso que el FR dado que su recompensa es mayor, pero mientras el requisito de PR incrementa, su valor disminuye.

Se estimó el requisito de PR en los puntos de indiferencia ajustando los datos experimentales a una función binaria (dado que solo hay dos alternativas) que representa las elecciones: la función sigmoide de Boltzmann. El número del ensayo de indiferencia se estimó en el punto en el que la curva sigmoide cruzó la línea media, y luego se extrajo de los datos el número de presiones a la palanca en ese ensayo. Los resultados mostraron que el punto de indiferencia era menor en 2xFR6 y mayor en 5xFR12, mientras que 2xFR12 y 5xFR6 tenían valores intermedios. Esto indica que los ratones podían ajustar su decisión de cambio de acuerdo con el valor de las alternativas.

***The cost parameter contributes more strongly to suboptimality.*** Dada la estructura de ensayos discretos de la tarea, se estudió la conducta de los ratones con referencia al modelo EoR en el que un agente racional maximiza la tasa de ganancia sobre costo promedio por ensayo. Este EoR es una versión modificada para dar cuenta de la

naturaleza discreta de las presiones a la palanca. EoR está dado por la ecuación

$$Expectation\ of\ ratios\ (EoR) = \frac{1}{N} \sum_{k=1}^N \left( \frac{r_k}{p_k} \right), \quad (1)$$

donde  $N$ ,  $r_k$ , y  $p_k$  denotan el número de ensayos, la recompensa en el ensayo  $k$  y el costo de presionar la palanca en el ensayo  $k$ , respectivamente. Con esta ecuación se calcula el número óptimo de ensayos PR que un ratón debería completar antes de cambiar. Los números estimados son 10, 22, 28, y 58 ensayos para 2xFR6, 2xFR12, 5xFR6, y 5xFR12, respectivamente. Este EoR óptimo ( $EoR^{opt}$ ) se comparó con lo observado ( $EoR^{mice}$ ). La optimalidad de EoR se definió como la razón

$$EoR\ optimality = \frac{EoR^{mice}}{EoR^{opt}} \quad (2)$$

que va de 0 a 1, donde 1 indica que los ratones se comportan de manera óptima.

La media de optimalidad EoR fue de 0.9, 0.83, 0.92, y 0.88 para las condiciones en el orden descrito antes. Esto indica que los ratones son óptimos en todas las condiciones, aunque una tendencia indica que el cambio en el parámetro de costo tuvo un efecto mayor en la desviación de la optimalidad, es decir, dado un beneficio fijo (2x, por ejemplo), la optimalidad era mayor con FR6 que con FR12; pero con un costo fijo, los cambios en el beneficio tuvieron efectos menos notorios. Una explicación es que el incremento en el costo aumentó la cantidad de ensayos incompletos, lo que disminuyó la tasa de obtención de recompensas. Además, los ratones revisitaban el lado PR más para las condiciones FR12 que para las FR6, lo que también contribuyó a la suboptimalidad incrementando indirectamente el número de ensayos incompletos. En resumen, el parámetro de costo fue la principal fuente que llevó a la suboptimalidad.

**Comparing behavioral optimality across experimental regimes.** Al usar EoR es fácil llegar a valores altos de optimalidad por aleatoriedad en distintos regímenes. Por ejemplo, eligiendo con 50 % de probabilidad es más fácil llegar a altos valores de optimalidad en 5xFR12 que en 2xFR6. La optimalidad EoR de elecciones aleatorias es de 0.6, 0.73, 0.72, y 0.92 para cada régimen. Por lo tanto se computó una métrica ajustada que corrigiera para las diferencias en optimalidad logradas por conducta aleatoria. Para ello se computó la distribución de valores de optimalidad EoR logrados por un agente aleatorio en un número total de 1 a 100 ensayos completados. Después, para cada sesión completada por un ratón, se computó la cantidad

$$corrected\ EoR\ optimality = \frac{EoR_{mouse}^{optimality}(n) - EoR_{random}^{optimality}(n)}{1 - EoR_{random}^{optimality}(n)} \quad (3)$$

donde  $EoR_{mouse}^{optimality}$  es la optimalidad EoR de la ecuación 2,  $EoR_{random}^{optimality}$  es la media de la distribución de puntajes equivalentes de optimalidad obtenida de comparar los EoRs de los agentes aleatorios con el EoR óptimo, y  $n$  es el número de ensayos que un ratón realizó en la sesión. Un valor de EoR corregida de 0 indica que el ratón ejecutó tan bien como un agente aleatorio, y un valor de 1 indica que fue óptimo. Valores negativos significan que el ratón se desempeñó peor que el azar, y positivos indican que fue mejor.

La optimalidades EoR corregidas fueron de 0.76, 0.27, 0.71, y -3.89 en el orden usual. En el último caso, el de 5xFR12, la media reportada es sesgada por outliers con menos de 200 ensayos. Si se eliminan, el valor resulta en -0.77.

***Susceptibility to sunk cost fallacy contributes to suboptimality.*** La falacia del costo hundido indica que las decisiones a menudo se basan en los recursos ya gastados en lugar de los beneficios potenciales. En esta tarea, después de pasar del punto de indiferencia los ratones frecuentemente revisitaban el lado PR, que requería una cantidad enorme de respuestas, y permanecían en él hasta recolectar la recompensa.

Para cuantificar la sensibilidad a estos costos se examinaron los ensayos PR recompensados e incompletos. Se parametrizó la proporción de ensayos PR exitosos en función del número de respuestas remanentes y el número de respuestas ya invertidas. Para cada condición, las proporciones de ensayos completados se ajustaron a una función lineal. Si los ratones mostraban susceptibilidad a la falacia del costo hundido, entonces se esperaría una mayor proporción de ensayos completados cuando se invierten más presiones. Esto se reflejaría en que la pendiente de la regresión se aproximaría a cero.

El análisis reveló efectivamente un efecto de costo hundido, aunque no en todas las condiciones. El efecto fue significativo solamente en contextos con un mayor parámetro de costo o de beneficio (es decir, solo estuvo ausente en la condición 2xFR6). Estos hallazgos muestran que los ratones tienen sensibilidad al costo hundido en esta tarea, lo que parece contribuir a su ejecución subóptima.

***Optimal and suboptimal regimes in the parameter space.*** Los resultados indican que la condición 2xFR6 es distinta de las demás. La sensibilidad al costo hundido no tuvo mucha influencia en 2xFR6, donde la opción de recompensa larga es menos atractiva y el costo de la recompensa pequeña es bajo. Tanto la conducta de exploración reducida como la falacia del costo hundido en 2xFR6 son causas probables de la alta optimalidad EoR en esa condición. En otras palabras, en ese régimen los ratones se comportan de forma casi racional.

## DISCUSIÓN

Se investigaron los factores que contribuyen a resultados subóptimos en conducta de toma de decisiones. Se encontró que costo y beneficio contribuyen diferencialmente a la suboptimalidad. Un costo incrementado reduce la tasa de recolección de recompensa y es el factor principal que contribuye a la suboptimalidad, pero el beneficio también contribuyó indirectamente mediante la sensibilidad a los costos hundidos.

La racionalidad de los animales está bajo intenso debate, desde la definición misma de racionalidad. Hay hallazgos contradictorios en la literatura. Parte de la riqueza de este estudio es proveer un paradigma en el que se encuentra conducta tanto racional como irracional en una misma tarea.

Existen algunos problemas con el diseño. Primero, que la comida y el agua como recompensas tienen un comportamiento distinto del de reforzadores que no generan saciedad: a lo largo de la tarea ocurren cambios en los niveles de motivación de los animales. Aun así, se argumenta que el efecto de saciedad del agua en esta tarea es despreciable.

Los puntos de indiferencia estimados en este estudio son solo aproximados del verdadero equilibrio dado que no se probó el efecto de histéresis en el punto de indiferencia. Para dar cuenta de ese efecto, los puntos de indiferencia estimados con el cambio de PR

a FR deberían compararse con el cambio de FR a PR. De acuerdo con otros estudios, la histéresis sí existe en el comportamiento de ratas.

El uso de EoR trajo conclusiones importantes, pero también dificultades: de acuerdo con el EoR, el número de ensayos de PR completados es un factor importante que determina la tasa óptima, pero el orden de las elecciones no importa. En esta tarea, los ratones casi siempre permanecieron en el lado de PR al comienzo y cambiaron a FR cuando el costo de PR incrementó. El modelo EoR no puede explicar eso, pero un modelo simple de aprendizaje por refuerzo en el que la recompensa se reemplaza con una tasa de recompensa por ensayo a las presiones a la palanca explica bien esta conducta. Segundo, fue necesaria la corrección del EoR dado que es fácil llegar a proporciones altas mediante el azar. En tercer lugar, el costo en tiempo y en esfuerzo son distintos y se piensa que se procesan de forma distinta en el cerebro. Este estudio no los puede separar. Cuarto, la optimalidad EoR se estimó mediante una presunción de que los costos y beneficios contribuyen de forma lineal al valor de la elección, lo que no siempre es correcto. La presunción no se sostiene para 5xFR12, por ejemplo. Por último, otros modelos que expliquen mejor esta conducta podrían existir.

Se podría argumentar que el efecto de costo hundido observado se debió más bien a descuento temporal. Sin embargo, según el descuento temporal, el valor de la recompensa incrementa según un ratón se acerca a recibirla mediante la presión de la palanca, de modo que la probabilidad de alcanzarla aumenta en consecuencia. Si se tratase de descuento temporal, ese patrón se encontraría en todas las condiciones probadas, pero no fue así para 2xFR6.