

THE $\Delta\Sigma$ HYPOTHESIS: HOW CONTRAST AND REINFORCEMENT RATE COMBINE TO GENERATE SUBOPTIMAL CHOICE

VALERIA V. GONZÁLEZ, ALEJANDRO MACÍAS,
ARMANDO MACHADO, MARCO VASCONCELOS

2020

Introducen una nueva terminología para subcho: IL_1 y IL_2 para los eslabones iniciales; $TL_{1,1}$, $TL_{1,2}$, $TL_{2,1}$ y $TL_{2,2}$ para los eslabones terminales; r_1 , $1 - r_1$, r_2 y $1 - r_2$ para las probabilidades de entrar a cada eslabón terminal; d para la demora de los eslabones terminales; y $p_{1,1}$, $p_{1,2}$, $p_{2,1}$ y $p_{2,2}$ para las probabilidades de reforzamiento de cada eslabón terminal. La probabilidad global de reforzamiento Σ asociada a cada alternativa es el promedio ponderado de las probabilidades de reforzamiento de cada eslabón terminal: $\Sigma_1 = r_1 \times p_{1,1} + (1 - r_1) \times p_{1,2}$ y $\Sigma_2 = r_2 \times p_{2,1} + (1 - r_2) \times p_{2,2}$. Los pesos son las probabilidades de entrar en cada eslabón terminal. En la tarea estándar de subcho las probabilidades globales de reforzamiento son de $\Sigma_1 = .2$ y $\Sigma_2 = .5$.

Han surgido algunas explicaciones:

- El contraste de Zentall: el paso de la probabilidad esperada de reforzamiento de 0.2 en la alternativa discriminativa a 1 cuando se presenta el estímulo positivo incrementa el valor de la alternativa completa.
- Modelo de descuento hiperbólico de Mazur. Si un estímulo anuncia una demora nunca seguida por recompensa, este nunca aumenta en valor y es ignorado, por lo que funcionalmente no existe. Así, el valor de IL_1 equivale al valor de $TL_{1,1}$, y no el promedio de $TL_{1,1}$ y $TL_{1,2}$. En cambio, el valor de IL_2 sí es afectado por los ensayos negativos, lo que resulta en preferencia por la alternativa discriminativa.
- Reinforcement Rate Model (RRM): Con predicciones similares a las del modelo de Mazur, presupone que los animales valoran la información que los estímulos dan sobre la disponibilidad de comida. $TL_{1,1}$ proporciona información, pero $TL_{1,2}$, no. En ambientes naturales, indica a los organismos que deben iniciar una nueva búsqueda, por lo que los animales nunca aprenden a atenderlo y efectivamente lo ignoran.
- Modelo de información temporal de Cunningham y Shahan: de acuerdo con esta explicación, los animales escogen la alternativa subóptima debido a que proporciona

información sobre cuándo y dónde se entregará comida. Esta propuesta también parte del supuesto de que un estímulo que no señala comida no tiene ningún impacto en la elección.

Las características comunes de estas explicaciones son (1) suponen que los animales ignoran el estímulo que nunca va seguido por comida, (2) como consecuencia de ello, asumen que $TL_{1,1}$ tiene un impacto desigual en la elección, y (3) lidian bien con situaciones en las cuales los estímulos de la alternativa discriminativa correlacionan perfectamente con la resencia de comida, pero no con situaciones en las cuales la probabilidad de entrega es distinta de 0 o 1.

La hipótesis $\Delta\Sigma$ propuesta intenta abarcar esas situaciones en las cuales las probabilidades de reforzamiento de la alternativa discriminativa son distintas de 0 o 1, y condensa las seis probabilidades de la tarea (dos r y cuatro p) en dos.

Δ se refiere a la diferencia entre las probabilidades de reforzamiento asociadas con los eslabones terminales de una alternativa: $\Delta_1 = p_{1,1} - p_{1,2}$ y $\Delta_2 = p_{2,1} - p_{2,2}$. En la tarea estándar, $\Delta_1 = 1$ y $\Delta_2 = 0$. La hipótesis asume que, si todo lo demás es igual, los organismos mostrarán preferencia por la alternativa con mayor Δ . Σ se refiere a la probabilidad global de reforzamiento de una alternativa. De esto se sigue que el valor total de la alternativa será la conjunción de Δ y Σ . Si V denota el valor de la alternativa, y f una función de Δ y Σ , entonces $V = f(\Delta, \Sigma)$, con derivadas parciales positivas, $\delta f/\delta\Delta > 0$ y $\delta f/\delta\Sigma > 0$.

Ahora, hace falta una función que relacione al valor con la preferencia. Se asume que la preferencia por la alternativa 1, P_1 , es igual a $g(V_1, V_2)$, para una función g que está siempre entre 0 y 1, incrementa con la tasa V_1/V_2 , y, sin contar un sesgo, es igual a .5 cuando $V_1 = V_2$. Unir ambas funciones resulta en

$$P_1 = g(V_1, V_2) = g(f(\Delta_1, \Sigma_1), f(\Delta_2, \Sigma_2))$$

En este estudio probaron la hipótesis $\Delta\Sigma$ en dos experimentos: en uno variaron las probabilidades de reforzamiento en los eslabones terminales (p). Mantuvieron todo constante salvo por los valores de Δ . En el otro, mantuvieron todo constante salvo por los valores de Σ .

Experimento 1

En el experimento 1 las probabilidades r de todos los eslabones terminales fueron iguales a .5, pero sus probabilidades p de reforzamiento fueron variadas para resultar en valores de $\Delta_{\text{diff}} = \Delta_1 - \Delta_2 = -0.5, 0$, o $+0.5$, cada uno de los cuales ocurrió en dos condiciones, una por cada asignación posible de Δ_1 y Δ_2 a las alternativas 1 y 2. En todas las condiciones, la probabilidad global de reforzamiento, Σ_1 y Σ_2 , fue igual a .5.

Según la hipótesis, las palomas deberían preferir la alternativa con mayor Δ , y debería haber el mismo grado de preferencia cuando el valor de Δ_{diff} es igual, sin importar los valores particulares de los que se origina.

La preferencia por la alternativa 1 incrementó con Δ_{diff} .

Analizaron las latencias de respuesta como una medida indirecta de preferencia, esperando que las latencias menores se presentaran asociadas a las Δ mayores. En todos los casos, las latencias fueron menores para la alternativa preferida, con mayor Δ . Sin

embargo, quedaba por ver si la latencia dependía de la Δ de la alternativa, o de la diferencia entre las dos Δ . Un análisis de regresión múltiple reveló que la latencia no estaba modulada por el efecto del contexto, sino por la Δ de la alternativa escogida.

Las preferencias fueron similares entre condiciones, lo que sugiere que dependieron del valor de Δ_{diff} , y no de los valores separados de las Δ de cada alternativa. Sin embargo, las latencias sí sugieren un efecto de las Δ individuales: los sujetos mostraron latencias más grandes para $\Delta = 0$ cuando estaba pareada con $\Delta = .5$ que para $\Delta = .5$ pareada con $\Delta = 1$, es decir, las latencias eran distintas aunque los valores de Δ_{diff} eran iguales. **Parece ser que en el procedimiento de elección subóptima, la latencia y la preferencia pueden depender de variables distintas.**

Los resultados del experimento 1 son consistentes con la hipótesis $\Delta\Sigma$. Los resultados indican preferencia por la Δ más alta, y también que la sensibilidad a las probabilidades de reforzamiento no se limita a los casos extremos de $\Delta = 1$ y $\Delta = 0$.

Experimento 2

Examina el efecto de la segunda variable de la hipótesis: Σ . Para ello, se fijaron los valores de $\Delta_1 = \Delta_2 = 1$, pero se variaron los valores de Σ correlacionalmente, es decir, al aumentar Σ_1 , disminuyó Σ_2 de modo que $\Sigma_1 + \Sigma_2 = 1$. Según la hipótesis, las palomas deberían preferir la alternativa con mayor Σ .

La forma general del experimento fue la misma. Los animales escogieron entre dos alternativas con TL de 10s. Las probabilidades de reforzamiento de los TL de la alternativa 1 fueron $p_{1,1} = 1$ y $p_{1,2} = 0$. Las mismas probabilidades se usaron en la alternativa 2, con lo que $\Delta_1 = \Delta_2 = 1$, y $\Delta_{\text{diff}} = 0$. Se manipularon las probabilidades $\Sigma_1 = r_1$ y $\Sigma_2 = r_2$.

Al mantener a $\Sigma_1 + \Sigma_2 = r_1 + r_2 = 1$, la probabilidad de ocurrencia de $\text{TL}_{1,1}(r_1)$ fue igual a la probabilidad de ocurrencia de $\text{TL}_{2,2}(1 - r_2)$. De igual modo, la probabilidad de ocurrencia de $\text{TL}_{1,2}(1 - r_1)$ fue igual a la probabilidad de ocurrencia de $\text{TL}_{2,1}(r_2)$. Entre las cinco condiciones, r_1 fue igual a 0.1, 0.3, 0.5, 0.7, y 0.9.

Se analizaron las latencias de la misma manera que en el experimento anterior, esperando que se relacionaran negativamente con los valores de Σ . Los resultados confirmaron esta predicción: la alternativa con el mayor valor de Σ tuvo menores latencias y más elecciones. La diferencia entre las latencias entre la alternativa preferida y la no preferida varió conforme a la diferencia entre Σ_1 y Σ_2 .

Discusión general

Previamente se ha evaluado el efecto de otras variables en el procedimiento de elección subóptima, como las demoras de los eslabones terminales, el grado de contigüidad entre la elección y el encendido del estímulo del eslabón terminal, y el requisito de respuesta del eslabón inicial. Todos los modelos propuestos necesitan de la predictabilidad de las señales de la alternativa subóptima y la diferencia en la tasa general de reforzamiento.

La hipótesis $\Delta\Sigma$ muestra un buen ajuste con los datos. Captura adecuadamente las variaciones en preferencia producidas por manipulaciones en el contraste intra alternativa (Δ) y en la probabilidad global de reforzamiento (Σ).

Otros modelos tienen problemas con los resultados de estos experimentos. El modelo de Cunningham y Shahan, por ejemplo, predice los resultados del segundo experimento, pero no puede aplicarse al primero dado que no hay una correlación perfecta entre los estímulos de los eslabones terminales y la consecuencia final.

La versión más reciente del Reinforcement Rate Model tiene las dificultades contrarias. Puede lidiar con correlaciones imperfectas, pero no puede lidiar con los resultados del segundo experimento. En él, los dos eslabones terminales de cada alternativa correlacionan perfectamente con comida y su ausencia, y dado que el modelo propone que los animales no atienden a S^- , los animales deberían funcionalmente escoger alternativas que siempre terminan en comida, de lo que se sigue la predicción de indiferencia en todas las condiciones del experimento 2, que están en desacuerdo con los resultados observados. El modelo hiperbólico de Mazur es matemáticamente muy similar a RRM, y por lo tanto tiene los mismos éxitos y fracasos..

La explicación del contraste de Zentall hace predicciones acertadas en el primer experimento, pero no en el segundo, pues en ese caso el contraste positivo y el valor de Σ están inversamente relacionados, y una preferencia por el contraste mayor implicaría preferencia por la Σ menor.

Para farolear, los autores hicieron predicciones de 28 experimentos diferentes usando la hipótesis $\Delta\Sigma$. Calcularon Δ_{diff} y Σ_2/Σ_1 , asumiendo que $\Delta = 0$ cuando hay un solo eslabón terminal. El modelo $\Delta\Sigma$ captura bien las preferencias observadas, y predice la proporción de preferencia observada exacta para cada estudio.

La conclusión obvia de estos estudios es que la exploración paramétrica está muy limitada. Típicamente se ha usado $\Delta_1 = 1$ y $\Delta_2 = 0$, lo que limita no solo las conclusiones sobre el efecto de Δ_{diff} , sino que también ensombrece las interpretaciones del efecto de Σ . Esto podría ser porque, con 6 diferentes estímulos, la exploración paramétrica es complicada. La hipótesis $\Delta\Sigma$ simplifica la nomenclatura y reduce todo a dos variables. Por ahora, todo indica que, como se predijo, el valor de una alternativa aumenta con Δ y con Σ .