

## **ASIGNATURA: PSICOLOGÍA DEL APRENDIZAJE**

**CURSO: 2023/2024.**

### **ACTIVIDADES FORMATIVAS COMPLEMENTARIAS OBLIGATORIAS (AFCO)**

A sugerencia del equipo docente se realizarán una serie de actividades formativas que serán complementarias al estudio del texto base. Se irán desarrollando a lo largo del curso y serán supervisadas por los Profesores Tutores de los Centros Asociados y evaluadas en el propio examen. Estas actividades están organizadas en dos bloques y a continuación se presenta una breve descripción de estas.

#### **PRIMER BLOQUE: ANÁLISIS DE VÍDEOS DIDÁCTICOS**

Los estudiantes tendrán a su disposición en la página web de la asignatura una serie de vídeos didácticos que podrán utilizar como ayuda al estudio de la asignatura. Los Profesores Tutores en los Centros Asociados, o en su espacio de tutoría del curso virtual, podrán orientar a los estudiantes sobre los conocimientos que han adquirido a partir de la visualización de estos vídeos. El equipo docente facilitará, a través del curso virtual, los cuestionarios mediante los cuales se realizará el seguimiento del progreso de los estudiantes. Durante este curso se evaluará el vídeo: "Programas de reforzamiento".

##### **1. Explique en qué consisten los programas de razón fija y variable, y las diferencias en el registro acumulativo.**

Los programas de razón requieren que, para obtener el reforzador, el sujeto deba emitir un número previamente determinado de respuestas. Por ejemplo, en el vídeo, a partir del minuto 9:13, se muestra como una rata debe presionar la palanca tres veces para obtener comida. Los programas de razón se pueden dividir en dos tipos:

- a) **Programas de Razón Fija (RF):** En estos programas, el número de respuestas requeridas para obtener el reforzador es siempre el mismo. Si el número de repuesta requeridas es elevado, la rata tiende a realizar una pausa tras la obtención del reforzador. En cambio, cuando el requisito es bajo, esta pausa es menos notable. La duración de la pausa post-reforzamiento está en función directa del requisito del programa de razón.
- b) **Programas de Razón Variable (RV):** En estos programas, el número de respuestas necesarias para obtener el reforzador es diferente en cada ensayo, de acuerdo con un número medio de respuestas fijado. Por ejemplo, en un programa RV10, en algunos ensayos el sujeto obtendrá el reforzador tras 13 presiones de palanca, en otro tras 7, o en otros tras 10.

Los registros acumulativos registran diferencias entre ambos tipos de programas. En programas de RF se observa una pausa post-reforzamiento, especialmente cuando la RF requiere un número elevado de respuestas. Esta pausa post-reforzamiento implica que no se registran respuestas una vez entregado el reforzador y antes de iniciar de nuevo la razón requerida para obtener el siguiente reforzador. En cambio, en el registro

acumulativo de los programas de RV no se observa esa pausa post-reforzamiento. En su lugar, se evidencia un patrón de respuestas rápidas, constantes y estables.

## **2. Explique en qué consiste y cómo se interpreta un registro acumulativo.**

El registro acumulativo consiste en un sistema que nos permite registrar la conducta del sujeto en función del tiempo, ofreciéndonos así una representación visual de cuándo y con qué frecuencia responde el sujeto durante la sesión. Este registro se presenta comúnmente en forma de gráfica, donde el eje vertical representa el número de respuestas del sujeto, mientras que el eje horizontal representa el paso del tiempo.

Para facilitar la explicación de cómo se interpreta, podemos utilizar el ejemplo sobre el registro de presión de palanca que aparece en el vídeo en el minuto 4:05. En la gráfica, se observa una curva acumulativa continua que refleja la tasa de presión de palanca en cualquier momento del estudio. Cuando esta curva está en crecimiento, indica que la rata está presionando la palanca. Por el contrario, cuando se mantiene estable, indica que el sujeto no está emitiendo ninguna respuesta de presión de palanca.

El registro acumulativo no sólo nos permite representar exclusivamente las respuestas que estamos midiendo, sino que también posibilita registrar otros elementos ambientales que ocurren durante la sesión, como la aparición de estímulos discriminativos o la entrega de reforzadores.

## **3. Indique las diferencias que existen entre los programas de tiempo, de intervalo y de reforzamiento diferencial de tasa baja. Puede crear una tabla para compararlos.**

En estos tres programas de reforzamiento se establece un intervalo de tiempo que determina cuándo y cómo se va a entregar el reforzador, pero es la configuración del intervalo lo que hace que estos programas sean diferentes. Como consecuencia observaremos diferencias en las curvas mostradas en los registros acumulativos.

En primer lugar, el programa de tiempo resulta el más sencillo. En este programa de reforzamiento, el reforzador se entrega pasado un cierto intervalo de tiempo, independientemente del comportamiento del sujeto, es decir, no es necesaria la aparición de una respuesta operante para que el reforzador se entregue. Por ejemplo, en el minuto 7:56 del video se puede observar que en un experimento con ratas se establece un programa de tiempo fijo (TF) 30 segundos. En este caso, la rata recibirá una bolita de comida cada 30 segundos, independientemente de la respuesta que esté emitiendo durante ese intervalo. En el video podemos observar que durante ese intervalo la rata desarrolla multitud de comportamientos repetitivos que no son contingentes con la entrega del reforzador como puede ser la bebida inducida por programa, acicalarse, husmear o acercarse al comedero cuando faltan pocos segundos para que se entregue la bolita de comida. Pero de nuevo, no es necesario que se emitan esos comportamientos para que el reforzador se entregue.

En segundo lugar, los programas de intervalo establecen que debe transcurrir un tiempo desde la entrega del último reforzador para que la respuesta operante del sujeto vuelva a ser reforzada. Este programa implica que se reforzará la primera respuesta que emita el sujeto pasado dicho intervalo. Aquí, la aparición de una respuesta operante es necesaria, a diferencia de los programas de TF. En el minuto 5:30 del vídeo podemos

observar un experimento con ratas en el que se establece un intervalo fijo (IF) 38 segundos con presión de palanca como conducta operante. En este ejemplo podemos apreciar que el reforzador se entrega cuando la rata presiona la palanca una vez una vez han pasado 38 segundos desde el último reforzador que se le entregó. Hay que resaltar que, si la rata emite alguna presión de palanca antes de que transcurran esos 38 segundos, no ocurre nada.

Por último, los programas de reforzamiento diferencial de tasa baja establecen que debe pasar un intervalo de tiempo desde la última respuesta emitida por el sujeto antes de volver a responder. Además, en estos programas se establece una contingencia de castigo si el sujeto responde antes de que ese tiempo transcurra, ya que el intervalo se reinicia. Al igual que en los programas de intervalo, se requiere la emisión de una respuesta operante. En el minuto 12:22 del vídeo se puede ver un experimento con ratas en el que se establece un programa de reforzamiento diferencial de tasa baja (RDB) de 7 segundos usando como conducta operante la presión de palanca. En este ejemplo la rata debe esperar 7 segundos desde la última vez que presionó para que esta vuelva a ser reforzada. Si la rata presiona a los 5 segundos no se entrega el reforzador y el intervalo se reinicia, por lo que la rata debe volver a esperar otros 7 segundos desde esa última presión no reforzada.

En resumen, en los programas de tiempo fijo, la rata solo debe esperar sin necesidad de emitir una respuesta operante. En los programas de intervalo fijo, la primera respuesta operante después de un intervalo específico desde el reforzador anterior será reforzada. Por último, en los programas de reforzamiento diferencial de tasa baja, la respuesta será reforzada si transcurre un tiempo determinado desde la última respuesta emitida.

	<b>Conducta Operante</b>	<b>Inicio del intervalo</b>	<b>Entrega del reforzador</b>
<b>Tiempo</b>	No	Entrega del último reforzador	Cuando transcurra el tiempo establecido
<b>Intervalo</b>	Sí	Entrega del último reforzador	Primera respuesta que se emita pasado el intervalo establecido
<b>Reforzamiento diferencial de tasa baja</b>	Sí	Emisión de la última respuesta	Respuesta operante una vez transcurrido el intervalo

**4. Ponga un ejemplo de la vida real de cada uno de los siguientes programas de reforzamiento: Razón Fija, Razón Variable, Intervalo Fijo, Intervalo Variable.**

Razón fija: Un ejemplo de razón fija se podría encontrar en las personas que se dedican a la venta de algún producto y se llevan una comisión por cada venta. Imaginemos que un comercial de coches se lleva una comisión por cada 5 coches que vende. En este ejemplo, la comisión estaría actuando como estímulo reforzador, la venta de un coche sería la respuesta operante, y el sujeto estaría sometido a un programa de razón fija 5 ya que tiene que vender 5 coches (emitir 5 veces la respuesta operante) para poder recibir la comisión (reforzador).

Razón variable: Se puede observar un ejemplo de programa de razón variable en aplicaciones de reservas de alojamiento. Estas aplicaciones suelen ofrecer incentivos en forma de bonificaciones después de alcanzar un número específico de reservas. En un ejemplo hipotético, la cantidad de reservas (conducta operante) que un usuario debe completar para obtener una bonificación (reforzador) varía, a veces, la bonificación se otorga cada 5 reservas, mientras que en otras ocasiones se concede cada 7 o cada 3. En este caso, la conducta de reservar para obtener bonificaciones quedaría bajo un programa de razón variable 5.

Intervalo fijo: Un ejemplo de intervalo fijo se puede encontrar en los trabajos en los que los empleados tienen que fichar a la entrada y a la salida. Imaginemos que en nuestra empresa tenemos una jornada laboral de 7 horas y tenemos que pasar nuestra tarjeta de empleado por un escáner para que dé comienzo la jornada laboral y se dé acceso a los ordenadores. Para poder dar por finalizada la jornada laboral debo pasar mi tarjeta de nuevo por el escáner. Si no han pasado 7 horas desde la última vez que pasé la tarjeta, no me va a leer la tarjeta por lo que no podré dar por finalizada la sesión de trabajo. En este ejemplo, la acción de pasar la tarjeta por el escáner (conducta operante) para concluir la jornada laboral (reforzador) está sujeta a un programa de intervalo fijo de 7 horas, ya que la primera vez que pasamos la tarjeta después de ese período, se refuerza la conducta. Si la pasamos antes, no ocurrirá nada.

Intervalo variable: Una persona se puede encontrar ante un programa de intervalo variable cuando, en época de exámenes, cada vez que publican las calificaciones de un examen, la página se bloquea debido a una alta demanda de usuarios simultáneos. En consecuencia, se ve obligada a recargar la página cada cierto tiempo. En determinadas ocasiones, logrará acceder a la página si la recarga transcurridos 3 minutos, en otras después de 5 minutos, e incluso, en algunas ocasiones, al cabo de un minuto. En este ejemplo, la conducta operante sería recargar la página, y el acceso a la misma actuaría como reforzador. El individuo que intenta comprobar su nota estaría sometido a un programa de intervalo variable ya que tendría acceso a la página tras la primera respuesta que dé pasado un tiempo variable.

**5. Indique las similitudes y diferencias entre los cuatro tipos de programas compuestos secuenciales: tándem, encadenado, mixto y múltiple. Puede crear una tabla para facilitar la comparación.**

Los programas complejos se configuran mediante la combinación de programas simples, clasificándose en secuenciales y simultáneos según su disposición. En esta pregunta nos enfocaremos en los programas secuenciales. Estos requieren la finalización de un programa para iniciar el siguiente. Se distinguen cuatro tipos de programas secuenciales:

- a) Tándem: Compuesto por al menos dos programas simples, donde el reforzador se entrega sólo después de cumplir sucesivamente los requisitos de cada programa simple. Es decir, se debe completar el primer programa simple antes de pasar al segundo.
- b) Encadenado: Similar al tándem, pero cada componente está asociado con un estímulo diferente, cosa que no ocurría en el tándem.

- c) Mixto: En estos programas, el cambio de un programa simple a otro ocurre independientemente de la conducta del sujeto. Por ejemplo, durante 10 minutos se activa un programa IV 60 segundos, luego se cambia a un programa de razón fija 50 durante otros 10 minutos. Los reforzadores se obtienen de manera independiente en cada componente.
- d) Múltiple: Similar al mixto, pero cada componente está señalado con un estímulo diferente, permitiendo al sujeto discriminar en todo momento en qué componente se encuentra.

Estos programas se diferencian en dos características principales. Primero, si la conducta del sujeto determina o no el cambio de un programa simple a otro. En tándem y encadenado, el sujeto debe completar cada componente para avanzar, mientras que en mixto y múltiple, el cambio está predeterminado y no depende de las respuestas del sujeto. La otra característica distintiva es si los componentes del programa complejo están señalados con una clave discriminativa distinta. En encadenado y múltiple, se utilizan estímulos discriminativos como luces o sonidos para señalar los componentes, mientras que en tándem y mixto no hay señales distintivas. El siguiente cuadro resume las características de cada programa:

	Sin estímulo discriminativo	Con estímulo discriminativo
<b>Dependiente conducta</b>	Tándem	Encadenado
<b>Independiente conducta</b>	Mixto	Múltiple

## SEGUNDO BLOQUE: ANÁLISIS DE ARTÍCULO CIENTÍFICO EN INGLÉS

El artículo que se propone analizar este curso es “*Suboptimal choice by pigeons is eliminated when key-pecking behavior is replaced by treadle-pressing*”, de Rodrigo González-Torres, Julio Flores y Vladimir Orduña, publicado en 2020. Se trata de un artículo experimental en el que se combinan aspectos básicos del estudio del aprendizaje en la evaluación de la elección por la alternativa óptima en palomas.

**1. En el artículo de González-Torres et al., se usa un procedimiento al que se le llama de elección subóptima. En dicho procedimiento, se presentan dos alternativas: subóptima - discriminativa y óptima - no discriminativa. Explique la diferencia entre ambas alternativas. Puede hacer una tabla comparativa para ayudar a elaborar la respuesta.**

En la alternativa subóptima - discriminativa se presentan dos estímulos: el 20% de las veces, es decir, con una probabilidad de 0,2, se presenta una luz roja que va a ser seguida por un reforzador en todas las ocasiones ( $p(\text{reforzamiento}) = 1$ ); mientras que en el 80% de las veces, es decir, con una probabilidad de 0,8, se presenta una luz verde que nunca será seguida de comida ( $p(\text{reforzamiento}) = 0$ ). Es decir, la probabilidad global de reforzamiento de la alternativa subóptima - discriminativa es del 0,2, pues sólo se entrega reforzador en el 20% de los ensayos. A esta alternativa se le conoce como subóptima, porque globalmente entrega menos reforzadores que la otra alternativa, y se le conoce como discriminativa porque uno de los estímulos predice de forma

confiable la entrega del reforzador (estímulo discriminativo) y el otro estímulo predice de forma confiable la ausencia del reforzador (estímulo delta).

Por otro lado, en la alternativa óptima - no discriminativa se van a presentar dos estímulos, uno con una probabilidad de 0,2, y otro con una probabilidad de 0,8; sin embargo, independientemente del estímulo que se presente, se entregará reforzador con una probabilidad de 0,5, es decir, se entregará comida en el 50% de los ensayos. A esta alternativa se le llama óptima, porque globalmente entrega más reforzadores que la otra alternativa, y se le conoce como no discriminativa porque sus estímulos no predicen de forma confiable la entrega de reforzador.

Tabla comparativa de alternativas:

	Alternativa subóptima - discriminativa		Alternativa óptima - no discriminativa	
	Estímulo 1	Estímulo 2	Estímulo 1	Estímulo 2
<b>Probabilidad de ocurrencia</b>	0,2	0,8	0,2	0,8
<b>% de ocurrencia</b>	20%	80%	20%	80%
<b>Probabilidad de entrega de reforzador</b>	1	0	0,5	0,5
<b>¿Cuándo entrega reforzador?</b>	Siempre	Nunca	La mitad de las veces	La mitad de las veces
<b>% total de ensayos reforzados</b>	20%		50%	
<b>Probabilidad global de reforzamiento</b>	0,2		0,5	

[Nota: Consulte la Figura 1 del artículo para ver una representación gráfica del procedimiento]

## 2. ¿Cuál fue la justificación teórica por la que González-Torres et al. (2020) decidieron probar el procedimiento de elección subóptima usando la respuesta de presionar pedales (*treadle pressing*) en lugar de picar teclas (*key pecking*) con palomas?

En la introducción del artículo se señala que existen diferencias en los resultados de esta tarea cuando se estudia con palomas o con ratas. En concreto, las palomas eligen más la alternativa subóptima, mientras que las ratas eligen más la alternativa óptima. La causa de estas diferencias no se conoce del todo, pero una de las hipótesis es que las teclas iluminadas usadas para probar el procedimiento con palomas tienen mayor saliencia de incentivo, mientras que las palancas usadas para probar el procedimiento con ratas no tienen tanta saliencia. La saliencia de incentivo es una propiedad de los estímulos condicionados que se convierten en reforzadores secundarios. Es decir, según esta hipótesis, para las palomas, la tecla que señala que siempre se entregará comida, estaría teniendo un mayor valor reforzante. En otros estudios citados en la introducción, se ha visto que usar un pedal cuando se trabaja con palomas es comparable a usar una palanca cuando se trabaja con ratas, en términos de conducta operante, como se señala con la cita de Hemmes (1975). De esta forma, González-Torres et al. (2020) sugieren que utilizar pedales y luces ambientales, en lugar de teclas,

debería disminuir la elección subóptima en palomas, al haberse disminuido la saliencia de incentivo.

**3. Haga un diagrama con las condiciones y subcondiciones del experimento, indicando el título de cada una, y las principales características/cambios.**

<b>Condición</b>	<b>Subcondición</b>	<b>Características</b>
<b>Condición de presión de pedales</b>	Habitación	Una sesión de 20 minutos. Se dio tiempo a los sujetos a habituarse a las cajas, además de entrenarles para que comieran de ambos comederos.
	Moldeamiento de presión de pedal	Se usó el método de aproximaciones sucesivas durante 1 a 5 sesiones para enseñar a las palomas a presionar los pedales.
	Pre-entrenamiento	Se presentaban ambos pedales (también llamados palancas, <i>levers</i> ), hasta que la latencia para su presión fuera similar entre ambos. Es decir, que las palomas tardaran aproximadamente el mismo tiempo en presionar el pedal del panel delantero, que el del panel trasero.
	Pre-exposición a estímulos	En esta subcondición se entrenó a las palomas para responder en presencia de los cuatro estímulos ambientales. Para ello, primero tenían que responder a la luz de la caja (house light), y después responder acorde a un programa de razón fija (RF) 3, tras el cual se entregaba comida.
	Entrenamiento. Procedimiento de elección subóptima.	En esta fase había 40 ensayos de elección forzada (20 por cada alternativa) y 20 ensayos de libre elección. En los ensayos de elección forzada, sólo se presentaba un pedal (palanca) y su luz de la caja, que al ser presionado daba paso a alguno de los estímulos asociados con esa alternativa y su respectiva consecuencia. En los ensayos de libre elección se presentaban ambos pedales (palancas) y sus respectivas luces de las cajas; el primero en ser presionado determinaba qué alternativa (óptima/subóptima) entraba en vigor.
	Reversión del entrenamiento	La fase era igual a la anterior, excepto que la localización de las alternativas cambiaba, es decir, si la alternativa discriminativa estaba en el panel delantero en la condición anterior, en esta condición pasaba a estar en el panel trasero.
<b>Fase de control: picar teclas</b>	Entrenamiento para picar la tecla	Se llevaron a cabo 7 sesiones de automoldeamiento para picar la tecla.
	Pre-entrenamiento	Se entrenó a los sujetos para que picaran la tecla con cualquiera de los colores utilizados en el procedimiento. Se continuó en esta fase hasta que la latencia para picar ambas teclas fuese similar.
	Entrenamiento. Procedimiento	Aquí se usó un procedimiento de elección subóptima clásico, que fue similar al procedimiento con pedales. La

	de elección subóptima clásico	diferencia consistió en que se usaban teclas, en lugar de pedales, y se usó únicamente el panel delantero, de forma que cada una de las alternativas correspondía a un lado (izquierdo o derecho) del panel.
	Reversión del entrenamiento	Se utilizó el mismo procedimiento que en la subcondición anterior, excepto que el lado asignado para cada alternativa fue el contrario que en la subcondición anterior. Es decir, si la alternativa discriminativa había estado en el lado derecho, en esta subcondición estaría en el lado izquierdo.

**4. ¿Por qué se llevaron a cabo fases de reversión de entrenamiento? ¿Qué resultados se observaron al respecto?**

Estas fases se llevaron a cabo para evaluar un potencial sesgo de posición, es decir, que los sujetos estuvieran respondiendo a un lado en concreto, y no a la alternativa asociada a ese lado. En las Figuras 2 y 4 del artículo se puede ver que los sujetos respondían a la alternativa, ya que la proporción de elección observada antes de la reversión fue la misma que se obtuvo al final de la fase de reversión para casi todos los animales.

**5. Explique de forma general los resultados del experimento, haciendo énfasis en las diferencias entre usar teclas y pedales, y a qué se atribuyeron esas diferencias según González-Torres et al. (2020).**

Tal como se describe en la sección de resultados, y se observa en las Figuras 2 y 4 del artículo, la mayoría de las palomas mostraron una preferencia por la alternativa óptima - no discriminativa en la condición de pedales, mientras que la mayoría de los sujetos mostraron una preferencia por la alternativa subóptima - discriminativa en la condición en la que se utilizaron teclas iluminadas. Estos resultados dieron soporte a la hipótesis de que la saliencia de incentivo es importante para determinar la elección de los sujetos, ya que cuando hubo una saliencia alta (condición de teclas), las palomas prefirieron la opción discriminativa, aunque fuera subóptima; mientras que cuando hubo una saliencia baja (condición de pedales), como en los experimentos con ratas, las palomas prefirieron la opción no discriminativa, que era la opción óptima.