

La integración de experiencias pasadas y recientes determina la elección en humanos

LUIS J. LÓPEZ-ROMERO, ANGÉLICA ALVARADO, CHRISTOPHER TAMAYO & JAVIER VILA

FES Iztacala Universidad Nacional Autónoma de México

Resumen

Estudios recientes han demostrado que cuando los animales buscan alimento, eligen entre diferentes parcelas después del paso del tiempo, de acuerdo con un promedio dinámico (Mazur, 1996). Este promedio dinámico ha servido para predecir la actuación basándose en la distancia temporal y al valor relativo de cada una de las experiencias de aprendizaje. Para comprobar estas ideas en humanos, se utilizó una tarea en donde los participantes elegían entre dos opciones similares en densidad de reforzamiento a lo largo de cinco ensayos (0.5 vs. 0.5), que cambiaban en el último o primer bloque de ensayos de entrenamiento (0.9 vs. 0.1) y que, además, se diferenciaban en el intervalo entrenamiento-prueba (0 h vs. 24 h). Los resultados mostraron que cuando el bloque de ensayos con mayor densidad de reforzamiento fue el más reciente, se observó preferencia por la opción con mayor densidad de reforzamiento en una prueba a las 0 h. En cambio, cuando el ensayo con mayor densidad de reforzamiento fue el primero, se observó indiferencia entre ambas opciones en la prueba a las 0 h y una preferencia por la opción con mayor densidad a las 24 h. Los resultados son coherentes con las predicciones de la Regla de Peso Temporal y suponen un promedio dinámico de las experiencias pasadas en la elección de los humanos.

Palabras clave: recuperación espontánea, regla de peso temporal, memoria, humanos.

Integration of past and recent experiences determines choices in human beings.

Abstract

Recent studies had showed that animals choose between patches base on time using a dynamic average of experiences (Mazur, 1996). This average has been used to predict performance based on the temporal distance and the relative value of learning experiences. To test these ideas in humans we used a task in which participants could choose between two options with similar reinforcement densities over 5 trials (0.5 vs. 0.5) but different in the last or first trials block (0.9 vs. 0.1) And also they differed in training-test interval (0h vs. 24h). When reinforcement density in last training block was higher, a test at 0h showed preference for the option with higher density. Whereas when higher density occurred during first trial block results showed indifference for both response options, but in a test at 24 hours they showed preference for the first higher option. These results are in agreement with Temporal Weighting Rule predictions and provide evidence for a dynamic average of past experiences in human choice.

Key words: Spontaneous recovery, temporal weighting rule, memory, humans.

Dirigir toda correspondencia sobre este artículo a: Javier Vila Carranza. División de Investigación y Posgrado. Facultad de Estudios Superiores, Iztacala. Universidad Nacional Autónoma de México. Tlalnepantla, Edo Mex. México. Teléfono: +55 56 23 12 96, ext. 133.
Correo electrónico: javila@campus.iztacala.unam.mx.
RMIP 113-120.
ISSN-impres: 2007-0926
www.revistamexicanadeinvestigacionenpsicologia.com
Derechos reservados ©RMIP

INTRODUCCIÓN

Recientemente se ha estudiado cómo las experiencias pasadas se almacenan en la memoria de los animales no humanos y humanos, para

observar después cómo estas se combinan e influyen en la conducta presente. Por ejemplo, los cambios en la elección entre diferentes opciones de respuesta que se presentan de manera alterna se pueden explicar si asumimos que un animal almacena la información de cada experiencia de manera separada y luego realiza un tipo de promedio entre todas ellas al inicio de una experiencia nueva, es decir, su elección se basa en la integración de toda la información aprendida previamente. (Devenport, 1998; Mazur, 1995).

Considerando esta la idea de evaluar la posible integración de la memoria de cada experiencia en la elección de opciones previamente reforzadas, Mazur (1996) registró las respuestas de palomas en un programa concurrente de reforzamiento en el que cambiaba el porcentaje de reforzamiento relativo a cada uno de los componentes del programa de manera sucesiva. En su estudio, las palomas respondían inicialmente más a la opción de respuesta que proporcionaba mayor reforzamiento y de manera similar a las dos opciones cuando estas fueron igualadas en su refuerzo relativo en fases posteriores. Sin embargo, después de un intervalo de 24 h los sujetos presentaron una preferencia temporal por la opción elegida inicialmente, lo que sugiere una recuperación espontánea de la respuesta de elección por la tecla "más rica" al inicio de la sesión de recuperación. Estos resultados se han interpretado como una integración de la información aprendida durante todas las sesiones y son consistentes con la idea de que la conducta de elección en el inicio de una nueva sesión se basa en una media ponderada de los eventos de las sesiones pasadas (Mazur, 1996; Devenport, 1998; Vila, López, & Alvarado, 2010).

Una explicación de estos resultados ha sugerido que los organismos utilizan un promedio dinámico definido entre la recencia de cada experiencia y el valor subjetivo de cada una de las experiencias de aprendizaje, que puede

ser expresada cuantitativamente en la Regla de Peso Temporal (RPT). De acuerdo con esta regla, las experiencias más recientes obtienen un mayor peso inmediatamente después de que ha ocurrido la experiencia, pero esta valoración cambiará con el paso del tiempo y tenderá a igualar la distancia temporal relativa de todas las experiencias pasadas, dando entonces un mayor peso al valor subjetivo de cada experiencia, lo que determinará que un organismo elegirá con base a la recencia de la experiencia si la prueba de elección es inmediata, o con base en su valor subjetivo si esta se realiza de manera demorada. La regla del peso temporal ha sido demostrada previamente en estudios realizados en animales no humanos como ratas (Devenport, 1998), ardillas (Devenport & Devenport, 1994), perros (Devenport & Devenport, 1993) y, recientemente, en humanos (López-Romero, García-Barraza, & Vila, 2010; Vila et al., 2010).

La ecuación 1 reproduce la propuesta de Devenport (1998) para la estimación temporal de una experiencia dada, en donde V_w corresponde al valor de

$$(1) V_w = \frac{\sum_{n=1}^{n=j} (Q * 1/T)}{\sum_{n=1}^{n=j} 1/T}$$

de estimación temporal de una experiencia, Q a su valor subjetivo estimado y $1/T$ al valor de recencia. Así, para calcular el valor de estimación temporal de la segunda experiencia B, se debe calcular $V_w B$, considerando la estimación temporal total de ambas experiencias, tal y como lo muestra la ecuación 2.

$$(2) V_w B = V_w B / V_w A + V_w B$$

Así, cuando los valores subjetivos de las experiencias A y B son iguales ($Q_A = Q_B$), después de un intervalo de tiempo se preferirá inicialmente la experiencia más reciente B, pero al paso del tiempo se tenderá a la indiferencia

Tabla 1. Diseño experimental donde el nombre de los grupos corresponde al momento en el que se presentó la densidad de reforzamiento mayor para una de las opciones de respuesta (densidad reciente: grupos DR y densidad distante: grupos DD) y el intervalo entrenamiento–prueba (0 h o 24 h).

GRUPO	BLOQUE 1	BLOQUE 2	BLOQUE 3	BLOQUE 4	INTERVALO	PRUEBA
DR 0	A, .5	A, .5	A, .5	A, .9	0 h	A vs. B
DR 24	B, .5	B, .5	B, .5	B, .1	24 h	
DD 0	A, .9	A, .5	A, .5	A, .5	0 h	
DD 24	B, .1	B, .5	B, .5	B, .5	24 h	

Nota: en negritas se marcan los bloques de ensayos donde la densidad de reforzamiento cambió.

entre ambas experiencias, debido a que ambas serán iguales en su distancia temporal relativa y en su valor subjetivo, tal y como ocurre en aquellos estudios en los que el valor subjetivo de ambas experiencias es similar (Alvarado, Jara, Vila, & Rosas, 2006; López et al., 2010). Por otro lado, cuando el valor subjetivo de la experiencia A es mayor que el de B ($Q_A > Q_B$), al pasar el tiempo, cuando disminuya la recencia de ambas experiencias y la distancia temporal relativa de A y B se iguale, se tenderá a elegir A por ser la experiencia con una valor subjetivo mayor, tal y como se ha observado cuando el valor de la experiencia A es mayor (Vila et al., 2010).

De este modo, según la RPT (Devenport & Devenport, 1994), durante la última elección es cuando se ponderan los pesos de la distancia relativa de cada alternativa, así como su valor subjetivo, teniendo un mayor peso inicial las experiencias recientes, pero al pasar el tiempo y perderse la recencia, será el valor subjetivo de cada alternativa el que tendrá entonces un mayor peso y, por ende, determinará la elección observada. La RPT supone así la integración de las experiencias pasadas con base en dos variables: su recencia o distancia temporal relativa ($1/T$) y su valor subjetivo (Q), que en la ecuación 1 es representada como ($Q * 1/T$).

La presente investigación trató de evaluar si en participantes humanos es posible observar una integración de la información aprendida previamente en una situación espacial con dos opciones, que afecte la conducta de elección

posterior a la exposición de cuatro experiencias reforzadas sucesivas, como una réplica sistemática del procedimiento empleado por Mazur (1996) con pichones. Lo anterior permitirá estudiar la generalidad de las ideas de integración de información planteadas originalmente en la RPT, pero utilizando un procedimiento diferente del empleado hasta ahora en el estudio de las predicciones de la RPT, en la que se han empleado situaciones de interferencia de dos fases en las cuales el aprendizaje de la segunda fase interfiere en lo aprendido inicialmente tal y como sucede en un procedimiento de extinción (ver Bouton, 1993).

En el presente estudio se utilizó una tarea similar a la empleada por Vila y cols. (2010), en la cual los participantes aprenden de manera alternada a responder a dos opciones de respuesta A y B reforzadas similarmente. Sin embargo, en este estudio los valores de reforzamiento para cada opción durante el entrenamiento cambiaron de densidad en una de cuatro fases (la primera o la última) y posteriormente se presentó una prueba de elección a las 0 h o a las 24 h. En la Tabla 1 se presenta el diseño experimental empleado. De acuerdo con las predicciones de la RPT, las cuatro experiencias deberán ser integradas de acuerdo con la distancia temporal de cada experiencia y su densidad de reforzamiento obtenida en cada opción considerada como su valor subjetivo. Así, la preferencia de A debe ser mayor a las 0 h cuando su densidad inmediata sea mayor (grupo DR 0) y debe ser mayor en las pruebas

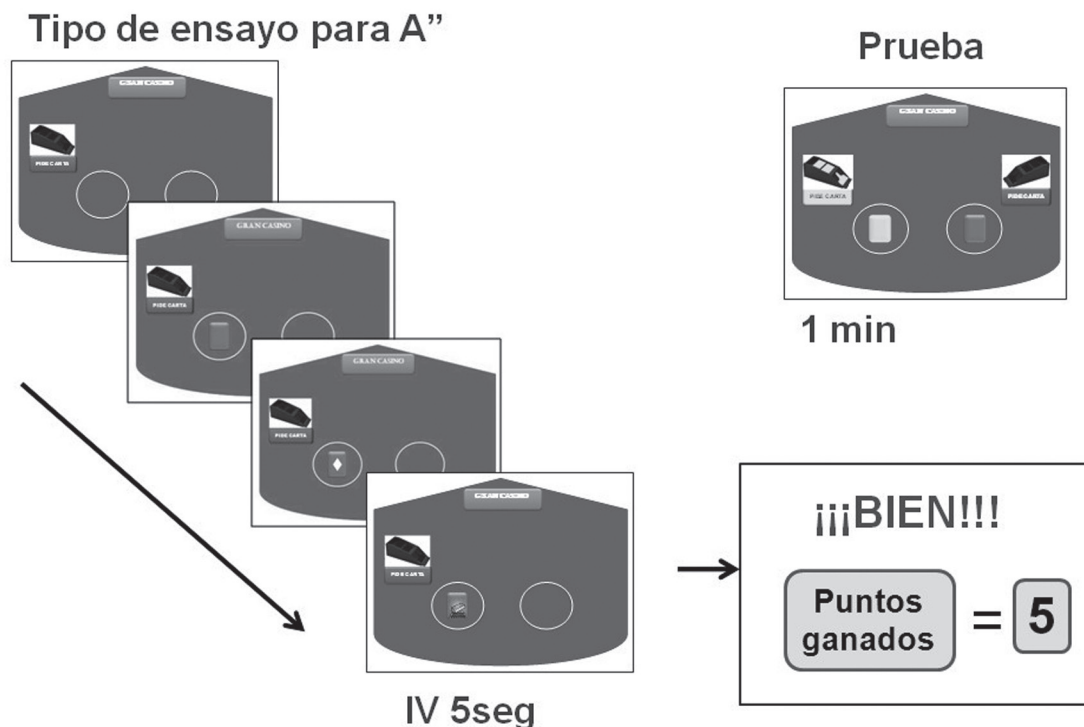


Figura 1. Se muestra en la parte izquierda la secuencia de un ensayo reforzado típico para la máquina dispensadora A, así como en parte derecha de la pantalla de prueba para todos los grupos que tuvo una duración de 1m.

de 24 h cuando el promedio de la densidad de reforzamiento para A en todos los bloques de entrenamiento sea mayor (grupo DD 24).

MÉTODO

Participantes

Se pidió la colaboración voluntaria e informada de cuarenta estudiantes universitarios de entre 19 y 21 años de edad (27 mujeres y 13 hombres), alumnos del Instituto Politécnico Nacional y de la Facultad de Estudios Superiores Iztacala de la UNAM de la Cd. de México, con buena salud y sin consumo de medicamentos al inicio del estudio. A todos los participantes se les informó de las condiciones éticas del estudio y se les otorgaron créditos en una de las asignaturas que cursan por su participación.

Aparatos y situación experimental

El experimento se llevó a cabo en un salón de

clases del Instituto Politécnico Nacional de 5 x 8 m. aproximadamente, que contaba con sillas-bancos, una mesa e iluminación artificial; además, se utilizó un cubículo de la FES Iztacala de 1.5 x 1.5 m., que contaba con iluminación artificial y aislado de ruidos externos. Se utilizó un ordenador portátil Gateway que tenía instalado el programa Super Lab Pro for Windows v 4.0.2 (Cedrus Co.) en donde todos los participantes resolvieron la tarea experimental individualmente y que, además, se utilizó para programar la tarea, presentar los estímulos y registrar las respuestas.

Tarea experimental

La tarea consistió en una situación donde los participantes tenían que escoger entre dos opciones de respuesta que se presentaban de manera alterna, y que se diferenciaban por su localización espacial y la proporción de ensayos reforzados-

no reforzados, así como por su valor de recencia. La pantalla de la computadora portátil simulaba ser una superficie de una mesa de juego donde se encontraban dos máquinas dispensadoras de cartas, una a la izquierda y otra a la derecha (amarilla y azul), que daban una carta cada vez que se pulsaba el ratón sobre el botón «Pedir carta». La posición de cada opción de respuesta fue contrabalanceada para cada grupo de participantes, la mitad de los cuales tenía como máquina A la de la izquierda y viceversa.

La tarea de los participantes consistió en las máquinas activar mediante clics del ratón hasta obtener una “carta ganadora” que otorgaba puntos bajo un programa de intervalo variable 5 s (ver Figura 1). Finalmente, se presentó una fase de prueba donde las dos máquinas se presentaban simultáneamente, una a la derecha y otra a la izquierda de la pantalla durante 60 s, sin reforzamiento, divididos en cuatro bloques de 15 s, en los que la posición de cada máquina dispensadora cambiaba de lugar para evitar sesgos de preferencia de lugar. Los participantes tenían que hacer una elección entre las opciones de respuesta A y B respondiendo sobre cada máquina.

Procedimiento

La asignación de los participantes a cuatro grupos se realizó de manera aleatoria. De acuerdo con el diseño de la Tabla 1, durante la fase de entrenamiento se presentaron cuatro bloques de diez ensayos, seguidos por un intervalo de retención y, finalmente, una prueba de elección donde se les presentaron a los participantes las dos opciones simultáneamente durante cuatro bloques de 15 s, en donde se registró su respuesta. Los grupos fueron similares en la densidad de reforzamiento a lo largo de cuatro bloques de diez ensayos (0.5 *vs.* 0.5), y solo se diferenciaban en el último o primer bloque de ensayos de entrenamiento (0.9 *vs.* 0.1). Así, para los grupos de densidad distante (DD 0 h y DD 24 h), la densidad de reforzamiento cambió durante

el primer bloque de ensayos, mientras que para los grupos de densidad reciente (DD 0 h y DD 24 h), la densidad de reforzamiento cambió durante el último bloque. Los grupos también fueron diferentes en el intervalo entrenamiento-prueba, así, la prueba de elección se realizó inmediatamente después de concluir el entrenamiento para dos grupos (grupos DR 0 h y DD 0 h) o después de 24 h para los grupos restantes (grupos DD 24 h y DD 24 h). Se registraron las respuestas emitidas durante la prueba de elección a cada una de las opciones de respuesta durante 60 s.

RESULTADOS

Para analizar los resultados se tomó en cuenta como una medida de la preferencia de los participantes la primera elección (A o B) durante la prueba, para posteriormente comparar los promedios grupales. De esta manera, se pudo analizar si la preferencia de los participantes sugiere recencia entendida como elegir la opción de mayor densidad de reforzamiento en el último bloque de diez ensayos (0.9 *vs.* 0.01) o primacía, entendida como la elección de esta opción durante el primer bloque de entrenamiento (0.9 *vs.* 0.01), o bien una integración de ambas opciones, considerada como el promedio de las densidades de reforzamiento en todos las fases de entrenamiento (0.6 *vs.* 0.4).

Los resultados obtenidos durante la prueba se presentan en la Figura 2 y muestran lo siguiente: a) un efecto de recencia en el grupo DR 0 h al preferir la opción A cuando la densidad de reforzamiento cambió en el último bloque y la prueba fue inmediata (0.9 > 0.1); b) indiferencia en la elección de ambas opciones cuando la densidad cambió en el primer bloque y la prueba fue inmediata para el grupo DD 0h (0.5 = 0.5); c) una preferencia por la opción B cuando la densidad cambió en el último bloque y la prueba se realizó a las 24 h en el grupo DR (promedio de A= 0.6); d) una preferencia por A cuando la densidad cambió en el primer

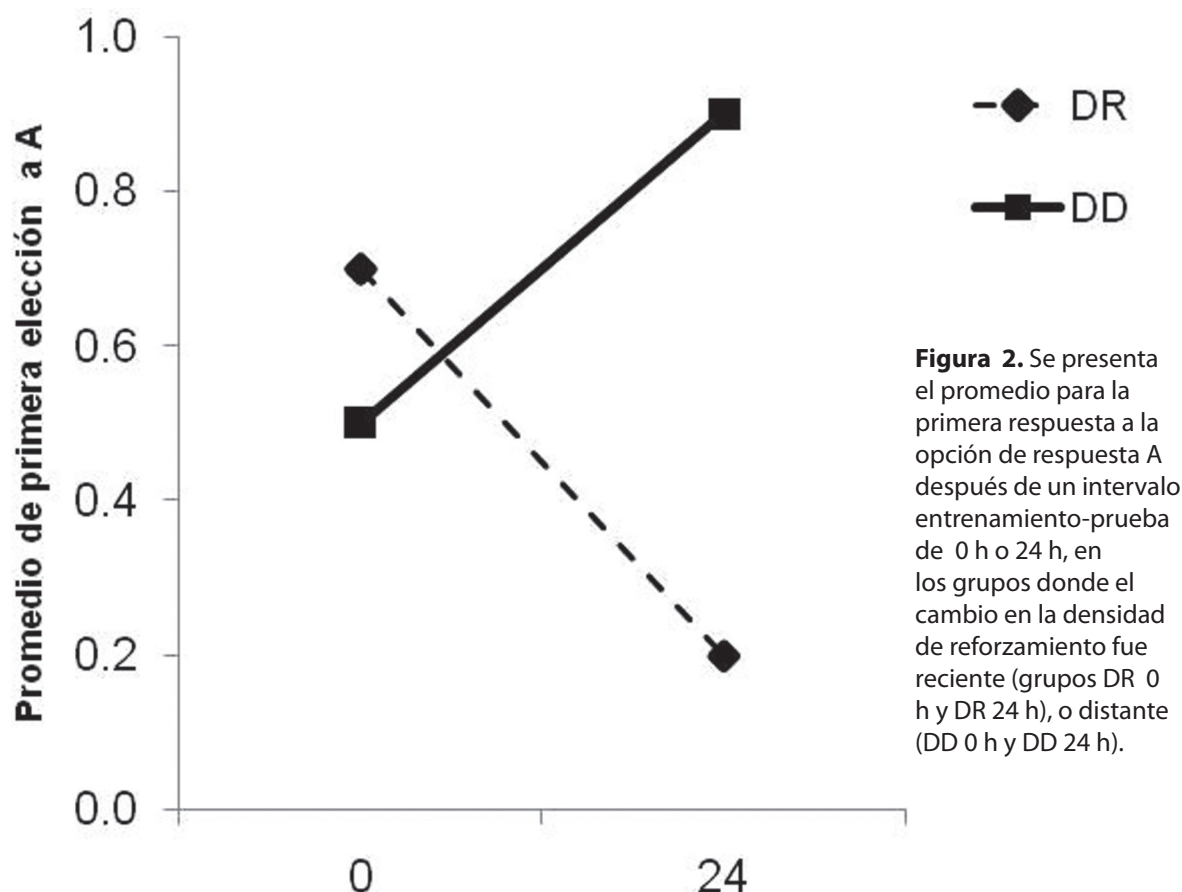


Figura 2. Se presenta el promedio para la primera respuesta a la opción de respuesta A después de un intervalo entrenamiento-prueba de 0 h o 24 h, en los grupos donde el cambio en la densidad de reforzamiento fue reciente (grupos DR 0 h y DR 24 h), o distante (DD 0 h y DD 24 h).

bloque de ensayos y la prueba de elección fue a las 24 horas para el grupo DD24 (promedio de $A = 0.6$). Lo anterior sugiere que la elección se pudo haber basado en la integración de información de todos los bloques de entrenamiento.

Los resultados se analizaron con un ANOVA de 2 (densidad) $\times$ 2 intervalo de retención) que mostró un efecto significativo de interacción ($F(1,36) = 6.04, p < .05$). Un ANOVA simple del factor densidad en ambos niveles de factor Intervalo mostró diferencias significativas en los dos valores del intervalo de retención de 0 y 24 h ($F(1,18) = 17.68, p < .05$), que sugiere una preferencia por A en el grupo DD 0, y una preferencia por A en el grupo DR 24.

DISCUSIÓN

Después de un entrenamiento donde los participantes eligieron entre dos opciones de respues-

ta (A y B) que se diferenciaban en el momento de cambiar la densidad de reforzamiento (DR y DD) y en el intervalo entrenamiento-prueba (0 y 24 h), se observó que los participantes eligen en una prueba inmediata la opción con una mayor densidad de reforzamiento (grupo DR 0), pero en una prueba demorada eligen aquella opción que en promedio tuvo una mayor densidad de reforzamiento (grupo DD 24). Lo anterior sugiere que en una prueba inmediata la recencia de cada experiencia tiene un peso mayor que su valor subjetivo, pero al paso del tiempo la recencia se desvanece y las opciones son elegidas, entonces, con base en su valor subjetivo, eligiéndose aquella opción con un promedio mayor de densidad de reforzamiento.

Estos resultados son coherentes con los resultados de estudios anteriores con humanos, que demuestran también que la elección se rea-

liza con base en la distancia temporal relativa y al valor subjetivo de cada experiencia de aprendizaje (López et al., 2010; Vila et al., 2010), tal y como lo predice la RPT (Devenport, 1998; Devenport & Devenport, 1994). Pero a diferencia de estos estudios, en el presente trabajo se empleó un procedimiento con más de dos fases de entrenamiento, que no incluye una fase de extinción, y similar al procedimiento de elección empleado originalmente por Mazur (1996).

Los resultados replican aquellos observados por Mazur (1996) en la conducta de elección en pichones con participantes humanos en una tarea de elección de cuatro fases. De manera similar a los datos observados en pichones, los resultados obtenidos sugieren que la información aprendida más recientemente tiene un mayor peso, pero después de un intervalo de tiempo la información aprendida de todas las experiencias se integra en un promedio dinámico entre el valor de recencia y el valor subjetivo de cada una. Así, los efectos de recencia e integración de información observados en la presente investigación pueden ser explicados por la RPT expresada en la ecuación 1.

De esta forma, los resultados obtenidos sugieren que en las pruebas inmediatas la elección de A se basó en la densidad de reforzamiento del último bloque del entrenamiento. Así, la selección de A fue mayor en el grupo DR 0 que en el grupo DD 0, ya que la densidad de reforzamiento de A fue mayor para este grupo en el último bloque (0.9 *vs.* 0.5), lo que es coherente con la idea de que los participantes, cuando elegían inmediatamente solo consideraban los bloques de entrenamiento más recientes. En cambio, el grupo DD 24 tuvo una preferencia mayor 24 por la opción A que el grupo DR en la prueba demorada, lo que sugiere que la elección de A se basó en su densidad de reforzamiento promedio durante los cuatro bloques de entrenamiento, más que en la densidad del último bloque (0.6 *vs.* 0.5). El grupo DR 24

mostró preferencia por B mayor a la esperada de 0.4, debido, quizás, a una falta de discriminación de las opciones que pudo alterar la preferencia de los participantes. Sin embargo, ese dato no altera las predicciones de la RPT mencionadas anteriormente en la introducción.

Así, para el caso del presente experimento, al variar la recencia 1/T (0 y 24 h) y el valor subjetivo QA (densidad mayor o igual) se obtiene que el valor de VwA es mayor cuando el valor de recencia 1/T es alto (0h) y determina, por lo tanto, la opción a elegir. Pero con el paso del tiempo, será el valor de QA (densidad 0.9) el que provoque que el valor de VwA se incremente, y se escoja entonces la opción A después de 24 h, tal y como ocurre en los grupos DR 0 y DD 24. De este modo, la ecuación 1 permite representar y predecir de forma adecuada los resultados observados en humanos en el presente estudio y aquellos observados en pichones (Mazur, 1996).

Esta dinámica de la elección nos permite predecir futuras elecciones al cambiar el peso de cada alternativa de acuerdo con su valor subjetivo (Q) y su distancia temporal relativa (1/T), pudiendo así predecir una recuperación de la experiencia inicial (primacía) después de un intervalo de tiempo cuando la primera experiencia tiene un mayor valor (Vila et al., 2010) o hay indiferencia entre ambas opciones cuando las experiencias tienen un valor similar (López et al., 2010). Hasta la fecha no existe ninguna otra aproximación teórica que permita explicar adecuadamente este cambio dinámico de recencia, primacía o indiferencia en la elección debido al paso del tiempo.

Los cambios en las preferencias que se observan en distintos momentos en un mismo individuo pueden ser explicados si suponemos que los animales humanos y no humanos almacenan una memoria para cada elección y que pasadas varias experiencias con cada opción estas se promedian al inicio de una nueva elección. De esta manera, empleando la RPT se puede

predecir la dirección y magnitud de la elección entre las opciones que será observada al final.

Así, por ejemplo, si consideramos el caso en el cual intentamos llegar a un destino determinado en un punto de la ciudad, evitando el tráfico denso a partir de dos posibles rutas con características distintas (A y B), una buena opción es elegir aquella que el día anterior haya presentado un tráfico menor, pero si esta elección es realizada una semana después, la mejor opción será considerar el tráfico promedio de las dos rutas durante la semana y elegir aquella que en promedio haya tenido menor tráfico. Y en el caso de que este promedio fuese similar para ambas, las características particulares de cada ruta (baches, paisaje, etc.) serán entonces las que determinen la elección.

Sin embargo, es necesario realizar investigación más detallada (por ejemplo, un mayor número de opciones de respuesta) que permita evaluar más ideas derivadas del supuesto de integración de información basado en la distancia temporal que propone la RPT.

REFERENCIAS

- Alvarado, A., Jara, E., Vila, J., & Rosas, J. M. (2006). Time and order effects on causal learning. *Learning and Motivation*, 37, 324-345.
- Bouton, M. E. (1993). Context, time and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, 114, 880-99.
- Devenport, L. D. (1998). Spontaneous recovery without interference: Why remembering is adaptive. *Animal Learning and Behavior*, 26(2), 172-181.
- Devenport, L. D., & Devenport, J. A. (1993). Time dependent decisions in dogs (*canis familiaris*). *Journal of Comparative Psychology*, 107(2), 169-173.
- Devenport, L. D., & Devenport, J. A. (1994). Time-dependent averaging of foraging information in least chipmunks and golden-mantled ground squirrels. *Animal Behavior*, 47, 787-802.
- López-Romero, L. J., García-Barraza, R., & Vila, J. (2010). Spontaneous recovery in human instrumental learning: Integration of information and recency to primacy shift. *Behavioral Processes*, 84, 617-621.
- Mazur, J. E. (1995). Development of preference and spontaneous recovery in choice behavior with concurrent variable-interval schedules. *Animal Learning & Behavior*, 23, 93-103.
- Mazur, J. E. (1996). Past experience, recency, and spontaneous recovery in choice behavior. *Animal Learning & Behavior*, 24(1), 1-10.
- Vila, J., Lopez, L. J., & Alvarado, A. (2010). Spontaneous recovery as a dynamic average of previous experiences in human instrumental conditioning. *International Journal of Psychology and Psychological Therapy*, 10(3), 403-413.

Recibido el 16 de diciembre de 2010

Revisión final 13 de abril de 2011

Aceptado el 14 de abril de 2011