



Revista Mexicana de Análisis de la Conducta
ISSN: 0185-4534
ISSN: 2007-0802
editor@rmac-mx.org
Sociedad Mexicana de Análisis de la Conducta
México

Castiello, Santiago; Burgos, José E.; Buriticá, Jonathan;
Valerio dos Santos, Cristiano; Alcalá, Jaime Emmanuel
Interacción entre magnitud y probabilidad de reforzamiento en la elección automoldeada
Revista Mexicana de Análisis de la Conducta, vol. 46, núm. 1, 2020, -, pp. 23-66
Sociedad Mexicana de Análisis de la Conducta
México

DOI: <https://doi.org/10.5514/rmac.v46.i1.76949>

Disponibile en: <https://www.redalyc.org/articulo.oa?id=59365747003>

- ▶ [Cómo citar el artículo](#)
- ▶ [Número completo](#)
- ▶ [Más información del artículo](#)
- ▶ [Página de la revista en redalyc.org](#)



Sistema de Información Científica Redalyc
Red de Revistas Científicas de América Latina y el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso
abierto

**INTERACCIÓN ENTRE MAGNITUD
Y PROBABILIDAD DE REFORZAMIENTO
EN LA ELECCIÓN AUTOMOLDEADA**

***INTERACTION BETWEEN MAGNITUDE AND
PROBABILITY OF REINFORCEMENT ON
AUTOSHAPED CHOICE***

Santiago Castiello

Centro de Estudios e Investigaciones en Comportamiento,
Universidad de Guadalajara
Department of Experimental Psychology,
University of Oxford

José E. Burgos

Jonathan Buriticá

Cristiano Valerio dos Santos

Jaime Emmanuel Alcalá

Centro de Estudios e Investigaciones en Comportamiento,
Universidad de Guadalajara

Santiago Castiello. Anna Watts Building, Radcliffe Observatory Quarter, Woodstock Rd, Oxford OX2 6HG, UK. Department of Experimental Psychology, University of Oxford. Correos electrónicos: santiagocdo@gmail.com; santiago.castiello@psy.ox.ac.uk
Agradecimientos: El proyecto fue financiado parcialmente por CONACYT INFR-281265, y la beca CONACYT de maestría No. 736730 a SCdO. Agradecemos al equipo editorial y los dos revisores por sus valiosos comentarios.

Resumen

Este estudio paramétrico evalúa el efecto de la magnitud y probabilidad de reforzamiento en elección automoldeada con redes neurales artificiales y palomas. El Experimento 1 fue una simulación con un modelo de redes neurales artificiales entrenadas separada e independientemente en distintas magnitudes y probabilidades de reforzamiento de dos señales, y posteriormente presentadas de manera concurrente en pruebas de elección en extinción sin aprendizaje (cambio de pesos). Las redes prefirieron una opción certera con poca recompensa que otra seguida 20% y 50% con mucha recompensa. Sin embargo, prefieren mucha recompensa cuando ésta se presentó el 80% de las veces. El Experimento 2 fue una prueba de estos resultados con palomas bajo condiciones análogas. Los resultados fueron en general consistentes con el Experimento 1, excepto que, a diferencia de las redes, las palomas mostraron indiferencia cuando la señal predictora de mayor recompensa lo hizo el 80 % de las veces. Se discute la interacción observada entre ambas contingencias y se plantea la posibilidad de que el contraste conductual se observe en contingencias pavlovianas.

Palabras clave: redes neurales artificiales, elección automoldeada, contingencias pavlovianas, preferencia, contraste conductual pavloviano, modelos computacionales de aprendizaje.

Abstract

This parametric study assessed the effects of the magnitude and probability of reinforcement in autoshaped choice with artificial neural networks and pigeons. Experiment 1 was a simulation with an artificial neural network model where networks were trained in different magnitudes and probabilities of Pavlovian reinforcement of two cues trained separately and independently, and afterwards presented concurrently in extinction choice tests without any learning. Results showed that the networks preferred a certain option with a small reward than another followed 20% and 50% with a big reward, but preferred the big

reward option when it was followed 80 % of trials. Experiment 2 was a test of these predictions with pigeons under analogous conditions. The results were generally consistent with those of Experiment 1, except that, unlike the networks, pigeons showed indifference when the cue that predicted a big reward did it 80% of trials. We discuss the interaction observed between contingencies and raise the possibility that behavioral contrast can arise under Pavlovian contingencies.

Key words: artificial neural networks, autoshaped choice, Pavlovian contingencies, preference, Pavlovian behavioral contrast

La extensa investigación en condicionamiento, tanto pavloviano como operante, muestra que los animales aprenden relaciones entre señales (estímulos sensoriales o exteroceptivos, e.g., luces o tonos, llamados “condicionados” en condicionamiento pavloviano, y “discriminativos” o “delta” en condicionamiento operante) y recompensas (estímulos biológicamente significativos, e.g., agua o comida, también llamados “incondicionados” en condicionamiento pavloviano, y “reforzadores primarios” en condicionamiento operante). Este aprendizaje no se limita a contingencias operantes (en las cuales el reforzador depende de alguna respuesta; e.g., Baum, 1974; Herrnstein, 1961), sino que también se ha observado bajo contingencias pavlovianas, en las cuales la recompensa no depende de respuesta particular alguna. Tampoco está restringido a respuestas viscerales o glandulares, sino que también se ha observado en respuestas de musculatura esquelética o estriada.

Un ejemplo paradigmático, central en el presente trabajo, es el autotomdeamiento de la respuesta de picoteo de una tecla en palomas hambrientas, reportado inicialmente por Brown y Jenkins (1968). El procedimiento básico consiste en presentar ensayos de una luz proyectada en una tecla de respuesta de una cámara de condicionamiento operante para palomas durante un tiempo (e.g., entre 4 s y 8 s), seguida por acceso breve al comedero lleno de grano. En este procedimiento, la luz funge de señal, mientras que la comida funge de “recompensa” o “reforzador” (en sentido pavloviano). El resultado principal, bien conocido, es que luego de varios apareamientos luz-comida, la mayoría

de las palomas empiezan a picar la tecla de respuesta. Este resultado muestra que una respuesta operante prototípica (picar una tecla en palomas) puede ser adquirida bajo contingencias pavlovianas. Este responder también se mantiene bajo las mismas contingencias (apareamientos luz-comida), sin necesidad de contingencia operante *explícita* alguna. Este mantenimiento puede llamarse automantenimiento, a falta de un mejor término (no confundir con automantenimiento negativo, en el cual hay una contingencia operante de omisión y no será objeto del presente trabajo).

El hecho de que durante el automantenimiento las respuestas a la señal son seguidas por comida en virtud de la contingencia pavloviana (señal-reforzador) ha llevado a la suposición de que en el automantenimiento también hay contingencias respuesta-reforzador que, aunque accidentales, implícitas o adventicias, pueden ejercer una influencia en el responder (ver Williams, & Williams, 1969; Yin, & Knowlton, 2006). Sin embargo, tal posible influencia aún no ha sido caracterizada experimentalmente de manera directa y en detalle, por lo que no la supondremos en este trabajo.

El automoldeamiento y automantenimiento muestran que al menos una forma de respuesta esquelética (picar una tecla iluminada) no provocada incondicionalmente por la recompensa (al menos en el sentido usual) puede ser adquirida y mantenida bajo contingencias pavlovianas. Esta observación ha sido muy discutida en relación con la distinción operante-respondiente (e.g., Schwartz & Gamzu, 1977), arguyéndose que cuestiona una distinción tajante. A pesar de esto, la mayoría de los analistas conductuales continúan haciendo tal distinción. Esta situación desafortunada es especialmente aguda en el estudio de la elección en el análisis experimental de la conducta, donde aún no se ha considerado la posibilidad de que contingencias pavlovianas puedan ejercer un efecto sustancial. Prueba de ello es que el único estudio experimental al respecto, reportado por Picker y Poling (1982), nunca ha sido citado en la literatura sobre elección y preferencia en análisis experimental de la conducta desde su publicación.

En su procedimiento básico, Picker y Poling (1982) primero entrenaron palomas privadas de alimento con un arreglo anterógrado demorado de automoldeamiento/automantenimiento. En este arreglo, el acceso a la comida fue contingente y contiguo (apareado) a dos luces de distintos colores, presentadas por separado y en diferentes teclas de respuesta. La comida fue siempre entregada independientemente de la conducta de las palomas, por lo cual no hubo contingencia operante explícita. Durante el entrenamiento, 100 % de los ensayos con un color fueron apareados con comida, y 50 % de los ensayos con otro color fueron apareados con comida. Luego, en una prueba de elección en extinción, se presentaron ensayos con ambos colores concurrentemente en teclas de respuesta separadas. Los resultados mostraron una tendencia a elegir el color apareado el 100% de las veces con el alimento durante el entrenamiento.

El objetivo principal de Picker y Poling (1982) fue mostrar que la elección era una variable dependiente sensible a ciertas variables independientes que parecían no afectar el automoldeamiento, en particular la probabilidad y magnitud de reforzamiento, pero su estudio va más allá. El automoldeamiento, estrictamente como la adquisición de la respuesta (de acuerdo a cierto criterio: tres ensayos de cuatro consecutivos con al menos una respuesta), no es afectado por la probabilidad de reforzamiento pavloviano, o $p(S^*|S)$, donde S^* denota la ocurrencia de la recompensa y S la señal. Sin embargo, esta variable sí afecta la preferencia por señales diferencialmente reforzadas en ensayos de elección luego del automantenimiento.

Burgos y García-Leal (2015) fueron los primeros en señalar que ese estudio representaba otro cuestionamiento a una separación tajante entre los dos tipos de condicionamiento, ya que implicaba que la preferencia podía emerger de contingencias Pavlovianas, sin necesidad de contingencias instrumentales explícitas (i.e., no accidentales). Tal reconocimiento llevó a estos autores a acuñar una nueva expresión para nombrar ese fenómeno, a saber, “elección automoldeada.” Este nombre expresa precisamente el procedimiento básico usado (automoldeamiento) y las condiciones bajo las cuales emerge la preferencia.

Burgos y García-Leal (2015) usaron el modelo de redes neurales propuesto por Donahoe, *et al.*, (1993), para simular los resultados básicos de Picker y Poling (1982, Experimento 1). Este modelo, justamente, es conocido por eliminar una distinción entre *aprendizaje* operante y respondiente, donde “aprendizaje” se refiere estrictamente a cambios en eficacias sinápticas, simulados en el modelo por cambios de pesos en las conexiones entre elementos o unidades neurales virtuales. El modelo hace esto preservando las diferencias operacionales y conductuales entre los dos tipos de condicionamiento, en especial, respuestas (evocadas versus emitidas) y contingencias (pavlovianas, en las cuales el reforzamiento no depende de respuesta alguna, versus instrumentales, en las cuales el reforzamiento depende de alguna respuesta). Al haber simulado exitosamente aquellos resultados con el modelo, Burgos y García-Leal (2015) propusieron una primera aproximación teórica al posible papel de las contingencias pavlovianas en la elección, señalando, además, su posible relevancia para la economía conductual y la neuroeconomía.

El presente trabajo propone expandir esta línea de investigación caracterizando el papel de las contingencias pavlovianas en la elección y preferencia paramétricamente en las dimensiones de magnitud (2 valores) y probabilidad (3 valores) de la recompensa, tanto en redes neurales artificiales y en palomas. La investigación se inició con una simulación computacional inspirada por el estudio de Burgos y García-Leal (2015).

Específicamente, el objetivo fue simular una de las implicaciones del Experimento 2 de Picker y Poling (1982) con el modelo, para determinar sus predicciones al respecto, y luego probar los resultados de la simulación con palomas en condiciones análogas. Estos autores reportaron que aumentando la magnitud de recompensa de la señal parcialmente reforzada (50% de los ensayos) y disminuyendo la de la señal totalmente reforzada (100% de los ensayos), las palomas cambiaban su preferencia; es decir, durante una prueba de elección prefirieron la señal parcialmente reforzada. También observaron que, al regresar las magnitudes de recompensa a las originales, la preferencia

se revertía. En el presente experimento se estudió el efecto de la probabilidad en lugar de la magnitud de reforzamiento. Los resultados de estos estudios podrían llevar a una mejor comprensión de la elección y la preferencia bajo contingencias operantes, al mostrar una posible contribución de contingencias pavlovianas.

El presente estudio fue inspirado por el modelo de redes neurales propuesto por Donahoe, *et al.*, (1993) como una interpretación conexionista de un principio unificado del reforzamiento (Donahoe, *et al.*, 1982). Según esta interpretación, las diferencias empíricas entre condicionamiento pavloviano y operante, en especial la distinción entre dos tipos de respuesta y contingencias de reforzamiento, no implican que no tengan un mecanismo subyacente común de aprendizaje, entendido de manera conexionista (i.e., cambios en pesos de conexiones, según una regla de aprendizaje descrita más adelante).

El modelo ha sido exitosamente usado para simular (y en este sentido explicar) varios fenómenos de condicionamiento, tanto Pavloviano como operante. Los primeros incluyen adquisición, extinción, y readquisición más rápida (Donahoe *et al.*, 1993); funciones de intervalo entre estímulos (Burgos, 1997; Donahoe & Burgos, 2000); inhibición latente (Burgos, 2003); efectos de la razón C/T (Burgos, 2005); especificidad y renovación contextual (Burgos & Murillo, 2007); condicionamiento simultáneo (Burgos *et al.*, 2008); condicionamiento de segundo orden y resistencia a la extinción (Sánchez *et al.*, 2010); bloqueo y ensombrecimiento (Burns *et al.*, 2011; Burgos & Donahoe, 2016).

Simulaciones de condicionamiento operante han sido menos numerosas, pero no menos importantes. Hasta ahora, el modelo ha simulado adquisición, extinción, readquisición más rápida, generalización, y discriminación bajo contingencias operantes (Donahoe, *et al.*, 1993); ejecución en programas de intervalo fijo (Burgos & Donahoe, 1999; Donahoe & Burgos, 1999); y reevaluación del reforzamiento (Donahoe & Burgos, 2000). También ha mostrado simular conducta estable en programas de intervalo variable (Calvin & McDowell, 2015). Aún está por verse si el modelo puede simular otros fenómenos de condicionamiento operante (e.g., moldeamiento por aproximacio-

nes sucesivas, contraste conductual, reforzamiento negativo, castigo positivo, ejecución en programas de razón fija, ley de igualación, igualación a la muestra, etc.). El modelo, sin embargo, también simula automoldeamiento y automantenimiento, los cuales son centrales para el presente trabajo. Ello permite teorizar sobre el posible papel de las contingencias pavlovianas en el condicionamiento de respuestas emitidas (i.e., no provocadas incondicionalmente por el reforzador), algo que ningún otro modelo permite.

El modelo consta de dos componentes, cada uno correspondiente a un nivel de organización neural distinto. En el componente de redes neurales se especifican distintos tipos de unidades que pueden constituir una red y pautas generales para conectar las unidades entre sí, inspiradas por principios generales de neuroanatomía (ver Donahoe, *et al.*, 1993). El otro componente es la parte computacional, que especifica una regla de activación de las unidades (nivel celular) y una regla de aprendizaje, la cual determina el cambio de pesos en las conexiones (nivel sináptico), cambio que en este tipo de modelos define el aprendizaje.

El hecho de que el modelo proponga un mismo mecanismo de aprendizaje (cambio de pesos de conexiones) para el condicionamiento bajo contingencias pavlovianas y operantes no significa que elimine otros aspectos de la distinción. El modelo preserva dos elementos centrales de la distinción. Por una parte, distingue entre respuestas emitidas (no provocadas incondicionalmente por el reforzador) y respuestas provocadas incondicionalmente por el reforzador (o puramente pavlovianas; ver Gormezano & Kehoe, 1975). El modelo también distingue entre contingencias pavlovianas, en las cuales el reforzamiento es independiente de la conducta, y contingencias operantes, en las cuales el reforzador depende explícitamente de alguna respuesta. Ambas distinciones han sido tradicionalmente tratadas como si fueran lógicamente codependientes. Sin embargo, el presente modelo es el único que cuestiona tal tratamiento, lo cual permite posibilidades teóricas que no se encuentran en otros modelos, en especial la adquisición y manteni-

miento de respuestas emitidas bajo contingencias pavlovianas, precisamente lo que ocurre en el automoldeamiento y automantenimiento.

En la Figura 1 se muestra la red neural usada en la simulación en el Experimento 1 (la misma usada por Burgos & García-Leal, 2015). Es una red de conectividad unidireccional, es decir, una conexión va de una unidad a otra pero no viceversa. Una unidad, entonces, puede ser activada solo por unidades de las cuales recibe conexión. Tal conectividad solo simula una relación *espacial* entre unidades (análoga a las relaciones espaciales descritas en neuroanatomía) y no determina un orden temporal particular de cómputo de activaciones.

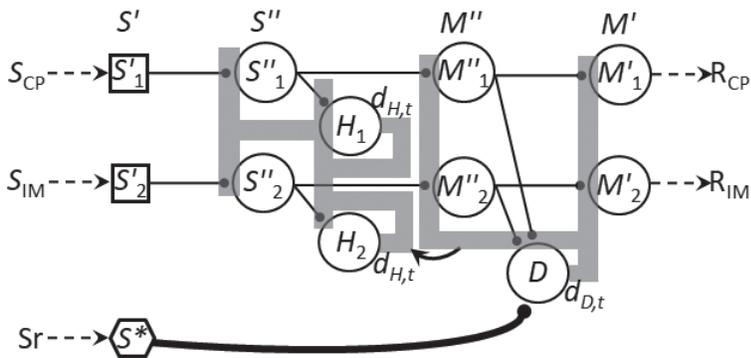


Figura 1. Arquitectura de red utilizada en Experimento 1. S_{CP} e S_{IM} : señales exteroceptivas correlacionadas con contingencias Certera-Poca (CP) e Incierta-Mucha (IM), respectivamente, explicadas más adelante. Sr: Estímulo apetitivo (e.g., comida), usado como reforzador. Flechas punteadas de S_{CP} , S_{IM} y Sr: proceso de transducción, el cual no se simula. S' : capa de entrada la cual simula grupos sensoriales primarios de neuronas. Cuadrados con S'_1 y S'_2 : unidades que su activación simula la presencia de estímulos S_{CP} e S_{IM} respectivamente. Hexágono con S^* : unidad que simula la presencia de Sr. Línea negra gruesa de S^* a D : conexión fija con peso máximo. Círculos: representan unidades neurocomputacionales (su activación se calcula con base en la regla de activación del modelo). Líneas delgadas con terminaciones circulares: conexiones iniciales débiles (modificables de acuerdo a la regla de aprendizaje). S'' : capa sensorial asociativa. S''_1 y S''_2 : unidades que simulan grupos neuronales polisensoriales. H_1 , H_2 : unidades que simulan áreas hipocámpicas (e.g. ca1). D : unidad tipo dopaminérgica. Cuadros grises: señal difusa provenientes de unidades H o D , ésta modula los pesos de las conexiones (de acuerdo a regla de aprendizaje). M'' : Capa motora asociativa. M''_1 y

M''_2 : unidades motoras secundarias. M' : capa (de salida) motora primaria. M'_1 y M'_2 : unidades primarias motoras precursoras de respuestas R_{CP} y $R_{IM'}$ respectivamente. Flechas punteadas hacia R_{CP} y $R_{IM'}$: transducción de área motora primaria a efectores (no simulada). R_{CP} y $R_{IM'}$: representan respuestas a distintos estímulos. Los valores de los parámetros libres son los expuestos en la sección del modelo en introducción; los pesos iniciales para todas las conexiones iniciales débiles son de 0.01.

La red está constituida por dos subredes: sensorial y motora. El modelo requiere que una red tenga al menos cuatro capas de unidades: S' (capa de entrada, sensorial-primaria), S'' (capa oculta polisensorial), M'' (capa oculta motora asociativa) y M' (capa de salida motora primaria). Por tener al menos dos capas ocultas (S'' y M''), se trata de una red de *aprendizaje profundo* (o "deep learning"). El procesamiento en la red ocurre de izquierda a derecha, empezando con las activaciones de entrada (S') que simulan eventos ambientales, como estímulos sensoriales o reforzadores.

La subred sensorial consta de las unidades S' (tres en la presente red, incluyendo la unidad S^*), S'' (dos unidades en esta red) y dos H . Las unidades H simulan áreas hipocampales que han mostrado estar críticamente involucradas en el condicionamiento operante y condicionamiento pavloviano (e.g. CA1), así como también en el automoldeamiento (e.g., Richmond & Colombo, 2002). Activaciones de las unidades S' simulan efectos sensoriales primarios de eventos ambientales usados como estímulos en estudios de condicionamiento (e.g., luces, tonos, etc.). Para el presente estudio, activaciones de S'_1 simulan efectos sensoriales primarios de una señal (S_{CP} , en la simulación descrita más adelante, e.g., una luz roja), mientras que activaciones de S'_2 simulan efectos sensoriales primarios de una señal distinta ($S_{IM'}$, e.g., una luz verde). Otra simplificación estratégica, para facilitar la simulación e interpretación de sus resultados, es que no se incluyó una unidad de entrada cuya activación simulara señales contextuales. Queda aún por investigar el papel que, según el modelo, juegan esas señales en la elección automoldeada, aunque se esperaría que el modelo predijera correctamente al menos una especificidad contextual de esta conducta

(e.g., una preferencia menos pronunciada cuando la prueba de elección se hace en un contexto distinto del de entrenamiento).

Las unidades H simulan el sistema hipocampal generando una señal difusa de discrepancia que modula cambios en las fuerzas o pesos de las conexiones $S'-S''$ y $S''-H$ (así como $S''-S''$ en redes que tengan más de una capa S'' , que no es el caso en la presente red), de acuerdo con la regla de aprendizaje (ver componente computacional en la siguiente sección). Burgos y Galeazzi (en prensa) han mostrado que, según el modelo, esta señal hipocampal aminora el efecto deteriorante de señales relativamente débiles (menos intensas) y la ausencia del reforzador (activación de S^* de 0.0).

La subred motora consta de las unidades M'' , las cuales simulan corteza prefrontal y premotora (o asociativa); D , la cual simula áreas dopaminérgicas (e.g. área ventral tegmental -VTA), que también han mostrado estar críticamente involucradas en automoldeamiento (ver Ikemoto, 2007), y M' , que simula corteza motora primaria. La unidad D genera una señal difusa de discrepancia que modula cambios en los pesos de las conexiones $S''-M''$, $M''-D$ y $M''-M'$ (así como también $M''-M''$ en redes con más de una capa M'' , que no es el caso en la presente red).

La distinción antes mencionada entre respuestas emitidas (no provocadas incondicionalmente por el reforzador) y respuestas provocadas incondicionalmente por el reforzador se hace en el modelo sobre la base de una distinción entre dos tipos de unidades M' . Puesto que el énfasis del presente estudio es sobre respuestas emitidas, la red solo tiene unidades M' , para simplificar la simulación e interpretación de sus resultados. Las activaciones de ambas unidades de salida en esta red, entonces, simulan precursores motores primarios solo de respuestas emitidas. Se conciben como emitidas en tanto las unidades precursoras de respuesta, M'_1 y M'_2 , solo pueden ser activadas por las unidades S'_1 y S'_2 , respectivamente, a través de las capas internas (S'' y M''), lo cual es posible solo mediante un entrenamiento que permita que las conexiones correspondientes ganen suficiente peso para permitir tales activaciones. En el presente estudio, tal entrenamiento consistió de contingencias pavlovianas entre activaciones de S'_1 o S'_2 , por una parte, y S^* , por otra.

Unas unidades activan a otras mediante conexiones, análogas a sinapsis o grupos sinápticos en circuitos neuronales naturales. Las conexiones se representan en la Figura 1 por líneas con botones terminales. La línea gruesa representa una conexión fija máximamente fuerte (con un peso de 1.0; este peso *nunca* se usa para el cómputo de activación o peso alguno en el componente neurocomputacional). Las líneas finas representan conexiones variables, inicialmente débiles cuya fuerza cambia de acuerdo con la regla de aprendizaje (ver componente computacional más adelante). La fuerza de una conexión es numéricamente representada por un peso.

Las unidades están conectadas de la siguiente manera. S'_1 está conectada a S''_1 , mientras que S'_2 está conectada a S''_2 . Por su parte, S''_1 está conectada a H_1 y M''_1 , mientras que S''_2 está conectada a H_2 y M''_2 . Ambas unidades M'' están conectadas a D , pero M''_1 también está conectada a M'_1 y M''_2 también está conectada a M'_2 . El resultado es una arquitectura en la cual una señal (activación de S'_1 o S'_2) afecta una vía sensoriomotora particular (e.g., S'_1 - S''_1 - M''_1 - M'_1). Este tipo de arquitectura, entonces, es máximamente especializada en su estructura y funcionamiento sensoriomotor, donde distintas señales afectan distintas vías sensoriomotoras independientes (con excepción de D). Por supuesto, ello no es el caso en animales, por lo que esta arquitectura es una posibilidad teórica límite como punto de partida conveniente para facilitar la teorización.

El nivel computacional del modelo consta de reglas para determinar cómo se activan las unidades y cambian los pesos de las conexiones de una red. Las reglas son, respectivamente, de activación y de aprendizaje.

La regla de activación es una función condicional con al menos tres modos mutuamente excluyentes en cualquier momento t en tiempo discreto: activación incondicional, reactivación y decaimiento. Un cuarto modo involucra unidades inhibitorias, pero se ignorará aquí porque la red usada en las simulaciones carece de tales unidades. Por tanto, solo se mostrará la versión abreviada de la regla de activación, sin inhibición y sin unidades de salida incondicionalmente activadas por

S^* . A continuación de muestra su forma algebraica, como una función condicional de tres posibles estados mutuamente excluyentes en t :

$$a_{j,t} = \begin{cases} a_{S^*,t}, & \text{si } a_{S^*,t} > 0 \text{ y } j \text{ es } D \text{ (activación incondicional de } D), \text{ de otro modo:} \\ L(exc_{j,t}) + \tau_j L(exc_{j,t-1})[1 - L(exc_{j,t})], & \text{si } L(exc_{j,t}) \geq \theta_{j,t} \text{ (reactivación), o} \\ a_{j,t-1} - \kappa_j a_{j,t-1}(1 - a_{j,t-1}), & \text{si } L(exc_{j,t}) < \theta_{j,t} \text{ (decaimiento)} \end{cases} \quad (\text{Ec. 1})$$

donde $a_{S^*,t}$ en el modo de activación incondicional denota la activación de la unidad S^* (reforzador) en t . En la red mostrada en la Figura 1, este modo se aplica solo si la unidad cuya activación se desea calcular en t es D y si la activación de S^* en t es mayor que cero (lo cual simularía efectos sensoriales-primarios de comida con cierta magnitud simulada por el nivel de activación de S^* , pero definida para las palomas en el Experimento 2 como tiempo de acceso al comedero). En otras simulaciones, este modo incluye unidades de salida activadas incondicionalmente por S^* , pero la presente arquitectura, como ya se ha aclarado, carece de tales unidades, como simplificación estratégica para facilitar la teorización. Por ello, también se eliminó la parte correspondiente del modo de activación incondicional.

Si $a_{S^*,t} = 0$ o j no es D , entonces la regla entra en modo de reactivación o de decaimiento, dependiendo de las condiciones. En la Figura 2 se muestra una unidad genérica j oculta o de salida con todos los factores de ambos modos. La unidad j (cuya activación se desea calcular para un momento t) recibe los aferentes $a_{1,t}, \dots, a_{i,t}, \dots$ y $a_{n,t}$ desde unidades presinápticas, de entrada u ocultas, cada una conectada por las líneas delgadas con conexiones variables inicialmente débiles en cada momento temporal t ; estos aferentes son elementos del vector de activaciones que afectan j en t ($a_{j,t}$). Cada conexión tiene una fuerza representada numéricamente por pesos $w_{1,t}, \dots, w_{i,t}, \dots$ y $w_{n,t}$ en t , una conexión por cada aferente y un peso por cada conexión. Cada peso es un elemento del vector de pesos para j en t ($w_{j,t}$). Tanto los pesos (w) como las activaciones (a) son números reales entre 0.0 y 1.0.

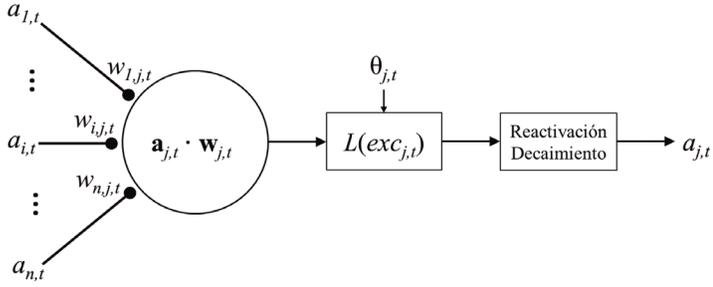


Figura 2. Unidad neurocomputacional genérica (oculta o de salida) para los modos de reactivación y decaimiento de la regla de activación (ver Ec. 1). i : unidad aferente (presináptica). j : unidad objetivo genérica (S'' , M'' , H , D o M' ; ver Figura 1). $a_{i,t}$: activación aferente excitatoria. $w_{ij,t}$: peso de la conexión de i a j . $exc_{j,t}$: producto interno de los vectores $\mathbf{a}_{j,t}$ y $\mathbf{w}_{j,t}$ para j en el momento t . L : Función logística con $exc_{j,t}$ como argumento. $\theta_{j,t}$: umbral gaussiano. $a_{j,t}$: activación condicional eferente de j en t .

Para ambos modos,

$$exc_{j,t} = \mathbf{a}_{j,t} \cdot \mathbf{w}_{j,t} = \sum_{i=1}^n a_{i,t} w_{i,j,t} \quad (\text{Ec. 2})$$

donde $\mathbf{a}_{j,t}$ denota el vector de activaciones de los aferentes que recibe j (j permanece constante en esta ecuación, ya que denota la unidad particular cuyos pesos se van a actualizar) en t , $\mathbf{w}_{i,j,t}$ denota el vector de pesos de las conexiones de todas las unidades conectadas a j ($1 \leq i \leq n$), y n es el número total de unidades conectadas a j . $exc_{j,t}$ es el argumento de la función logística L , definida como:

$$L(x) = \frac{1}{1 + e^{\frac{-x + \mu}{\sigma}}} \quad (\text{Ec. 3})$$

donde $x = exc_{j,t}$, $\mu = 0.5$ y $\sigma = 0.1$. Los parámetros libres se mantuvieron constantes en todas las simulaciones y permiten un valor de la

función mayor que cero (aproximadamente 0.006) cuando $exc_{j,t} = 0.0$, lo cual simula una activación espontánea cercana a cero.

La regla entra en modo de reactivación o de decaimiento dependiendo de $\theta_{j,t}$ un umbral aleatorio que sigue una distribución gaussiana con media de 0.2 y desviación estándar de 0.15. El umbral aleatorio se calcula en cada momento temporal t para cada unidad computacional, oculta o de salida, en la red. Otros parámetros libres son el de sumación temporal (τ_j) y el de decaimiento (κ_j), donde $\tau_j = 0.1$ y $\kappa_j = 0.1$ para todas las unidades, redes y simulaciones.

Esta regla se utiliza para cambiar el peso (w) de una conexión de una unidad presináptica i a una unidad postsináptica j en cada momento t . Un peso se puede entender como la eficacia con la cual una unidad activa a otra. En este modelo, todos los pesos, así como las activaciones, son números entre 0.0 y 1.0. Por lo tanto, no hay pesos negativos, a diferencia de otros modelos. La regla de aprendizaje se define en términos del cambio de (Δ) de los pesos y tiene la siguiente forma:

$$\Delta w_{i,j,t} = \begin{cases} \alpha_j a_{j,t} d_t p_{i,t} r_{j,t}, & \text{si } d_t \geq 0.001; \text{ si no,} \\ -\beta_j w_{i,j,t-1} a_{i,t} a_{j,t} \end{cases} \quad (\text{Ec. 4})$$

donde $\alpha = 0.5$ (es la tasa de incremento de los pesos) y $\beta = 0.5$ (es la tasa de decremento de los pesos) son parámetros libres que permanecieron constantes para todas las unidades, redes y simulaciones. Los otros términos son $a_{i,t}$ (la activación de una unidad presináptica), $a_{j,t}$ (activación de una unidad postsináptica) y

$$p_{i,t} = \frac{a_{i,t} w_{i,j,t-1}}{x}, \text{ donde } x = exc_{j,t}; \quad (\text{Ec. 5})$$

$$r_{j,t} = 1 - \sum_{i=1}^n w_{i,j,t} \quad (\text{Ec. 6}).$$

El factor central de la regla es d_t , una señal difusa que modula los cambios en todos los pesos en cada momento t . Es una diferencia tem-

poral entre activaciones de las unidades H y D en momentos sucesivos t y $t - 1$, y se calcula de la siguiente forma:

$$d_t = \begin{cases} |a_{H,t} - a_{H,t-1}| + d_{D,t}(1 - d_{H,t}), & \text{si } j \text{ es } S'' \text{ o } H; \\ a_{D,t} - a_{D,t-1}, & \text{si } j \text{ es } M'', D, \text{ o } M' \end{cases} \quad (\text{Ec. 7})$$

Los factores $p_{i,t}$ y $r_{j,t}$ introducen una competencia entre las conexiones que convergen en j por una cantidad limitada de peso (1.0). Además, $p_{i,t}$ denota qué proporción de x es el producto de la activación de i (conectada a j) y el peso de su conexión correspondiente e involucra un componente de aprendizaje hebbiano. Por su parte, $r_{j,t}$ hace que el cambio de peso también dependa de la cantidad disponible de peso en j , de tal manera que a menor peso disponible en la unidad j , menor será el cambio de peso en todas las conexiones que confluyen en j . Exactamente la misma regla de aprendizaje se usó para cambiar los pesos de todas las conexiones en todas las unidades y redes en el Experimento 1. Aunque este estudio se concentró sobre contingencias pavlovianas, la misma regla se usa en cualquier simulación que involucre contingencias operantes. Es en este sentido que el presente modelo postula un mecanismo neural de aprendizaje común a la adquisición y mantenimiento de respuestas emitidas y/o respuestas incondicionadas/condicionadas (ausentes en el presente estudio), bajo contingencias pavlovianas (él énfasis del presente estudio) o contingencias operantes (no usadas en este trabajo).

Las activaciones de todas las unidades y los pesos de todas las conexiones se actualizan en cada momento temporal según un procedimiento aleatorio asincrónico, en el cual el orden de cómputo de pesos y activaciones es aleatorio de un momento temporal a otro, con un muestreo sin reemplazo que introduce una posible desincronización entre activaciones de máximo 1 momento temporal (i.e., la activación de una unidad j en t puede, por azar, depender de la activación de otra i (conectada a j) en $t-1$).

Experimento 1

Para esta simulación, se adoptaron tres conjeturas, hipótesis teóricas aún por confirmar con animales, que también fueron adoptadas por Burgos y García-Leal (2015). Primera, cada estímulo exteroceptivo (e.g., tecla roja y tecla con cruz) afecta una vía de procesamiento sensoriomotor independiente (ver Figura 1). Segunda, las activaciones de distintas unidades de salida, que simulan precursores motores primarios de respuestas, no son mutuamente excluyentes y, por tanto, pueden activarse en el mismo momento; ello difiere de lo que sucede en la elección automoldeada con palomas, en la cual respuestas a distintas señales, presentadas concurrentemente, son mutuamente excluyentes. Aunque el efecto de simular respuestas mutuamente excluyentes en elección automoldeada en estas redes aún está por investigarse cabalmente, el estudio de Burgos y García-Leal (2015) implica que tal exclusividad no es necesaria para la conducta de elección, una implicación novedosa en espera de prueba empírica.

Tercera, ninguna unidad de salida es activada por S^* , por lo que la arquitectura neural no simula reflejo incondicionado alguno (los precursores motores primarios y efectos sensoriales primarios de un reflejo incondicionado se simularían, respectivamente, por la activación incondicional de alguna unidad de salida por la unidad S^*). Por ende, la arquitectura tampoco simula condicionamiento pavloviano, en la medida en que éste requiera de un reflejo incondicionado (ver Gormezano & Kehoe, 1975). Por supuesto, como otros mencionados, este es otro límite teórico como simplificación estratégica con el objeto de facilitar el uso del modelo para explicar y predecir. Es razonable suponer que la consumición de grano provoca algún reflejo incondicionado digestivo en las palomas, pero, de nuevo, esta posibilidad será ignorada en este trabajo, a manera de abstracción teórica, ya que el punto importante aquí es que el picoteo sobre la tecla iluminada no es un reflejo incondicionado, al menos no en el sentido estándar del término.

Esto es simulado en el modelo por unidades de salida que nunca son activadas por S^* y solo pueden ser activadas por las unidades S' ,

mediante las unidades S'' y M'' , luego de aumento de pesos que permita tales activaciones. Otra característica de las respuestas automoldeadas es que son dirigidas a partes específicas del ambiente, pero no se simulará esta característica en el presente trabajo (ver Burgos, 2007). La característica crucial será que la consumición de comida (simulada por la activación de S^*) no provoca incondicionalmente el picoteo sobre la tecla iluminada.

Lo anterior permite una distinción entre contingencias pavlovianas, en las cuales la operación de reforzamiento (entrega de un reforzador primario, e.g., comida) es independiente de la respuesta de interés (picar la tecla iluminada) y condicionamiento pavloviano, definido como el cambio en la función de una señal de neutra a condicionada, donde la respuesta condicionada usualmente se asemeja a la incondicionada, lo cual presupone un reflejo incondicionado (ver Gormezano & Kehoe, 1975). Aunque históricamente ambas distinciones han ido de la mano (tal cambio ocurre gracias a contingencias pavlovianas), este modelo es el único que permite separarlas teóricamente. Ello posibilita la simulación de contingencias pavlovianas sin condicionamiento pavloviano. Burgos y García-Leal (2015) propusieron que las respuestas mutuamente excluyentes, la competencia entre respuestas y el condicionamiento pavloviano, no son necesarios para la elección automoldeada.

En este experimento, se usó el modelo descrito en la sección anterior para simular el efecto de la probabilidad de reforzamiento de una señal S_{IM} (definida como la activación máxima de S'_2 ; ver Figura 1) sobre la elección entre ésta y otra señal S_{CP} (definida como la activación máxima de S'_1 ; ver Figura 1), donde “reforzamiento” se refiere a “reforzamiento pavloviano” (independiente de cualquier respuesta). En las etiquetas, “S” significa “señal,” “I” significa “incierto,” “M” significa “mucha,” “C” significa “certera,” y “P” significa “poca.” Las señales fueron entrenadas con distintas magnitudes del reforzador (tiempo de acceso al comedero). Picker y Poling (1982, Experimento 2) exploraron ambas variables, usando cinco valores de la magnitud del reforzador (2, 4, 8, 10 y 14 s de acceso al grano) y dos valores de la probabilidad reforzamiento (0.5 y 1.0). Sus resultados mostraron que la preferencia

fue reversible al cambiar la magnitud de reforzamiento entre condiciones y posteriormente regresar a la condición inicial. En el presente estudio, se usaron dos magnitudes de reforzamiento (M , definida como una activación de $S^* = 1$, vs. P , definida como una activación de $S^* = 0.7$) y tres valores de $p(\text{Sr}|S_{\text{IM}})$ (i.e., 0.2, 0.5 y 0.8). El objetivo fue simular el efecto de esta probabilidad y magnitud de la recompensa en elección automoldeada en redes neurales artificiales.

Método

Sujetos

Cuarenta redes con la arquitectura mostrada en la Figura 1 formaron dos grupos de 20 redes cada uno, llamados “Ascendente” (As) y “Descendente” (De), en referencia al orden de presentación de los valores de $p(\text{Sr}|S_{\text{IM}})$ en un diseño intrasujeto. Tal orden sirvió de contrabalanceo de los valores de $p(\text{Sr}|S_{\text{IM}})$.

Materiales e Instrumentos

Computadora con Windows® y con el simulador SelNet1 instalado (diseñado y codificado por José E. Burgos).

Procedimiento

Se usó un diseño intrasujeto de cuatro fases (ver Tabla 1). Ambos grupos pasaron por condiciones distintas desde la primera hasta la tercera fase. En la última fase se repitió la condición de la Fase 1, según un diseño ABCA, donde cada letra denota una fase y corresponde a un valor de $p(\text{Sr}|S_{\text{IM}})$ (0.2, 0.5 o 0.8).

Tabla 1. Probabilidad de recompensa dada S_{IM} o $p(Sr|S_{IM})$

Grupo	Fases			
	1 (A)	2 (B)	3 (C)	4 (A)
De	0.8	0.5	0.2	0.8
As	0.2	0.5	0.8	0.2

Nota: Al lado del número de fase, entre paréntesis, se muestra la condición del diseño experimental. Probabilidad de recompensa (Sr) dada S_{IM} o $p(Sr|S_{IM})$ en la contingencia IM. Sr en esta contingencia se simuló como un nivel de activación de S^* de 1.0. Dicho nivel fue de 0.7 para S_{CP} en CP, donde $p(Sr|S_{CP}) = 1.0$ para ambos grupos y todas las fases. Los grupos **As** y **De** solo jugaron un papel de contrabalanceo parcial del orden de los valores de $p(Sr|S_{IM})$.

Cada fase tuvo dos etapas sucesivas: entrenamiento y prueba. Durante el entrenamiento, se presentaron las señales S_{CP} y S_{IM} individualmente y entremezcladas de manera aleatoria. S_{CP} se asoció con la contingencia CP, mientras que S_{IM} se asoció con la contingencia IM. En la contingencia CP, la señal S_{CP} tuvo una duración de cinco momentos temporales, en el último de las cuales se activaba S^* con un nivel de 0.7. Esta activación de S^* simuló una recompensa P (de baja, pequeña o poca magnitud) pero certera (C), puesto que $p(Sr|S_{CP})=1.0$. Esta probabilidad se mantuvo constante para todos los sujetos y condiciones.

Por su parte, S_{IM} (la señal asociada con la contingencia IM) consistió en la activación máxima de S'_2 también por cinco momentos temporales. Sin embargo, a diferencia de $p(Sr|S_{CP})$, $p(Sr|S_{IM})$ varió entre las condiciones A, B y C del diseño (ver Tabla 1). Para contrabalancear parcialmente el orden de presentación de los valores de $p(Sr|S_{IM})$, se formaron los grupos **As** (ascendente) y **De** (descendente). Para S_{IM} si en el último momento de un ensayo ocurría el reforzador, S^* se activaba con el nivel máximo de 1.0, para simular una recompensa mayor. Como otra simplificación estratégica, no se simuló explícitamente el intervalo entre ensayos (IEEn) sino que se asumió que fue lo suficientemente largo como para que las activaciones de todas las unidades computacionales de la red (ocultas, H , D y salida) decayeran a sus niveles mínimos determinados por la regla de activación (aproximadamente, 0.006).

El entrenamiento consistió en 100 ensayos de S_{CP} entremezclados aleatoriamente con 100 ensayos de S_{IM} . Luego del entrenamiento, las redes recibieron 30 ensayos de prueba de elección en los que ambas señales, S_{CP} y S_{IM} fueron presentadas (i.e., S'_1 y S'_2 en la Figura 1 fueron activadas) simultáneamente por cinco momentos, en extinción (i.e., con una activación S^* de 0.0). Durante la prueba, se deshabilitó la regla de aprendizaje para permitir que las activaciones de salida fueran debidas únicamente a los pesos finales del entrenamiento, sin que éstos se modificaran debido a la ausencia de S_r . El procedimiento gráfico se muestra en la Figura 3.

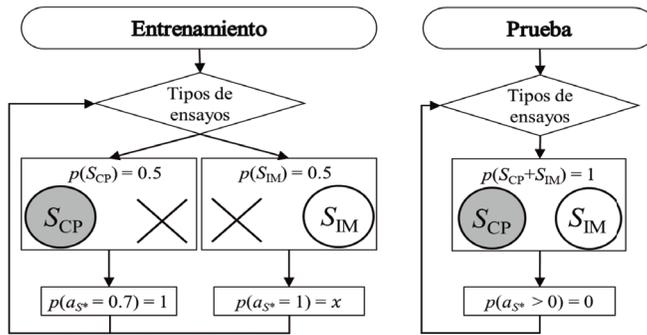


Figura 3. Esquema del procedimiento de Experimento 1. Entrenamiento tuvo 100 ensayos de cada tipo. Prueba tuvo 30 ensayos con regla de aprendizaje deshabilitada. $p(a_{S^*} = 1.0)$ después de S_{IM} fue de x , la variable independiente (ver Tabla 1), donde a_{S^*} denota la activación de S^* , la cual simula la ocurrencia del reforzador.

Análisis de datos

Se registró el nivel de activación de las unidades de salida (M'_1 y M'_2) en el cuarto momento temporal (penúltimo) de los ensayos de prueba de elección. La arquitectura de redes usada (ver Figura 1) consistió de dos vías neurales independientes: $S_{CP} \rightarrow R_{CP}$ (constituída por las conexiones $S'_1 - S''_1$, $S''_1 - M''_1$ y $M''_1 - M'_1$) y $S_{IM} \rightarrow R_{IM}$ (constituída por las conexiones $S'_2 - S''_2$, $S''_2 - M''_2$ y $M''_2 - M'_2$). Por lo tanto, cada señal puede controlar solo una unidad de salida. De este modo, S_{CP} (que activa S'_1)

puede controlar solo R_{CP} (cuyo precursor motor primario es simulado por la activación de M'_1), mientras que S_{IM} (que activa S'_2) solo puede controlar R_{IM} (cuyo precursor motor primario es simulado por la activación de M'_2). Una mayor activación de M'_1 en un ensayo de prueba de elección, entonces, indica una preferencia por S_{CP} , mientras que una mayor activación de M'_2 una preferencia por S_{IM} . Se promediaron los niveles de activación del cuarto momento temporal de los 30 ensayos de prueba para cada unidad de salida (M'_1 y M'_2 ; ver Figura 1) por separado y para cada sujeto. Debido a que la regla de aprendizaje fue deshabilitada durante los ensayos de prueba de elección, la variación por sujeto de la activación durante los ensayos de prueba fue mínima.

No se realizó estadística inferencial con pruebas de hipótesis debido a que los datos son simulaciones por computadora.

Resultados

La visualización de los resultados se muestra en la Figura 4, panel superior.

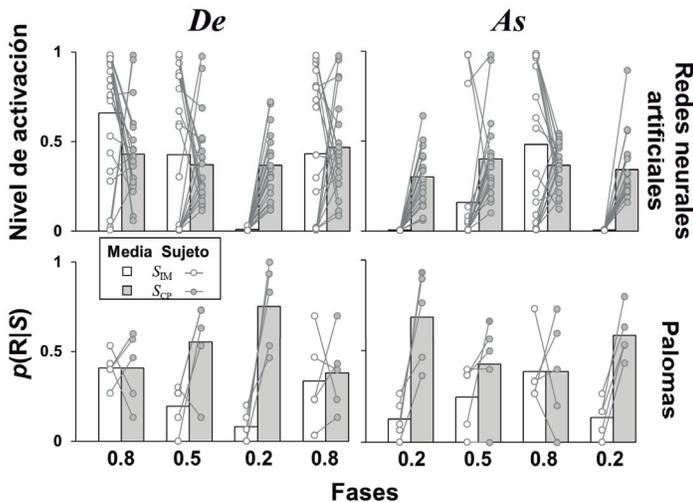


Figura 4. Panel superior: Nivel de activación promedio ($n=20$) durante la prueba por fase y para ambas unidades de salida. Panel superior izquierdo grupo *De*; panel

superior derecho grupo **As**. Las barras son la media del grupo para cada señal por separado, y los puntos son los niveles de activación para cada red. La línea que une los puntos es para indicar que esos dos niveles de activación son de un mismo sujeto. Panel inferior: $p(R|S)$ durante ensayos de elección por fase y para cada señal. Panel inferior izquierdo grupo **De**; Panel inferior derecho grupo **As**. Las barras son la media ($n=5$) del grupo, y los puntos son las $p(R|S)$ para cada paloma. Los puntos unidos por una línea indican que esas dos $p(R|S)$ son de un mismo sujeto.

Como lo muestra la Figura 4, panel superior, independientemente del grupo o el orden de entrenamiento con $p(Sr|S_{IM})$, cuando $p(Sr|S_{IM})=0.2$ se observa una mayor activación de M'_2 que de M'_1 , lo cual indica una preferencia por S_{CP} (ver Figura 4, panel superior, y Tabla 2). En la Fase 3 del grupo **De**, se observa que la activación para S_{CP} es mayor. Asimismo, para el grupo **As** en la primera y cuarta fases con $p(Sr|S_{IM})=0.2$ se observa un mayor nivel de activación para S_{CP} . También se encontró una diferencia creíble en $p(Sr|S_{IM})=0.5$ del grupo **As**, se observa mayor activación para la S_{CP} . Sin embargo, en el grupo **De** no se encuentra diferencia; esto puede sugerir que el orden de presentación puede afectar el nivel de activación. Para la primera condición $p(Sr|S_{IM})=0.8$ del grupo **De** se observa mayor nivel de activación para S_{IM} ; por otro lado, para la segunda $p(Sr|S_{IM})=0.8$, y la tercera fase del grupo **As**, también $p(Sr|S_{IM})=0.8$, pareciera que las activaciones son similares.

Estas diferencias en las activaciones de salida son explicadas por los pesos finales (en el último momento del entrenamiento) resultantes del entrenamiento, cuyos cambios definen el aprendizaje en este tipo de modelo. Los pesos finales resultan y, en esta medida, reflejan los efectos de las condiciones de entrenamiento en la fuerza de las conexiones de las redes. De especial interés aquí son las diferencias entre los pesos finales de las dos vías neurales de la arquitectura usada (ver Figura 2). Como se aclaró, una vía, $S'_1-S''_1-M''_1-M'_1$, fue afectada solo por S_{CP} , mientras que la otra vía, $S'_2-S''_2-M''_2-M'_2$, fue afectada solo por S_{IM} . Estas vías, entonces, mediaron las relaciones entrada→salida, $S_{CP} \rightarrow R_{CP}$ y $S_{IM} \rightarrow R_{IM}$, respectivamente. Por lo tanto, la diferencia entre los pesos finales de estas vías en cada condición es un indicador de cuánto aprendizaje (entendido a la manera conexionista como cambio de pesos) hubo ante cada señal

y bajo cada condición. La Figura 5 muestra la diferencia entre los pesos finales (en el último momento de la fase de entrenamiento) de las vías neurales de la arquitectura usada ($S_{CP} \rightarrow R_{CP}$ y $S_{IM} \rightarrow R_{IM}$) para *De* (panel superior) y *As* (panel inferior) y cada fase. Como lo muestra la Figura 1, la vía $S_{CP} \rightarrow R_{CP}$ estuvo constituida por las conexiones $S'_1-S''_1$, $S'_1-M''_1$, y $M''_1-M'_1$, mientras que la vía $S_{IM} \rightarrow R_{IM}$ estuvo constituida por las conexiones $S''_2-S''_2$, $S''_2-M''_2$, y $M''_2-M'_2$.

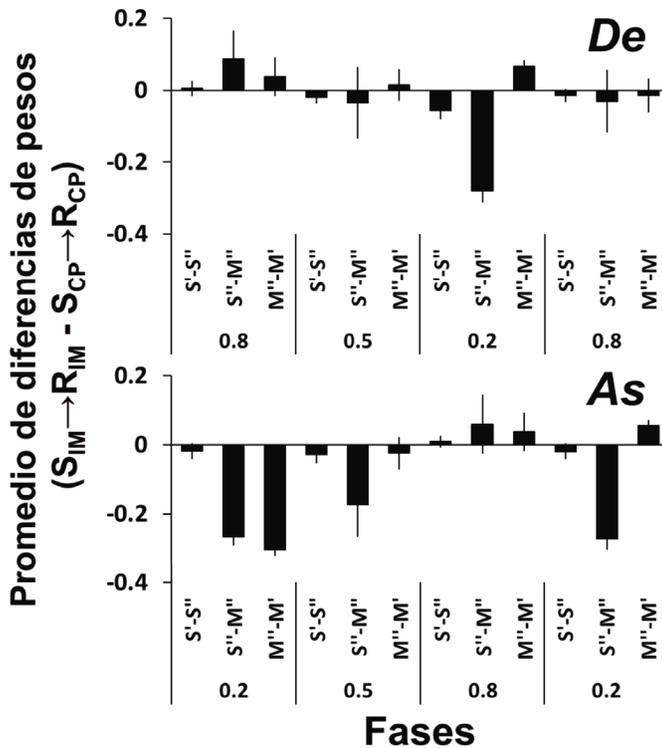


Figura 5. Promedio ($n=20$) de la diferencia de los pesos ($S_{CP} \rightarrow R_{CP} - S_{IM} \rightarrow R_{IM}$) de los tres pares de conexiones en la red: $S'-S''$, $S''-M''$, $M''-M'$. Panel superior grupo *De*, e inferior grupo *As*. Barras negras son la diferencia entre los pesos de las conexiones de $S_{CP} \rightarrow R_{CP}$ y $S_{IM} \rightarrow R_{IM}$. Las barras de error representan el 95% de intervalo de confianza (95% de IC). Si el 95% de IC no toca el cero, es sugerente a una diferencia estadísticamente significativa.

Como puede apreciarse en la Figura 5, las mayores diferencias en los pesos finales entre las dos vías se observaron en $p(\text{Sr}|S_{\text{IM}})=0.2$, lo cual indica que las redes fueron más sensibles a diferenciar las dos opciones cuando la $p(\text{Sr}|S_{\text{IM}})=0.2$. Para la mayoría de las conexiones a lo largo de las condiciones y en ambos grupos, los pesos de $S'-S''$ difirieron mínimamente (Figura 5). Esto sugiere que no aportaron mucho a las diferencias de salida activaciones observadas en la Figura 4. Por otro lado en las conexiones $S''-M''$ se observan mayores pesos en la vía $S_{\text{CP}} \rightarrow R_{\text{CP}}$ que en la vía $S_{\text{IM}} \rightarrow R_{\text{IM}}$ para $p(\text{Sr}|S_{\text{IM}})=0.2$ y $p(\text{Sr}|S_{\text{IM}})=0.5$ del grupo **As**. Por último, las conexiones $M''-M'$ difirieron mucho menos en $p(\text{Sr}|S_{\text{IM}})=0.5$ y 0.8 , con excepción de $p(\text{Sr}|S_{\text{IM}})=0.2$ para el grupo **As**.

Discusión

Estos resultados muestran las predicciones que hace el modelo de redes neurales previamente descrito respecto al efecto de la probabilidad y magnitud de reforzamiento sobre la elección posterior a un entrenamiento en contingencias pavlovianas. Los resultados muestran que las redes fueron sensibles a las distintas contingencias a lo largo de las cuatro fases, y modificaron su preferencia en función de la probabilidad de reforzamiento. Se observaron incrementos en la activación de la unidad de salida (M'_2) controlada por S_{IM} cuando la probabilidad de reforzamiento aumentaba, y viceversa. Como se aprecia en la Figura 4, las activaciones de salida controladas por S_{CP} se mantuvieron estables lo cual sugiere la ausencia de una interacción entre las contingencias CP e IM. Es posible que esto sea efecto de la arquitectura de la red empleada (ver Figura 1), en la cual las vías de procesamiento de las dos señales fueron casi completamente independientes (excepto las conexiones en D).

También se observó un efecto de acarreamiento del aprendizaje, es decir, la activación de salida en la condición $p(\text{Sr}|S_{\text{IM}})=0.5$ es distinto entre los grupos, siendo que lo único diferente es la fase 1.

Se simuló parcialmente reversibilidad de la preferencia (Picker & Poling, 1982) entre la primera y última fase del grupo **As**. Esta reversibilidad se observa en la preferencia por S_{CP} en ambas fases, ya que en el

diseño usado estas fases son condiciones iguales. Por otro lado en el grupo **De** no observamos consistencia en la preferencia entre las fases 1 y 4.

La preferencia que las redes muestran a S_{CP} en la $p(Sr|S_{IM})=0.2$ puede ser explicada por los pesos de las conexiones (ver Figura 5), sobre todo de la conexión $S''_2-M''_2$. Además de haber evaluado los pesos con la regresión múltiple, se observa que cuando la media de activación de ambas unidades de salida es similar i.e., no se acepta la h_1 (ver condiciones con $p(Sr|S_{IM})=0.8$ en Figura 4) y además la diferencia de los pesos es pequeña (ver condiciones con $p(Sr|S_{IM})=0.8$ en Figura 5). Por otro lado, los pesos de las conexiones $S'-S''$, varían poco en función de las condiciones, lo cual sugiere que esas conexiones explican poco la preferencia (y no fueron significativos en la regresión múltiple). Por otro lado, en $p(Sr|S_{IM})=0.2$, donde hubo preferencia por S_{CP} (Figura 4), se observa mayor activación en la vía $S_{CP} \rightarrow R_{CP}$ que en la $S_{IM} \rightarrow R_{IM'}$ así como una fuerza en las conexiones $S''-M''$ y $M''-M'$ (Figura 5). Esto sugiere que la activación de las unidades de salida fue mediada en mayor medida por esas conexiones, sobre todo las conexiones $S''-M''$ (donde se encuentra la interfaz sensorial-motor de las redes).

Puesto que estos resultados fueron obtenidos previos a experimentación con animales, se proponen como predicciones novedosas del modelo sobre algunos posibles efectos de una variante del procedimiento de Picker y Poling (1982, Experimento 2). La predicción principal fue una preferencia por S_{CP} cuando $p(Sr|S_{IM})=0.2$ y preferencia o indiferencia por S_{IM} cuando $p(Sr|S_{IM})=0.8$. Además es posible esperar una reversión de la preferencia en la redeterminación (i.e., Fase 4) de $p(Sr|S_{IM})=0.2$ para **As**. El objetivo del Experimento 2, descrito en la siguiente sección, fue probar las predicciones del modelo con animales. Siguiendo el estudio de Picker y Poling (1982), se utilizó para ello una preparación de automoldeamiento de la respuesta de picar en palomas, aunque se intentó que la simulación fuera análoga al experimento con palomas, como se discutirá al final, hubo diferencias obvias e importantes entre los dos experimentos.

Experimento 2

El objetivo de este experimento fue probar las predicciones del modelo descritas en los resultados del Experimento 1, usando una preparación de automoldeamiento de la respuesta de picar en palomas y un procedimiento análogo al simulado en ese experimento. A partir de esos resultados, se espera que en la condición $p(\text{Sr}|S_{\text{IM}})=0.2$ haya preferencia por S_{CP} y que en la condición $p(\text{Sr}|S_{\text{IM}})=0.8$ haya una preferencia por S_{IM} . También se espera reversibilidad de esta preferencia en la Fase 4 en relación la Fase 1, ya que la condición es igual en estas fases. Se espera esta reversibilidad por dos razones, (1) Picker y Poling (1982) lo observaron, y (2) un grupo de la simulación también lo presentó.

Método

Sujetos

Se utilizaron diez palomas *Columba livia* con peso promedio de 386.2 g y desviación estándar de 47.7 g. Antes de iniciar el experimento, los sujetos fueron privados de alimento hasta alcanzar entre 80 y 85% de su peso *ad libitum*. Cada día, al terminar la sesión experimental, se pesaron todos los sujetos y se alimentaron aquellos que no estuvieran en el porcentaje de privación establecido. Cinco sujetos habían tenido experiencia previa en automoldeamiento, y los demás en tareas de estimación temporal con contingencias operantes. Los primeros cinco sujetos fueron obtenidos de un criadero local en el cual se usaron para competencias de vuelo en distancias largas. Se desconoce la procedencia del resto de las palomas, así como su sexo. Todos los sujetos, por lo menos seis meses antes y durante el experimento, permanecieron en sus mismas cajas y en el criadero de palomas del Centro de Estudios e Investigaciones en Comportamiento de la Universidad de Guadalajara. Los sujetos vivían en habitáculos independientes; ocho de ellos con tamaño de 30(base) x 30(altura) x 35(profundo) cm y el resto en jaulas de 50 x 7 x 40 cm. La temperatura del criadero se mantuvo a 25°C, bajo un ciclo de 12h luz y 12h oscuridad, la luz se encendía a la 07:00.

Materiales e Instrumentos

Cuatro cajas operantes (MED ENV-007, 25.4 cm × 21 cm × 31.8 cm, ancho × alto × largo), dentro de una cámara aislante de sonido (ENV-018V) equipada con un ventilador (VF80A11- AC 115 v). El panel frontal de las cuatro cajas operantes se compone de tres subpaneles. Cada subpanel tenía una tecla de 2.5 cm de diámetro (ENV-123AM) a 20.5 cm del piso de rejilla de barras. Para este experimento solo utilizamos las teclas laterales. Debajo de la tecla del subpanel central estaba el comedero (ENV-205M). Este contenía pellets compactados Pichonina (Purina®); el mismo alimento que comían las palomas en sus jaulas. En el mismo cuarto que las cajas experimentales, estaba la computadora con el programa MED-PC IV, utilizado para programar la presentación de eventos y registrar los datos.

Procedimiento

Los sujetos fueron aleatorizados por bloques (se hace esto para que no todas las palomas con historia de reforzamiento similar queden en el mismo grupo) para conformar dos grupos de cinco sujetos. Estos grupos fueron etiquetados como en el Experimento 1 (**De** y **As**). Se usó la preparación de automoldeamiento, con dos colores como Estímulo Condicionado, que eran seguidos por 4 u 8 segundos de acceso al comedero (Estímulo Incondicionado) dependiendo de si el estímulo era certero poco (CP) o incierto mucho (IM). Se usó el mismo diseño que el Experimento 1 (ABCA). Cada fase duró 18 sesiones, 15 de entrenamiento y tres de prueba. Se llevó a cabo una sesión por día, siete días a la semana. Cada sesión de entrenamiento (ver Figura 6, Entrenamiento) consistió en la presentación aleatorizada de dos tipos de ensayos, cada uno presentado 20 veces.

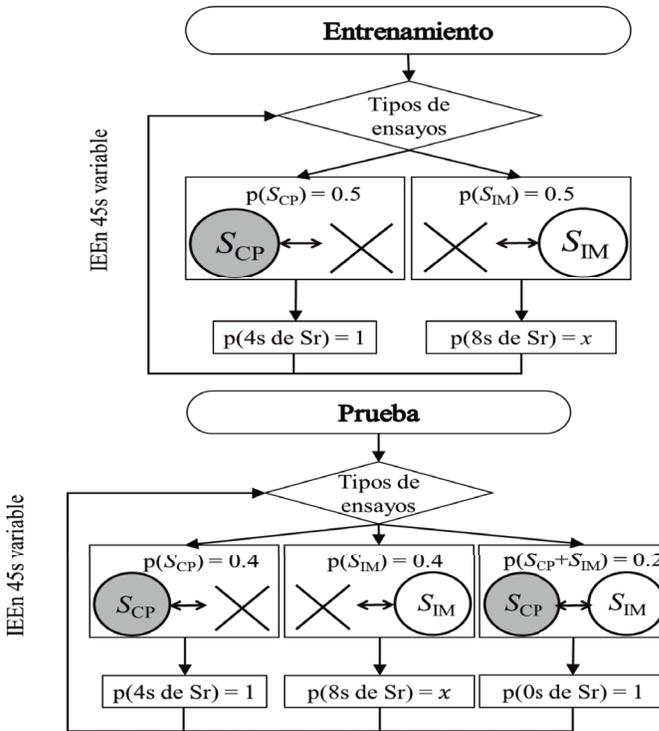


Figura 6. Esquema del procedimiento de las dos etapas durante una fase experimental. Etapa de Entrenamiento tuvo una duración de 15 días por fase. Etapa de Prueba tuvo una duración de 3 días por fase. La $p(8s \text{ de Sr})$ después de S_{IM} es de x , este valor es la variable independiente (ver Tabla 1). El signo \leftrightarrow representa que la ubicación derecha-izquierda de las teclas fue aleatorizada entre los ensayos. Sr = acceso al comedero.

Los tipos de ensayos difirieron en el color del estímulo condicionado; rojo o verde, y en la contingencia que cada color señalaba i.e., CP (Certo-Poco) e IM (Incierto-Mucho). La relación color-contingencia se mantuvo a lo largo del experimento para todos los sujetos, pero fue contrabalanceada, de tal modo que, para la mitad de los sujetos, por grupo, S_{IM} era verde y S_{CP} roja, y viceversa. En cada ensayo la señal se presentó por 6 s, aleatoriamente en la tecla derecha o izquierda, de forma equitativa seguida por el reforzador correspondiente, dependiendo de si

era S_{CP} (señal de la contingencia CP) o S_{IM} (señal de la contingencia IM). Si era S_{CP} , al terminar el ensayo (6 segundos) se dio acceso por 4 s. Si era S_{IM} el acceso fue de 8 s, dependiendo de $p(Sr|S_{IM})$ (ver Tabla 1).

Después del entrenamiento, siguieron tres sesiones de prueba (ver Figura 6, Prueba), las cuales tenían 20 ensayos de S_{CP} y 20 de S_{IM} , además de diez ensayos de elección donde se presentaban S_{CP} y S_{IM} simultáneamente durante 6 s sin reforzamiento. El IEEEn fue variable con media de 45 s con cuatro escalones (i.e., cada 10 ensayos se cumple la media de 45 s). El rango de los IEEEn fue entre 2 y 148 s, siguiendo una distribución exponencial negativa según Fleshler y Hoffman (1962).

Análisis de datos

Para analizar los datos se utilizó un modelo general linear mixto (MGLM) por medio de la función *lmer()* del paquete *lme4* (Bates, & Machler, 2014). El resultado del MGLM se pasó por la función *anova()* de R studio base con el fin de obtener valores F y p y determinar que factor tiene significancia estadística; para esto es necesario el paquete *lmerTest* (Kuznetsova, et al., 2015). Los factores utilizados como efectos fijos fueron: probabilidad de respuesta ($p(R|S_{IM})$ y $p(R|S_{CP})$), condición (0.2, 0.5 y 0.8), grupo (**De** y **As**). Como efecto aleatorio se utilizó el factor sujeto, donde cada sujeto tuvo un intercepto distinto. Por último se utilizó como covariable el factor fase, con el fin de determinar si no hubo un efecto dependiente al orden de las fases. Los resultados del ANOVA del MGLM se escribieron de forma automática (solo se tuvieron que traducir por los autores) con la función *analyze()* del paquete *psycho* (Makowski, 2018).

Por otro lado, para pruebas secundarias (i.e., a partir de la Figura 7). Cuando se comparó dos medias se utilizó una t de student a dos colas para muestras pareadas, y si no se cumplió normalidad y homocedasticidad se utilizó la prueba de los rangos con signo de Wilcoxon. También se llevaron a cabo ANOVAs de una vía para comparar más de dos medias, para esto se probó normalidad y homocedasticidad, cuando no se cumplió algún criterio, se usó una prueba Kruskal Wallis. Para

evaluar normalidad se utilizó Shapiro Wilk, la cual tiene evidencia de ser la más confiable que otras pruebas (Yap y Sim, 2011). Para homocedasticidad se utilizó la prueba de Levene.

Todos los análisis estadísticos fueron realizados en R (R Core Team, 2017). El nivel de significancia estadística fue $\alpha = 0.05$ para todas las pruebas.

Resultados

Para obtener $p(R|S)$, se dividió la suma de los ensayos con la primera respuesta dado S entre el número total de ensayos de prueba, i.e., 30. De modo que cada ensayo podría tener: (1) una respuesta a S_{IM} , (2) una respuesta a S_{CP} o (3) ninguna respuesta.

En la Tabla 2 se muestran los resultados del ANOVA aplicado al modelo lineal general mixto. Escrito con *psycho*: El efecto de la $p(R|S)$ es significativo ($F(1, 59) = 57.13, p < .001$). El efecto de condición no es significativo ($F(2, 59) = 0.56, p > .1$). El efecto de grupo no es significativo ($F(1, 8) = 0.07, p > .1$). El efecto de fases no es significativo ($F(2, 59) = 0.82, p > .1$). La interacción entre la $p(R|S)$ y condición es significativa ($F(2, 59) = 21.02, p < .001$). La interacción entre $p(R|S)$ y grupo no es significativa ($F(1, 59) = 2.57, p > .1$). La interacción entre condición y grupo no es significativa ($F(1, 59) = 0.07, p > .1$). La interacción entre $p(R|S)$, condición y grupo no es significativa ($F(2, 59) = 0.45, p > .1$).

Tabla 2. Resultados de ANOVA sobre el modelo general lineal mixto

Efecto	gl	gl residuales	Suma de cuadrados	Media cuadrada	F	p
$p(R S)$	1	59	1.520	1.520	57.134	< .001***
Condición	2	59	0.030	0.015	0.565	> .1
Grupo	1	8.382	0.002	0.002	0.073	> .1
Fase	2	59	0.044	0.022	0.818	> .1

Efecto	gl	gl residuales	Suma de cuadrados	Media cuadrada	F	p
$p(R S) \times$ Condición	2	59	1.118	0.559	21.024	< .001***
$p(R S) \times$ Grupo	1	59	0.068	0.068	2.574	> .1
Condición \times Grupo	1	59	0.002	0.002	0.069	> .1
$p(R S) \times$ Condición \times Grupo	2	59	0.024	0.012	0.445	> .1

Nota: gl = grados de libertad; $p(\text{respuesta})$ = probabilidad de respuesta.

Similar a los resultados del Experimento 1, en la condición $p(R|S_{IM})=0.2$ durante los ensayos de elección, el valor de $p(R|S_{CP})$ fue mayor que $p(R|S_{IM})$, es decir preferencia por S_{CP} (ver Figura 4, panel inferior). Esta preferencia se observó cuando $p(R|S_{IM})=0.2$ para ambos grupos. Ambos grupos en $p(R|S_{IM})=0.5$, parecen mostrar preferencia por S_{CP} . Por último, en $p(R|S_{IM})=0.8$, las medias de las $p(R|S)$ parecen iguales, lo cual sugiere una indiferencia. En los datos individuales $p(R|S_{IM})=0.8$ en la Figura 4, panel inferior, se observa que, aunque las medias de los grupos son similares, algunos sujetos prefieren S_{IM} y otros S_{CP} . Esta heterogeneidad de preferencia dentro del grupo no se observa en $p(R|S_{IM})=0.2$ ni 0.5 ; con excepción de un sujeto en **De** y dos en **As**. Las distribuciones de respuestas dentro de cada ensayo para ambas señales se muestran en la Figura 7 y revelan que conforme el ensayo transcurre el número promedio de respuestas aumenta. Para las tres condiciones se observa una distribución monótonica y creciente similar. También se observó que en $p(R|S_{IM})=0.8$ no hay una distribución de respuestas diferenciada entre S_{CP} e S_{IM} . Sin embargo, en $p(R|S_{IM})=0.2$, se observa mayor número de respuestas promedio y pendiente para S_{CP} que para S_{IM} .

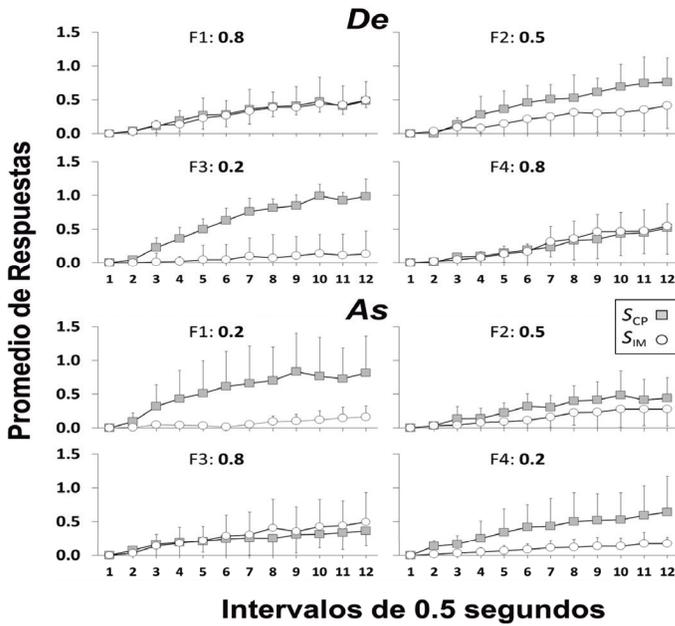


Figura 7. Distribuciones de respuesta promedio ($n=5$) en ensayos de elección tanto para S_{CP} como para S_{IM} . Los cuatro gráficos superiores son del grupo *De*, y los cuatro inferiores son del grupo *As*. Cada punto es la media del grupo en cada intervalo de clase, y la barra de error representa la DE. El orden de las fases y la condición se muestra en la parte superior de cada gráfico (e.g. F1: 0.2, que quiere decir, Fase 1 condición 0.2).

Para comparar las pendientes de las distribuciones de respuestas (Figura 7) entre S_{CP} y S_{IM} , se llevaron a cabo regresiones lineales por cada sujeto en cada fase. Se promediaron las pendientes de los todos los sujetos por condiciones, ver Figura 8. Para la condición 0.2 se obtuvo una diferencia significativa; $W(14) = 195, p < 0.05$, sin embargo para las condiciones 0.5 y 0.8 las diferencias en las pendientes no fueron significativas; $W(9) = 73, p > 0.05$; y $W(14) = 93, p > 0.05$.

Discusión

Lo que se observó en los ensayos de elección fue consistente con la reversibilidad de la preferencia mostrada por Picker y Poling (1982). Hay similitud entre la primera y cuarta fase para ambos grupos, Figura 4, panel inferior. Esto indica que no hubo acarreo de los efectos de condiciones anteriores sobre las posteriores a lo largo del experimento. Esta es diferente al Experimento 1, donde se observó acarreo de la Fase 1 a la Fase 2. Las palomas regresan a la misma preferencia en Fase 1 y 4; independientemente de las contingencias en fases intermedias. Se observó una distribución de respuestas durante los ensayos creciente y monotónica. La pendiente de la distribución producida por S_{CP} fue mayor que la producida por S_{IM} , aunque con significancia estadística solo para $p(R|S_{IM})=0.2$ (ver Figura 8). A diferencia de Picker y Poling (1982, Fig. 2), que muestran que en la señal 100% seguida de alimento, la mayor cantidad de respuestas se acumula en la primera mitad del ensayo, y para la otra señal (50%), la mayor cantidad de respuesta se acumula al final de la duración del ensayo.

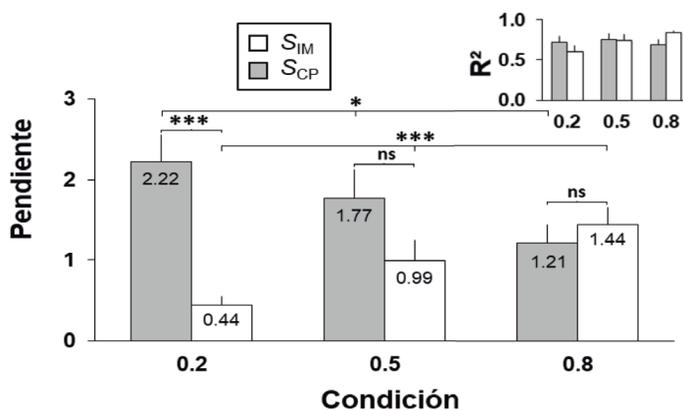


Figura 8. Promedios (condiciones 0.8 y 0.2: $n = 15$, y condición 0.5: $n = 10$) de pendientes obtenidos por regresiones lineales de las distribuciones de respuesta de cada sujeto en cada fase para las dos señales. Se compararon las medias de S_{IM} y S_{CP} condición con pruebas Wilcoxon. Las medias por señales entre las condiciones se compararon con Kruskal-Wallis. El * significa un valor $p < 0.05$; *** significa un valor $p < 0.001$; y ns denota “no significativo”.

Interacción entre contingencias

En la Figura 7 se observa que la $p(R|S_{CP})$ varía entre las condiciones, aunque esa contingencia no haya cambiado a lo largo del experimento. Esta es distinta al Experimento 1 en el cual el nivel de activación para S_{CP} fue estable a lo largo del experimento (Figura 4, panel superior). En las palomas, se observó que cuando aumenta $p(Sr|S_{IM})$ incrementa $p(R|S_{IM})$; pero también al aumentar $p(Sr|S_{IM})$ disminuye la $p(R|S_{CP})$. Lo interesante aquí es que a diferencia de $p(Sr|S_{IM})$, la $p(Sr|S_{CP})$ fue estable a lo largo de todo el experimento. Esto podría sugerir que las probabilidades de responder a las señales no son independientes entre sí. A esto se le llamará una interacción entre contingencias.

Para evaluar esta interacción, se combinaron los sujetos de los grupos por condiciones, solo de las primeras tres fases (ver Figura 9). Se llevaron a cabo comparaciones de medias de respuesta ante cada señal entre las tres condiciones y se encontró un efecto significativo para la $p(R|S_{CP})$ en las palomas; $F(2,27) = 6.35, p < 0.05$, más no para las redes; $H(2) = 2.39, p > 0.05$. Esto apoya la idea de que las activaciones de salida ante S_{CP} y S_{IM} son independientes en el caso de las redes, pero no en las palomas. Por otro lado, la $p(R|S_{CP})$ depende de la condición en las palomas. Las medias $p(R|S_{IM})$ y el nivel de activación de S_{IM} entre las condiciones muestran suficiente evidencia para rechazar la hipótesis de que dichas medias sean iguales y esto fue significativo; palomas: $F(2,27) = 6.73, p < 0.05$; redes: $H(2) = 62.20, p < 0.05$, esto apoya la idea que tanto las respuestas de las redes como las de las palomas a S_{IM} varían en función de la $p(Sr|S_{IM})$.

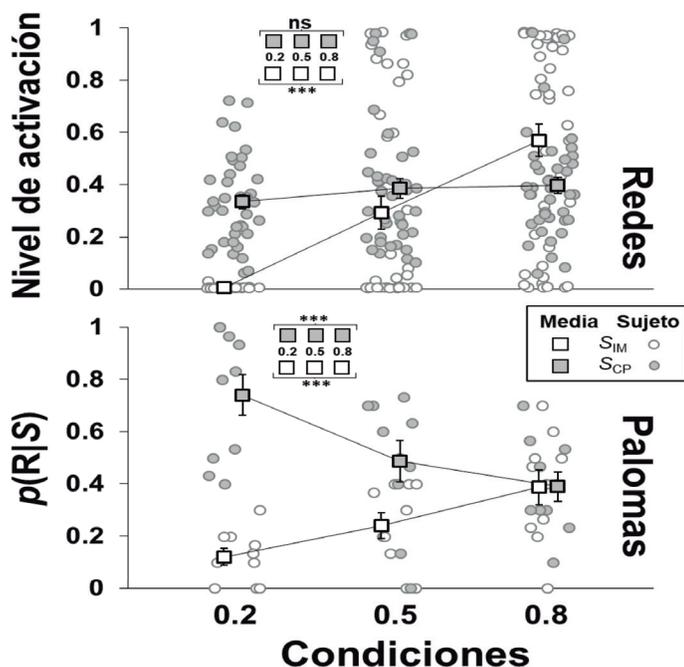


Figura 9. Comparaciones de los resultados de los Experimentos 1 (panel superior; $n=40$ por condición) y 2 (panel inferior; $n=10$ por condición). Para cada condición se juntaron los sujetos de ambos grupos (solo primeras tres fases). Los cuadros representan las medias y las barras el error estándar de la media. Cada círculo representa un sujeto individual. Cuadros y puntos grises corresponden al responder a S_{CP} y los blancos al responder a S_{IM} . En la parte superior de cada panel se muestran las 6 medias con su respectiva condición y se señala la significancia estadística con *** (valor $p < 0.001$) y *ns* (no significancia). Se usaron ANOVAs para las palomas y Kruskal Wallis para las redes.

Discusión general

La comparación entre las predicciones del modelo en el Experimento 1 y los resultados con las palomas en el Experimento 2 son parcialmente consistentes entre sí (ver Figura 4), excepto la interacción entre las contingencias que se observó con las palomas y no en las redes (ver Figura 9). Pese a esta diferencia, la predicción del modelo fue razonablemente acertada. El resultado general principal es que tanto

redes como palomas al haber sido expuestas a contingencias pavlovianas previas a las pruebas de elección, ajustan su responder según la probabilidad y magnitud de reforzamiento, y de maneras semejantes. Los sujetos (de ambos experimentos) incrementan su respuesta o activación a la S_{IM} cuando aumenta la probabilidad de esa alternativa.

Por supuesto, como ya se anticipó, hay diferencias sustanciales, alguna obvias, entre los dos experimentos que podrían explicar algunas de las diferencias observadas en los resultados entre las redes neurales y las palomas. Empezando con lo obvio, ciertamente una red neural artificial no es lo mismo que una paloma. Una red artificial neural es una abstracción considerable de ciertas estructuras cerebrales gruesas que tienen en común varias especies de vertebrados, incluyendo las palomas. Entonces, el objetivo de las redes usadas en el Experimento 1 no fue simular el cerebro de una paloma como tal, sino más bien algunas características gruesas que siguen ciertos principios básicos neuroanatómicos comunes en todos vertebrados, incluyendo palomas. El Experimento 2, entonces, pudo muy bien haberse realizado con ratas y la explicación del modelo hubiera sido la misma, si en efecto se hubieran obtenido resultados comparables a los que se obtuvieron con las palomas. Si no, ello sugeriría, según el modelo, diferencias en ciertos aspectos de la organización neuroanatómica gruesa de ratas y palomas, lo cual representaría un estudio comparativo. Una replicación sistemática sugerida por este estudio, entonces, sería repetirlo con ratas.

Además, la conectividad de las redes fue casi completamente desconectada, otra abstracción adoptada como simplificación estratégica con fines de teorización. Excepto por la unidad D , la cual recibió conexiones de todas las unidades M'' (y, en esta medida, podía ser afectada por ambas señales mediante sus respectivas unidades de entrada y S''), en general las redes tenían vías autónomas dedicadas a cada señal. Para ello se supuso que una señal activaba solo una vía y , de ese modo, realizaba una sola relación entrada-salida. Sin embargo, este tipo de conectividad es mucho más simple que el cerebro de una paloma. De nuevo, el presente estudio es en parte un ejercicio en modelamiento matemático de sistemas complejos, el cual inevitablemente involucra

una simplificación sustancial de aquello que se pretende modelar. Pero el objetivo de tal simplificación es la claridad y la precisión, medios que facilitan evaluar cuán consistentes son los modelos con la evidencia.

Por razones semejantes, es difícil equiparar los reforzadores entre redes y palomas. No es obvio cuánto vale un nivel de activación de 0.7 en tiempo de acceso al comedero, la variable que usualmente define la magnitud del reforzamiento en la preparación de automoldeamiento. Esta puede ser la razón por la cual en la condición de 0.8 en las palomas se encontró indiferencia en comparación de las redes que mostraron preferencia por S_{IM} . Eso se puede deber a los parámetros elegidos para el experimento con las palomas (4 s como poca recompensa) o las redes. Quizás si se hubiera usado como recompensa CP un tiempo de acceso más corto y/o como recompensa IM un acceso más prolongado al comedero, se hubiera observado una preferencia por S_{IM} en la condición 0.8 con las palomas, porque la diferencia entre 4 y 8 segundos es el doble y Picker y Poling (1982) encontraron reversibilidad de preferencia cuando las diferencias entre señales eran de 3:1 o 6:1 (12 vs 4 s o 12 vs 2 s). Otra posible limitación, en el caso de las palomas fue que quizás éstas no experimentaban 8 s como el doble de reforzamiento que 4 s. Puede que durante 8 s las palomas no siempre estén comiendo, y esto disminuya el valor reforzante de S_{IM} .

Interacción entre contingencias

A diferencia de los resultados del Experimento 1 con las redes, las palomas mostraron un efecto de interacción entre las dos señales asociadas a las distintas contingencias pavlovianas. Aunque la contingencia CP fue estable a través de las condiciones, la $p(R|S_{CP})$ no lo fue, a diferencia de las redes. Para las palomas, la $p(R|S_{CP})$ y $p(R|S_{IM})$ se correlacionan inversamente con -0.276, es decir a mayor $p(R|S_{CP})$, menor $p(R|S_{IM})$ (e.g. condición 0.2) y viceversa (e.g. condición 0.8). Por otro lado, la correlación entre los niveles de activación de salida para S_{CP} y S_{IM} en las redes fue casi nula con un 0.021.

La interacción entre contingencias cambiantes (IM) y contingencias estables (CP) pareciera contraste conductual positivo con contin-

gencias pavlovianas. El contraste conductual es una interacción entre dos o más componentes; se ve cuando la tasa de respuesta aumenta en un componente A debido a que la tasa de reforzamiento de otro componente B disminuye, siendo que la tasa de reforzamiento de A se mantiene estable (Mackintosh, 1974). En las Figuras 4, panel inferior, y Figura 9 se observó que $p(R|S_{CP})$ aumenta cuando la tasa de reforzamiento de S_{IM} baja. Al igual que un procedimiento típico de contraste, en el Experimento 2 las dos señales se presentaron simultáneamente. A esto se le pudiera llamar contraste conductual pavloviano. Hasta donde los autores de este documento saben, este fenómeno no ha sido reportado. Pudiera ser un nuevo hallazgo.

El tipo de interacción aquí descrita se asemeja a la descrita por Reynolds: “Un cambio en el comportamiento durante la presentación de un estímulo, provocado por el cambio del programa asociado con un estímulo diferente, se denomina *interacción*.” (Reynolds, 1961, pp. 57) [Traducido en Google Translate, fiabilidad de Google Translate en Balk, *et al.*, 2013]. Esa misma definición captura razonablemente bien los resultados del Experimento 2. Además, se supone que la interacción solo puede verse entre componentes de programas múltiples (Schwartz & Gamzu, 1977), debido a la intención de establecer claridad conceptual basada en procedimientos; aunque como se acaba de sugerir, podrían explorarse las implicaciones de ampliar la noción a la interacción entre contingencias pavlovianas en la elección automoldeada. Habrá que ver en otras preparaciones pavlovianas si se observa esta interacción.

Implicaciones para la economía conductual

El presente estudio apoya y extiende los resultados básicos de Picker y Poling (1982). La elección, como base de la conducta económica, no solo no está restringida a contingencias operantes. También puede surgir a partir de contingencias pavlovianas y ser sensible a manipulaciones relativamente finas de tales contingencias, al menos en una preparación de automoldeamiento, la cual involucra respuestas emitidas, como el picoteo de la tecla en palomas.

Más específicamente relacionado con el presente estudio, un fenómeno investigado en economía conductual bajo contingencias operantes es el descuento probabilístico, la disminución hiperbólica (Rachlin, *et al.*, 1991; Green & Myerson, 2004) del valor subjetivo de una recompensa en función de la probabilidad de no obtener la recompensa por emitir cierta respuesta (Madden & Bickel, 2010). También es posible que haya un descuento probabilístico pavloviano, definido como una disminución de la respuesta en función de la probabilidad de reforzamiento de una señal, independientemente del responder. Una posible forma obtener esta función sería adaptar a contingencias pavlovianas el procedimiento de ajuste de cantidad (Mazur, 1987; Green, *et al.*, 2010). Este procedimiento consiste en manipular en tiempo real la probabilidad y la magnitud de reforzamiento dentro de una sesión para obtener puntos de indiferencia. En su versión pavloviana, se manipularía la probabilidad de reforzamiento condicional a ciertas señales. Esta posibilidad provee una base para investigar el descuento probabilístico con contingencias pavlovianas, algo que aún no ha sido investigado.

En conclusión, este trabajo presentó una predicción novedosa con el modelo de redes neurales previamente descrito. Dicha predicción, fue corroborada con palomas. Además, señalamos la importancia de las contingencias pavlovianas en comportamiento económico. Donde fenómenos como descuento probabilístico, quizás puedan emerger en situaciones donde exclusivamente existan contingencias pavlovianas. Y, por último, en las palomas observamos (en las redes no, dada la arquitectura utilizada) un fenómeno conductual nuevo, el cual argüimos es contraste conductual pavloviano; en el cual hay una interacción entre el responder ante dos opciones, siendo que una opción siempre fue fija.

Referencias

- Balk, E. M., Chung, M., Chen, M. L., Trikalinos, T. A., & Chang, L. K. W. (2013). Assessing the accuracy of google translate to allow data extraction from trials published in Non-English languages.

- Rockville (MD): *Agency for Healthcare Research and Quality* (US); 2013 Jan. Report No.: 12(13)-EHC145-EF.
- Bates, D., & Machler, M. (2014). Package lme4: Linear Mixed-Effects Models Using Eigen and S4. *Journal of statistical software*.
- Baum, W. M. (1974). On two types of deviation from the matching law: bias and undermatching. *Journal of the Experimental Analysis of Behavior*, 22(1), 231-242. <https://doi.org/10.1901/jeab.1974.22-231>
- Burgos, J. E., & Galeazzi, J. M., (in press). Neural-Network Simulations of a Possible Role of the Hippocampus in Pavlovian Conditioning. En P. A. Gargiulo y H. L. Mesones-Arroyo (Eds.), *Psychiatry and neuroscience update, Vol. IV*. Springer Nature. <https://doi.org/10.1007/978-3-319-17103-6>
- Burgos, J. E. (1997). Evolving artificial neural networks in Pavlovian environments. In J. W. Donahoe & V. P. Dorsel (Eds.), *Advances in psychology*, Vol. 121. Neural-network models of cognition: Biobehavioral foundations (p. 58-79). North-Holland/Elsevier Science Publishers. [https://doi.org/10.1016/S0166-4115\(97\)80090-8](https://doi.org/10.1016/S0166-4115(97)80090-8)
- Burgos, J. E. (2003). Theoretical note: simulating latent inhibition with selection neural networks. *Behavioural processes*, 62(1-3), 183-192. [https://doi.org/10.1016/s0376-6357\(03\)00025-1](https://doi.org/10.1016/s0376-6357(03)00025-1)
- Burgos, J. E. (2005). Theoretical note: the C/T ratio in artificial neural networks. *Behavioural processes*, 69(2), 249-256. <https://doi.org/10.1016/j.beproc.2005.02.008>
- Burgos, J. E., & Murillo-Rodríguez, E. (2007). Neural-network simulations of two context-dependence phenomena. *Behavioural processes*, 75(2), 242-249. <https://doi.org/10.1016/j.beproc.2007.02.003>
- Burgos, J. E. (2007). Autoshaping and Automaintenanc: a neural-network approach. *Journal of the experimental analysis of behavior*, 88(1), 115-130. <https://doi.org/10.1901/jeab.2007.75-04>
- Burgos, J. E., Flores, C., García, Ó., Díaz, C., & Cruz, Y. (2008). A simultaneous procedure facilitates acquisition under an optimal interstimulus interval in artificial neural networks and rats. *Behavioural processes*, 78(2), 302-309. <https://doi.org/10.1016/j.beproc.2008.02.018>

- Burgos, J.E. & García-Leal, Ó. (2015). Autoshaped choice in artificial neural networks: Implications for behavioral economics and neuroeconomics. *Behavioural processes*, 114:62-71. <https://doi.org/10.1016/j.beproc.2015.01.010>
- Burgos, J. E., & Donahoe, J. W. (2016). Unified principle of reinforcement in a neural-network model: Reply to NT Calvin and JJ McDowell. *Behavioural processes*, 126, 46-54. <https://doi.org/10.1016/j.beproc.2016.03.003>
- Burns, R., Burgos, J. E., & Donahoe, J. W. (2011). Pavlovian conditioning: Pigeon nictitating membrane. *Behavioural processes*, 86(1), 102-108. <https://doi.org/10.1016/j.beproc.2010.10.004>
- Brown, P. L., & Jenkins, H. M. (1968). Auto-shaping of the pigeon's key-peck. *Journal of the experimental analysis of behavior*, 11(1), 1-8. <https://doi.org/10.1901/jeab.1968.11-1>
- Calvin, N. T., & McDowell, J. J. (2015). Unified-theory-of-reinforcement neural networks do not simulate the blocking effect. *Behavioural processes*, 120, 54-63. <https://doi.org/10.1016/j.beproc.2015.08.008>
- Donahoe, J.W., Burgos, J.E., & Palmer, D.C., (1993). A selectionist approach to reinforcement. *Journal of the Experimental Analysis of Behavior*, 60:17-40. <https://doi.org/10.1901/jeab.1993.60-17>
- Donahoe, J. W., Crowley, M. A., Millard, W. J. & Stickney, K. A. (1982). A unified principle of reinforcement. En Commons, M.L., Herrnstein y R.J., Rachlin, H. (Eds.), *Quantitative Analyses of Behavior: Vol. 2: Matching and Maximizing Accounts* (pp. 493-521). Cambridge, MA: Balinger.
- Donahoe, J. W., & Burgos, J. E. (1999). Timing without a timer. *Journal of the experimental analysis of behavior*, 71(2), 257-263. <https://doi.org/10.1901/jeab.1999.71-257>
- Donahoe, J. W., & Burgos, J. E. (2000). Behavior analysis and reevaluation. *Journal of the experimental analysis of behavior*, 74(3), 331-346. <https://doi.org/10.1901/jeab.2000.74-331>
- Fleshler, M., & Hoffman, H. S. (1962). A progression for generating variable-interval schedules. *Journal of the Experimental Analysis of Behavior*, 5(4), 529. <https://doi.org/10.1901/jeab.1962.5-529>

- Gormezano, I., & Kehoe, E. J. (1975). Classical conditioning: Some methodological-conceptual issues. En Estes, W.K. (Ed). *Handbook of learning and cognitive processes*, 2, 143-179.
- Green, L., & Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin*, 130(5), 769. <https://doi.org/10.1037/0033-2909.130.5.769>
- Green, L., Myerson, J., & Calvert, A. L. (2010). Pigeons' Discounting of Probabilistic and Delayed Reinforcers. *Journal of the Experimental Analysis of Behavior*, 94(2), 113-123. <https://doi.org/10.1901/jeab.2010.94-113>
- Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, 4(3), 267-272. <https://doi.org/10.1901/jeab.1961.4-267>
- Ikemoto, S. (2007). Dopamine reward circuitry: two projection systems from the ventral midbrain to the nucleus accumbens-olfactory tubercle complex. *Brain research reviews*, 56(1), 27-78. <https://doi.org/10.1016/j.brainresrev.2007.05.004>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). Package 'lmerTest'. *R package version*, 2(0).
- Madden, G. J., & Bickel, W. K. (2010). Impulsivity: The behavioral and neurological science of discounting. *American Psychological Association*. ISBN: 978-1-4338-0477-9
- Mackintosh, N. J. (1974). *The Psychology of Animal Learning*. New York: Academic Press
- Makowski, D. (2018). The psycho Package: an Efficient and Publishing-Oriented Workflow for Psychological Science. *J. Open Source Software*, 3(22), 470. <https://doi.org/10.21105/joss.00470>
- Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. En Commons, M. L., Mazur, J. E., Nevin, J. A., & Rachlin, H. (Eds.), *Quantitative analyses of behavior: Vol. 5. The effect of delay and of intervening events on reinforcement value* (pp. 55-73). Hillsdale, NJ: Erlbaum

- Picker, M., & Poling, A. (1982). Choice as a dependent measure in autoshaping: Sensitivity to frequency and duration of food presentation. *Journal of the Experimental Analysis of Behavior*, 37(3), 393-406. <https://doi.org/10.1901/jeab.1982.37-393>
- R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior*, 55(2), 233-244. <https://doi.org/10.1901/jeab.1991.55-233>
- Reynolds, G. S. (1961). Behavioral contrast. *Journal of the Experimental Analysis of Behavior*, 4(1), 57-71. <https://doi.org/10.1901/jeab.1961.4-57>
- Richmond, J., & Colombo, M. (2002). Hippocampal lesions, contextual retrieval, and autoshaping in pigeons. *Brain Research*, 928(1-2), 60-68. [https://doi.org/10.1016/s0006-8993\(01\)03355-8](https://doi.org/10.1016/s0006-8993(01)03355-8)
- Sánchez, J. M., Galeazzi, J. M., & Burgos, J. E. (2010). Some structural determinants of Pavlovian conditioning in artificial neural networks. *Behavioural processes*, 84(1), 526-535. <https://doi.org/10.1016/j.beproc.2010.01.018>
- Schwartz, B. Gamzu, E. (1977). Pavlovian control of operant behavior: An analysis of autoshaping and its implications for operant conditioning. WK Honig & JE R. Staddon (Eds), *Handbook of operant behavior*, 53-97. ISBN: 0133805352
- Williams, D.R. and Williams, H. (1969), Auto-maintenance in the pigeon: sustained pecking despite contingent non-reinforcement. *Journal of the Experimental Analysis of Behavior*, 12: 511-520. <https://doi.org/10.1901/jeab.1969.12-511>
- Yap, B. W., & Sim, C. H. (2011). Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation*, 81(12), 2141-2155. <https://doi.org/10.1080/00949655.2010.520163>
- Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience*, 7(6), 464. <https://doi.org/10.1038/nrn1919>