

# Testing the $\Delta$ - $\Sigma$ hypothesis in the suboptimal choice task: Same delta with different probabilities of reinforcement

Valeria V. González, Alejandro Macías, Armando Machado,  
Marco Vasconcelos

2020

El fenómeno de elección subóptima ocurre confiablemente con palomas en condiciones diversas, y amplios esfuerzos se han llevado a cabo para su entendimiento. Sin embargo, aún no se conocen las condiciones necesarias ni suficientes para que ocurra.

En el procedimiento, las palomas escogen consistentemente la alternativa discriminativa a pesar de que ésta entrega 2.5 veces menos comida que la no-discriminativa.

La explicación usual de este fenómeno apunta a la correlación perfecta entre los estímulos de la alternativa discriminativa, y la ausencia de correlación en la no-discriminativa. Es decir, los estímulos del eslabón terminal son diferencialmente informativos, y la elección de las palomas puede señalar una preferencia por la información relacionada con la comida.

De acuerdo con la hipótesis  $\Delta\Sigma$ , el valor  $V$  de una alternativa incrementa con dos variables asociadas con ella: la diferencia en las probabilidades de reforzamiento en los eslabones terminales de la alternativa ( $\Delta$ ) y la probabilidad global de reforzamiento obtenida al elegir la alternativa ( $\Sigma$ ). De forma concreta:

$$V = \Sigma^c \times e^{-\beta\Delta}$$

donde  $\beta$  y  $c$  son dos parámetros del modelo mayores que cero. Las  $\Delta$  de las alternativas 1 (discriminativa) y 2 (no-discriminativa) son:  $\Delta_1 = (p_{1,1} - p_{1,2})$ , y  $\Delta_2 = (p_{2,1} - p_{2,2})$ . Donde  $p_{1,1}$ ,  $p_{1,2}$ ,  $p_{2,1}$ , y  $p_{2,2}$  son las probabilidades de reforzamiento ante los dos estímulos de la alternativa discriminativa y no-discriminativa. Las  $\Sigma$  de cada alternativa son:  $\Sigma_1 = r_1 \times p_{1,1} + (1 - r_1) \times p_{1,2}$ , y  $\Sigma_2 = r_2 \times p_{2,1} + (1 - r_2) \times p_{2,2}$ , donde  $r_1$  y  $r_2$  son las probabilidades de presentación de una de las luces en cada alternativa (la probabilidad de la luz complementaria se representa simplemente como  $1 - r_1$  y  $1 - r_2$ ). Así, los valores de las alternativas 1 y 2 son

$$V_1 = (\Sigma_1)^c \times e^{-\beta\Delta_1}$$

$$V_2 = (\Sigma_2)^c \times e^{-\beta\Delta_2}$$

Para predecir la probabilidad  $P_1$  de elegir la alternativa 1, la hipótesis sugiere una regla simple de proporción:

$$P_1 = \frac{V_1}{V_1 + V_2}$$

que resulta en

$$\begin{aligned} P_1 &= \frac{(\Sigma_1)^c \times e^{-\beta\Delta_1}}{(\Sigma_1)^c \times e^{-\beta\Delta_1} + (\Sigma_2)^c \times e^{-\beta\Delta_2}} \\ &= \frac{1}{1 + \left(\frac{\Sigma}{\Sigma_1}\right)^c e^{-\beta(\Delta_1 - \Delta_2)}} \end{aligned}$$

Si la elección se realiza entre alternativas con distintos  $\Delta$  pero  $\Sigma$  iguales, la ecuación se simplifica a

$$P_1 = \frac{1}{1 + e^{-\beta(\Delta_1 - \Delta_2)}}$$

que tiene un solo parámetro. De forma similar, si la elección es entre dos alternativas con distinto  $\Delta$  pero  $\Sigma$  igual, entonces la ecuación se simplifica a

$$P_1 = \frac{1}{1 + \left(\frac{\Sigma}{\Sigma_1}\right)^c}$$

con un solo parámetro, también.

González (2020) mostraron que los animales son sensibles a ambas variables. En su experimento 1, la  $\Delta$  fue variada entre condiciones manteniendo a  $\Sigma$  igual. Consistente con la hipótesis, los animales prefirieron la alternativa con mayor  $\Delta$ . En su experimento 2 variaron  $\Sigma$  y mantuvieron a  $\Delta$  constante, y los animales prefirieron la alternativa con mayor  $\Sigma$ .

La hipótesis  $\Delta\Sigma$  predice que la preferencia dependerá de la diferencia entre las probabilidades asociadas a los eslabones terminales, no de las probabilidades absolutas. Así, alternativas con eslabones terminales con probabilidades de .75/.25 y .9/.4 deberían resultar en la misma preferencia, dado que ambas tienen una  $\Delta$  de .5. Al presentarlas juntas, los animales deberían ser indiferentes, y al comparar a ambas con la misma alternativa, deberían resultar en el mismo grado de preferencia.

Este experimento busca probar la hipótesis  $\Delta\Sigma$  de este modo, probando alternativas con probabilidades absolutas distintas pero  $\Delta$  constante de .5, y comparándolas con una alternativa con una  $\Delta$  de 0. Además, las  $\Sigma$  se mantuvieron iguales (o casi). Dado que  $\Sigma_1 \approx \Sigma_2$ , la hipótesis  $\Delta\Sigma$  predice preferencia por la  $\Delta$  mayor, pero el mismo grado de preferencia sin importar las probabilidades específicas.

Ningún otro modelo de elección subóptima predice el mismo resultado cuando  $\Delta$  permanece igual. Otros asumen que las palomas ignoran las malas noticias. De acuerdo con ellos, cuando las respuestas ante uno de los dos estímulos nunca son reforzadas, los animales se comportan como si esa alternativa estuviese compuesta por solo un eslabón terminal siempre reforzado. De ahí viene que cuando se compara una  $\Delta = .5$  obtenida con el par .5/0, con una  $\Delta = 0$  obtenida con el par .5/.5, los modelos predicen indiferencia.

Un segundo propósito del experimento es contrastar el modelo  $\Delta\Sigma$  contra otros modelos de elección subóptima.

En el experimento 1, se obtuvo una  $\Delta$  de .5 usando las probabilidades extremas de 1 y 0 de la tarea estándar. Específicamente, se usaron los pares 1/.5, .75/.25, y .5/0. En el experimento 2 se obtuvo la misma  $\Delta$  de .5 con los pares de probabilidades .9/.4 y .6/.1.

### EXPERIMENTO 1

Comenzó con preentrenamiento en RF 1, 5 y 10. El procedimiento fue similar al de elección subóptima con las posiciones variadas aleatoriamente. Las sesiones consistían en 120 ensayos: 40 de elección y 80 forzados. Ejemplo:

En la condición .75/.25 ( $\Delta = .5$ ) vs .5/.5 ( $\Delta = 0$ ), de los 40 ensayos forzados a la alternativa 1, 20 fueron seguidos del eslabón terminal  $TL_{1,1}$  y 20 del eslabón  $TL_{1,2}$ . 15 de los ensayos  $TL_{1,1}$  terminaron en reforzamiento (el 75 %), y 5 de los ensayos  $TL_{1,2}$  terminaron en reforzamiento (el 25 %). De los 40 ensayos forzados a la alternativa 2, 20 fueron seguidos por  $TL_{2,1}$  y 20 por  $TL_{2,2}$ . En ambos casos, 10 de esos ensayos fueron reforzados (50 %). En resumen, no se trata exactamente de elección subóptima. Una alternativa tiene estímulos más informativos (.75/.25) que la otra (.5/.5), pero ninguna da certeza. En ambos casos la probabilidad de ocurrencia de cualquiera de los dos estímulos es de  $r = .5$ , la diferencia está en que los estímulos de la alternativa .75/.25 eran un poco mejores como predictores.

En la condición 2, las probabilidades  $r_1$  y  $r_2$  fueron ajustadas a .9 para mantener el valor de  $\Sigma$  igual entre las alternativas.

La  $\Delta = .5$  fue asignada a alternativas distintas en condiciones consecutivas. La tasa global de reforzamiento  $\Sigma$  fue siempre de .5 para la alternativa con  $\Delta = 0$ , pero varió ligeramente para la alternativa con  $\Delta = .5$  entre .45 y .55.

### RESULTADOS Y DISCUSIÓN

Con los pares de 1/.5 y .75/.25, la mayoría de las palomas mostró preferencia por la alternativa con  $\Delta = .5$ . En el par .5/0, la preferencia estuvo levemente debajo de .5. Hubo un efecto significativo de condición, y diferencias significativas entre el par .5/0 y los dos restantes. Una prueba T mostró que la preferencia por el par .5/0 se encontraba significativamente debajo de la indiferencia. Las latencias mostraron un patrón similar: latencias más cortas en la alternativa con  $\Delta = .5$  en las condiciones 1 y 2, y latencias más largas en la condición .5/0.

El experimento 1 mostró que el valor de una alternativa parece cambiar con las probabilidades específicas de reforzamiento en los eslabones terminales, lo que es inconsistente con la hipótesis  $\Delta\Sigma$ . Se sugiere que quizá la ausencia de reforzamiento (en el par .5/0) podría ser un caso especial (nota de Daniel: si el modelo necesita incluir casos especiales, quizá no sea un gran modelo).

La diferencia  $\Delta$  entre probabilidades de reforzamiento podría no ser el factor determinante para la elección. Alternativamente, el efecto podría cambiar cuando se incluye un eslabón terminal sin reforzamiento. Esa idea aparece en otros modelos de elección subóptima (los animales ignoran los estímulos que anuncian malas noticias). Sin embargo, esos modelos sí predicen indiferencia, mientras que estos datos sugieren una leve preferencia por la alternativa  $\Delta = 0$ .

Algunos investigadores han sugerido que los casos de  $p = 1$  y  $p = 0$  podrían ser especiales evolutivamente: los animales son atraídos por el valor extremadamente alto, y

repelidos por el extremadamente bajo. Sin embargo, estos resultados sugieren que  $p = 1$  podría no ser un caso especial, sino solamente  $p = 0$ . Se sugiere que el valor de incertidumbre de  $p = .5$  en el par 1/.5 podría restarle valor a la alternativa porque a los organismos no les gusta la incertidumbre.

## EXPERIMENTO 2

Parece que el experimento fue pensado después para evaluar valores cercanos pero distintos a los extremos de 0 y 1. Los pares fueron .9/.4 y .6/.1. La hipótesis  $\Delta\Sigma$  predice una preferencia comparable por la alternativa con  $\Delta = .5$ .

Los organismos pasaron por tres condiciones: la primera fue la tarea estándar de elección subóptima con probabilidades .2 vs .5 en la que una alternativa ofrece predictores perfectos y la otra no ofrece información. Después de adquirir su preferencia, pasaron a las condiciones de .9/.4 vs .5/.5, y .6/.1 vs .5/.5. Sin embargo, por la manera en que decidieron la cantidad de ensayos reforzados y no reforzados, hubo mayor variabilidad en los reforzadores obtenidos al final de las sesiones.

## RESULTADOS Y DISCUSIÓN

En el par .9/.4, cuatro de cinco sujetos mostraron preferencia por la alternativa con mayor  $\Delta$ . En el par .6/.1, la preferencia de todas las palomas estuvo cerca o debajo de la indiferencia.

En cuanto a las latencias, estas fueron menores para la alternativa con  $\Delta = .5$  en el par .9/.4, pero lo contrario sucedió con el par .6/.1. Las diferencias no fueron significativas.

Los resultados estuvieron en línea con los del experimento 1 tanto en preferencia como en latencias.

Para determinar si  $p = 1$  y  $p = 0$  son casos especiales, se realizaron pruebas T pareadas comparando los pares 1/.5 vs .9/.4, y .5/0 vs .6/.1. No se encontraron diferencias significativas.

## DISCUSIÓN GENERAL

El modelo  $\Delta\Sigma$  sugiere que la preferencia por una alternativa incrementa con  $\Delta$ , sin importar cómo se obtenga su valor. Si dos distintos pares de probabilidades resultan en la misma  $\Delta$  deberían ser indistintos en términos de preferencia. Sin embargo, los datos se oponen a esta perspectiva: los pares con valores “extremos” de 1, .9, 0, y .1, resultaron en indiferencia o leve preferencia en contra de la alternativa con  $\Delta = .5$ .

Los resultados fueron consistentes con la idea de que las latencias son una buena medida del valor de las alternativas.

Los datos, dicen los autores, apoyan “parcialmente” el modelo  $\Delta\Sigma$ . Tres de sus cinco condiciones fueron consistentes, mostrando preferencia de magnitud comparable.

El papel de  $\Delta$  es similar al contraste intra-ensayos propuesto por Zentall (2005). Según esta idea, el estado hedónico de un sujeto cambia entre el final de un evento poco apetitivo y un estímulo que señala recompensa o la recompensa misma. En el procedimiento estándar de elección subóptima, la alternativa preferida tiene  $\Delta = 1$ , la máxima diferencia posible; y la menos preferida tiene  $\Delta = 0$ , la mínima diferencia posible. Cuando se escoge la alternativa preferida, con un 80 % de probabilidades de no obtener reforzamiento, se entra en un contexto local negativo en el que el contraste con las raras ocasiones en que se obtiene el reforzador las exalta.

Se propone que algo similar podría estar en efecto en esta situación: en el caso de la condición 1/.5, los sujetos experimentaban la mayoría de las veces el estímulo que anunciaba una probabilidad moderada de reforzamiento, lo que pudo establecer un contexto en el que las ocasiones en que se presentaba el estímulo que señala probabilidad de 1 reciben un contraste positivo que alza el valor de la alternativa completa. En el caso del par .5/0 sucedería lo contrario: en la mayoría de las ocasiones los sujetos experimentaban el estímulo con probabilidad de .5, que podría establecer un contexto en el que el estímulo con probabilidad de 0 recibía un contraste negativo.

Aunque esta explicación se ajusta bien a los pares de 1/.5, .5/0, .9/.4, y .6/.1, se encuentra con problemas para explicar el caso de .75/.25, dado que ambos eslabones terminales se presentaron con igual frecuencia, y no hay razón teórica para elegir al estímulo más pobre como el que conforma el contexto.

El modelo de información temporal de Cunningham y Shahan (2018) asume que existe una competencia entre la información temporal dada por los estímulos y las tasas relativas de reforzamiento de cada alternativa, por lo que enfatiza las mismas variables que la hipótesis  $\Delta\Sigma$ . Sin embargo, el modelo no especifica cómo cuantificar la información temporal dada por las señales asociadas con diferentes probabilidades de reforzamiento, sino que trata únicamente con las probabilidades extremas de 0 y 1 de la alternativa subóptima, y la probabilidad de .5 de la alternativa óptima. Así, todas las condiciones de este estudio salvo la de .5/0 están fuera del alcance del modelo, que predice en ese caso indiferencia, mientras que los datos revelaron una leve preferencia por la alternativa de  $\Delta = 0$ .

El modelo de Reinforcement Rate (RRM) se basa en el forrajeo óptimo, y sugiere que los animales siguen una estrategia de búsqueda de información. Los animales atienden al estímulo rojo porque provee información del reforzamiento, e ignoran el verde porque en el contexto natural, encontrarlo dispararía una nueva búsqueda. Sin embargo, ese modelo tampoco está hecho para probabilidades distintas de 1 y 0. Una versión nueva del modelo hecha para lidiar con tales situaciones predice preferencia por la alternativa con  $\Delta$  mayor para todas las condiciones salvo para .5/0, para la que predice indiferencia, es decir, que también se equivoca.

El modelo hiperbólico de Mazur es muy similar, así que falla en las mismas condiciones que RRM.

El modelo de Zentall de contraste no dice explícitamente qué sucede en situaciones donde el contraste es negativo. Si se asume que solo el contraste negativo afecta la elección, el modelo de Zentall tiene las mismas dificultades que los anteriores.

En resumen, todos los modelos, incluido  $\Delta\Sigma$ , tienen dificultades para explicar la conducta de elección subóptima.