

PRINCIPIOS de APRENDIZAJE Y CONDUCTA

Maria Jesús Fdez Aguirre

*Prof. Centro Asociado
Uned Las Palmas*

1



Tema 6.

PROGRAMAS DE
REFORZAMIENTO
Y
CONDUCTA DE ELECCIÓN

Objetivos:

- Conocer los principales **programas de reforzamiento**, tanto básicos como compuestos.
- Distinguir los programas de **razón e intervalo**, teniendo en cuenta si son **fijos o variables** y conocer los patrones conductuales q generan.
- Conocer los **programas de reforzamiento diferencial** de tasas de respuesta y la **Tª del reforzamiento diferencial** de los tiempos entre Rs.
- Conocer los fundamentos de la **Ley de Igualación**, las variables q relaciona y como predice la conducta en circunstancias de elección.
- Conocer los principales estudios sobre el **autocontrol** y el uso de **programas concurrentes encadenados**.
- Conocer las **relaciones asociativas** estímulo-respuesta y estímulo-consecuencia como responsables del aprendizaje instrumental.
- Conocer el **Principio de Premack** y los principales conceptos relacionados con la regulación conductual (punto de bienestar y distribución de Respuesta).

PROGRAMAS DE REFORZAMIENTO

¿Qué es un Programa de reforzamiento?

- Es la forma en que se programan los reforzadores contingentemente con la emisión de la respuesta.
- Especifica una regla que determina **cómo y cuándo** la ocurrencia de una respuesta irá seguida por la administración de un reforzador.

Los programas de reforzamiento influyen, tanto en la forma de aprender una respuesta, como en la forma en que se mantiene.

- **Programa de reforzamiento continuo:**
Cada respuesta correcta es reforzada (ej. Interruptor de la luz)
- **Programa de reforzamiento intermitente:**
No todas las respuesta correctas son reforzadas.(ej..)
Produce un aprendizaje más duradero.
Es más difícil de extinguir.

PROGRAMAS BÁSICOS DE REFORZAMIENTO

- **RAZÓN** → el reforzamiento depende sólo del **nº de respuestas** efectuadas por el sujeto.
- **INTERVALO** → la respuesta se refuerza sólo si ocurre cuando ha transcurrido un cierto **período de tiempo**.

Estos a su vez se subdividen según el criterio de fijo o variable:

RAZÓN FIJA	INTERVALO FIJO
RAZÓN VARIABLE	INTERVALO VARIABLE

Existe 4 tipos de programas de reforzamiento:

- **Razón Fija:** El reforzador se consigue después de un **nº fijo de respuestas**.
- **Razón Variable:** El **nº de respuestas varía** de una recompensa a la otra.
- **Intervalo Fijo:** El **intervalo de tiempo es constante**.
- **Intervalo Variable:** El **período de tiempo varía** de una recompensa a otra

PROGRAMAS BASICOS DE REFORZAMIENTO

Programas de razón

o RAZÓN FIJA (RF):

-Se entrega el reforzador tras un **nº fijo de respuestas** emitidas.

-En un programa de reforzamiento continuo el sujeto suele responder con una tasa constante pero moderada.

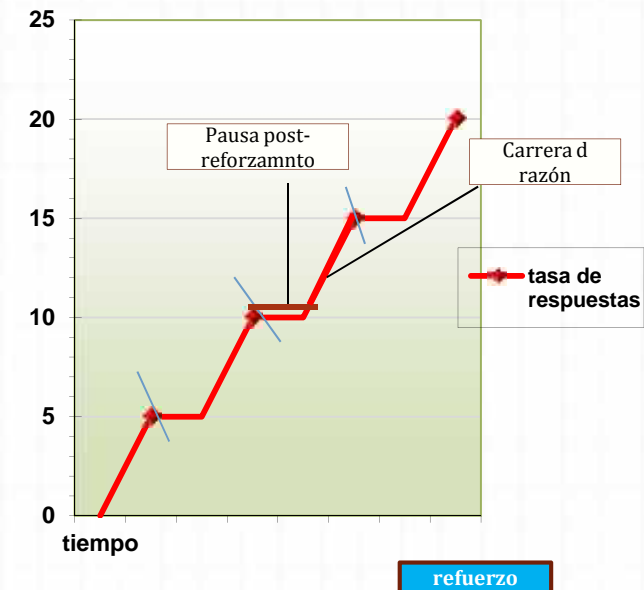
-En un programa de razón fija intermitente la **tasa de respuesta es alta y estable** una vez que la conducta se ha iniciado.

-Se produce una **pausa post-reforzamiento**, tras la entrega del reforzador.

-La tasa de respuesta alta y estable que completa cada requerimiento de la razón se denomina **carrera de razón**.

-**Tensión de la razón:** Si se aumenta demasiado el valor de la razón tienden a ocurrir pausas post-reforzamiento más largas y en casos extremos el sujeto puede dejar de responder.

Registro acumulativo



EJ. de RF:
TRABAJO EN CADENA DE LAS
FÁBRICAS

PROGRAMAS BASICOS DE REFORZAMIENTO

Programas de razón

o RAZÓN VARIABLE (RV):

-El **nº de respuestas** requeridas para obtener el reforzador **varía** de un ensayo a otro.

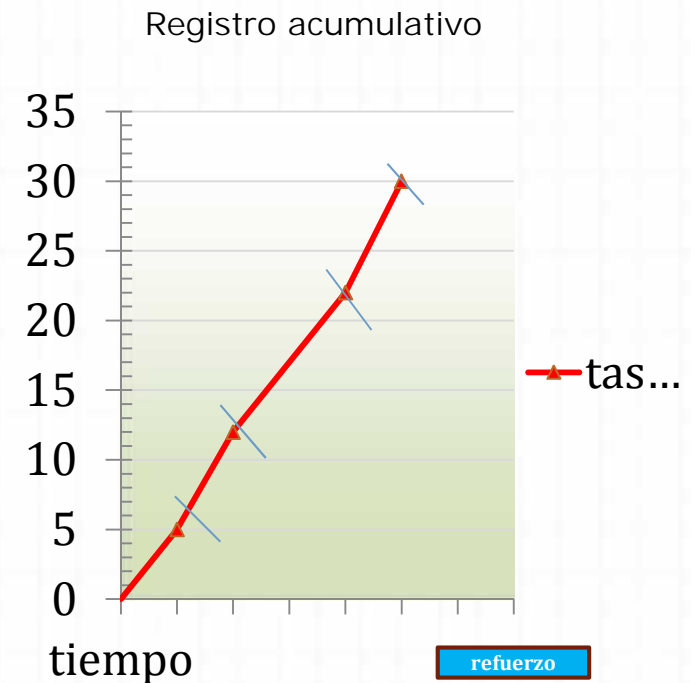
-Un programa de Razón Variable viene definido por el **nº medio de respuestas** requeridas para conseguir el reforzador.

(Ej: RV 7

1ºreforzador 5resp, 2ºreforzador 7 resp, 3ºreforzador 9resp....)

-Las pausas post-reforzamiento son menos probables que en los programas RF. El programa impide que el sujeto adquiera una impresión exacta del nº de respuestas requeridas para obtener el reforzador, reduciéndose así la pausa post-reforzamiento.

-El sujeto suele responder con una **tasa de respuesta muy alta y constante.**



Ej. de RV:
MAQUINAS TRAGAPERRAS

PROGRAMAS BASICOS DE REFORZAMIENTO

Programas de intervalo

o INTERVALO FIJO (IF)

-Las respuestas son reforzadas sólo si ocurren cuando ha transcurrido un **período de tiempo fijo**.

-El intervalo indica sólo cuándo está disponible el reforzador, no cuándo se entrega.

-Los sujetos aprenden a esperar para responder al final del intervalo, esto es, **la tasa de respuesta aumenta hacia el final del IF**.

-Este patrón de respuesta se denomina **festón del intervalo fijo**.

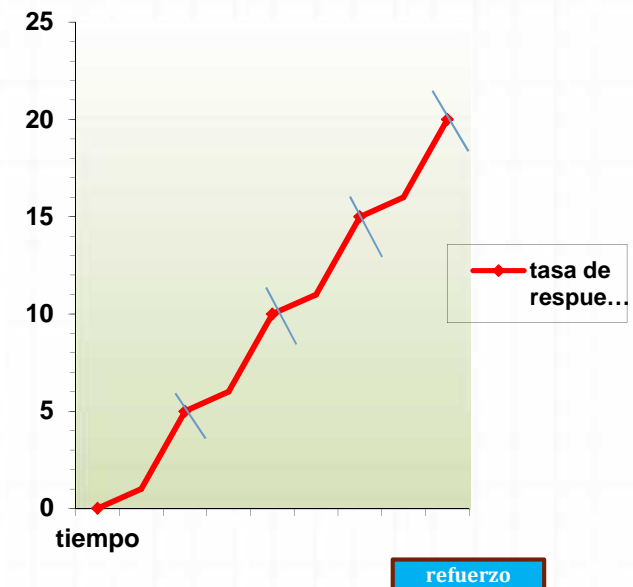
-La ejecución en un programa IF refleja la precisión del sujeto para contar el tiempo.

Evidencias:

-la pausa post-reforzamiento y la aceleración hacia el final del intervalo.

-A mayor duración del intervalo, mayor pausa.

Registro acumulativo



Ej. de IF:
PAUTAS DE ESTUDIO
EXÁMENES

PROGRAMAS BASICOS DE REFORZAMIENTO

Programas de intervalo

o INTERVALO VARIABLE (IV)

-Las respuestas son reforzadas si ocurren tras haber transcurrido un **período de tiempo variable**.

-El intervalo se define por el **tiempo medio** transcurrido entre reforzamientos sucesivos.

(Ej: IV 2min

1º reforzador 1mto, 2º reforzador 3mtos, 3º reforzador 2mtos)

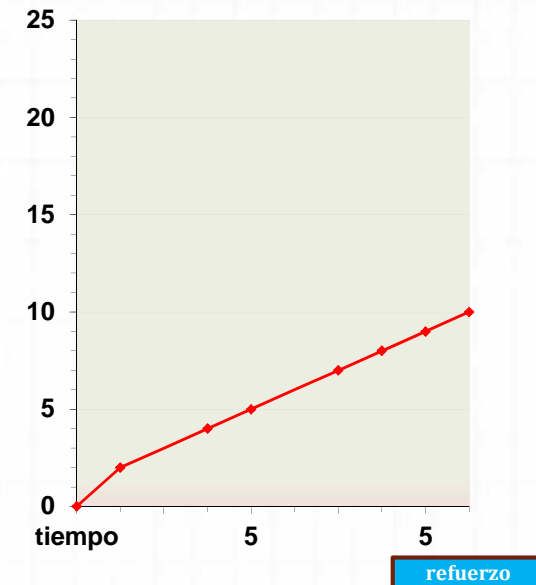
-El sujeto tiene que realizar la respuesta para obtener el reforzador.

-El registro acumulativo de respuestas presenta una ejecución con una **tasa de respuesta constante y estable, sin pausas**.

o En los programas de IF y IV el reforzador está disponible hasta que se realice la R, sin importar el tiempo que esto lleve.

o **Espera limitada:** se restringe el tiempo que está disponible el reforzador.

Registro acumulativo



Ej. de IV:
LA PESCA

PROGRAMAS BÁSICOS DE REFORZAMIENTO

Comparación de los programas de intervalo y razón

o Similitudes

- Los programas de RF e IF:
 - producen pausa postreforzamiento
 - producen altas tasas de respuesta tras la pausa
- Los programas de RV e IV:
 - mantiene tasas estables de respuesta
 - no presentan pausas

o Diferencias

- La conducta mantenida por los programas de intervalo está mediada por el sentido del tiempo
- Los programas de razón presentan una conducta instrumental más vigorosa, esto es, tasas más altas de respuesta

o **EXPLICACIÓN:** T^a centrada en el tiempo entre respuestas (TER)

- En los programas de razón** los tiempos entre respuestas largos posponen el reforzamiento. De hecho cuanto antes complete el requerimiento de la razón, más rápido recibirá el reforzador. Por tanto un programa de razón favorece **tiempos entre respuestas cortos** y por consiguiente una tasa de respuestas alta.
- En los programas de intervalo** la probabilidad de recompensa aumenta con **tiempos entre respuestas más largos** y esto implica tasas de respuestas más bajas.

PROGRAMAS BÁSICOS DE REFORZAMIENTO

Programas de reforzamiento diferencial

o Programas de tasas de respuesta:

Se requiere que el sujeto responda a una tasa particular para obtener el reforzador. El reforzamiento de una determinada respuesta vendrá determinado en función del tiempo transcurrido entre dicha respuesta y la respuesta anterior. (nº respuestas/tiempo)

o Reforzamiento diferencial de tasas altas (RDA):

En los programas RDA, una respuesta se refuerza sólo si ocurre **antes** de que haya transcurrido cierta cantidad de tiempo tras la respuesta precedente.
Se refuerzan los TER cortos.

Este procedimiento fomenta tasas altas de respuesta.

Ej:

o Reforzamiento diferencial de tasas bajas (RDB):

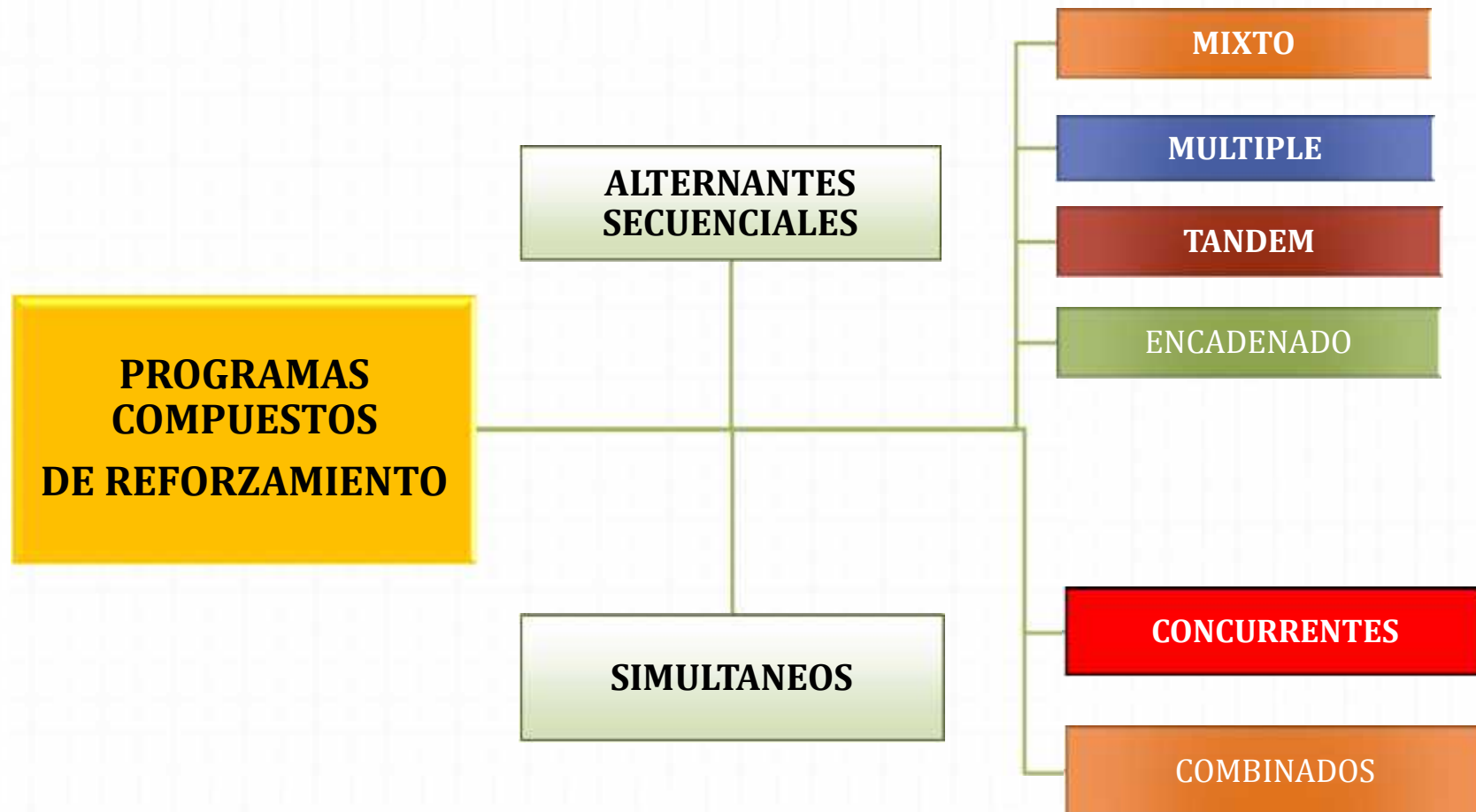
Se refuerza una respuesta sólo si ocurre **después** de que haya transcurrido cierta cantidad de tiempo desde la respuesta anterior.

El sujeto tiene que refrenar su respuesta hasta que ha pasado un cierto período de tiempo. **Se refuerzan los TER largos**

Ej:

PROGRAMAS COMPUESTOS DE REFORZAMIENTO

Los programas compuestos resultan de la aplicación de dos o más programas básicos de reforzamiento.



PROGRAMAS COMPUESTOS DE REFORZAMIENTO

Programas alternantes

o Mixto

- Dos o más programas básicos alternándose al azar.
- El cambio de uno a otro de los programas es independiente de la conducta del sujeto.
- Depende de la duración de un tiempo determinado por el experimentador.
- **Recibe un reforzador al finalizar cada uno de los componentes del programa.**

Ej: Cada 10 minutos se alternarán dos programas: IF 60 seg – RF 50

o Múltiple

- Igual q el mixto, pero cada componente se señala con una **clave discriminativa** distinta.
- El sujeto es informado en todo momento de en qué parte del programa se encuentran.

Ej: una luz (Ed) indica un programa IF 60 seg
la luz apagada (E) indica el programa RF50

- Estos programas se han utilizado en la investigación sobre discriminación y generalización y el estudio del control estimular.

PROGRAMAS COMPUESTOS DE REFORZAMIENTO

Programas secuenciales

o Tandem

Dos más programas simples q se presentan siempre en el mismo orden.

Se debe completar el primer programa para poder iniciar el segundo.

El reforzador se recibe al final del 2º programa,

Sólo es posible obtener el reforzador si se cumplen en los dos programas.

o Encadenado

Igual que el tándem pero cada uno de los programas está señalado

por un **E discriminativo (Ed)**

Estos programas se han utilizado para estudiar el reforzamiento condicionado.

El Ed que señala el inicio del 2º componente del programa adquiere propiedades reforzantes por su asociación con el reforzador q se consigue al final del 2º programa.

PROGRAMAS COMPUESTOS DE REFORZAMIENTO

Programas simultáneos

o Concurrentes

- Se presentan al menos dos programas básicos de forma simultánea.
- Se puede cambiar de un programa a otro sin necesidad de completar un tiempo o un nº d respuestas determinado en cada uno de los programas.
- Los programas concurrentes estudian los mecanismos de la **conducta de elección**
- Se mide el nº de respuestas en cada uno de los componentes del programa y se calcula la proporción en qu se distribuyen las respuestas en cada uno d ellos.

o Combinados

Combinaciones simultáneas de un programa de razón y uno de intervalo.

- **P. alternativos:** Se refuerza una respuesta cada vez que se cumple con el requisito del programa de razón o el de intervalo.
- **P. conjuntivos:** Se refuerza una respuesta cada vez que se cumple al **mismo tiempo** con el requisito de la razón y el intervalo
- **P. entrelazados:**

CONDUCTA DE ELECCIÓN



CONDUCTA DE ELECCIÓN

La conducta de elección se ha estudiado:

➤ **Tradicionalmente** utilizando un procedimiento de **premios concurrentes:**

- Se daba elegir entre dos actividades diferentes
- Se medía el **tiempo** que el animal empleaba en cada una de las dos opciones.
- Se inferían las preferencias relativas de los sujetos.
- No es adecuado pues se comparan **respuestas distintas**.

➤ **Actualmente** se utilizan los **programas concurrentes:**

- Se aplican dos programas de reforzamiento a la vez.
- Se elige entre distintas alternativas que requieren una **respuesta operante idéntica**.
- La conducta de elección se refleja en la **distribución de las respuestas** en las dos opciones alternativas.

El sujeto distribuirá las respuestas entre las dos opciones dadas en función del programa de reforzamiento de cada una de las alternativas.

CONDUCTA DE ELECCIÓN

MEDIDAS DE LA CONDUCTA DE ELECCIÓN

- o **Tasa relativa de respuesta** de cada respuesta alternativa.

$$\frac{RA}{(RA+RB)}$$

* RA: Tasa relativa de respuestas en la alternativa: A
(RA+RB): Tasa de respuestas total

- La razón será $R=0,5$ → Si el sujeto elige igualmente las dos teclas.
- La razón será $R >0,5$ → Si la tasa d respuesta a la tecla A es mayor q en la B.
- La razón será $R <0,5$ → Si la tasa d respuesta es mayor en la tecla B.

- o **Tasa relativa de reforzamiento** para cada alternativa de respuesta.

$$\frac{Fa}{(Fa+Fb)}$$

* Fa: Tasa de reforzamiento en la alternativa:A
(Fa+Fb): Tasa de reforzamiento total

- La razón será $r=0,5$ → Si recibe el mismo nº de reforzadores en los dos programas.
- La razón será $r >0,5$ → Si recibe más refuerzos en el programa de la tecla A.
- La razón será $r <0,5$ → Si recibe más refuerzos en el programa de la tecla B.

CONDUCTA DE ELECCIÓN

LA LEY DE LA IGUALACIÓN

-La ley de igualación fue enunciada por **Herrnstein**.

-Describe la relación entre la tasa de respuesta y la tasa de reforzamiento:

Las tasas relativas de respuesta se igualan con las tasas relativas de reforzamiento.

$$\frac{Ra}{(Ra + Rb)} = \frac{Fa}{(Fa + Fb)}$$

“La tasa relativa de respuesta entre dos alternativas de respuestas concurrentes es igual a la tasa relativa (frecuencia) de reforzamiento asociada con cada una de dichas alternativas de respuesta”.

-Si en la alternativa A se ofrece una mayor frecuencia de reforzamiento que en la alternativa B también se dará un mayor número de respuestas en esta alternativa, y viceversa.

-Si los reforzadores se distribuyen por igual en las dos alternativas, las respuestas también se distribuirán igual en las dos alternativas.



El sujeto distribuirá las respuestas entre las dos opciones dadas en función del programa de reforzamiento de cada una de las alternativas.

Tener en cuenta que existen también otros parámetros del reforzador (frecuencia, magnitud y demora)

CONDUCTA DE ELECCIÓN

LA LEY GENERALIZADA DE LA IGUALACIÓN

Baum amplió la ley de la igualación para corregir los sesgos e incorporar las desviaciones por sobreigualación e infraigualación.

$$\frac{R_a}{R_b} = k \left(\frac{F_a}{F_b} \right)^s$$

***k**: Sesgo de respuesta
S: Sensibilidad a las tasas relativas de reforzamiento

➤ **Igualación perfecta:** $s=1$

➤ **Sobreigualación:** $s>1$

Cuando la mejor alternativa es valorada por encima de lo que predice la igualación. (Responden **más**) (mayor sensibilidad a la tasa relativa de reforzamiento)

➤ **Infraigualación:** $s<1$

Cuando el sujeto responde **menos** de lo previsto por la igualación en el programa más ventajoso. (sensibilidad reducida a las tasas relativas de reforzamiento)

➤ **Sesgo de respuesta:** Ante dos alternativas igualmente favorables, preferencia por una d ellas

El sesgo se suele producir cuando:

- Las alternativas d respuesta son diferentes (picotear tecla/presionar pedal, color, fuerza requerida para responder...etc)
- Cuando los reforzadores son diferentes.
- Conducta de alternancia

Si no existen sesgos $k=1$, si hay sesgo $k>1$ ó $k<1$, dependiendo del sesgo.

CONDUCTA DE ELECCIÓN

TEORIAS
de la
IGUALACION

MOLECULARES

MOLARES

T^a MEJORAMIENTO

CONDUCTA DE ELECCIÓN

TEORIAS DE LA LEY DE IGUALACIÓN

➤ T^a maximización molecular:

Los sujetos eligen cualquier alternativa de respuesta con mejores probabilidades de obtener un **reforzamiento en ese momento** (**probabilidad momentánea**)

Explica bastante bien la distribución de respuestas en programas concurrentes de intervalo ajustándose a lo que predice la ley de igualación.

➤ T^a maximización molar:

Los sujetos distribuyen sus respuestas entre las diferentes alternativas con el fin de maximizar la cantidad de **reforzamiento total**.

(Eligen la alternativa que proporciona mayor **tasa global** de reforzamiento).

Esta T^a se formuló para explicar la elección en programas concurrentes d razón.

Dificultades:

- No puede explicar la conducta de elección en los programas RV-IV.
- Dificultad para explicar porqué los animales prefieren reforzadores variables a fijos, aunque se ofrezca igual tasa de reforzamiento.

CONDUCTA DE ELECCIÓN

TEORIAS DE LA LEY DE IGUALACIÓN

➤ T^a mejoramiento:

Se encuentran en una escala temporal entre las T^a molares y las T^a moleculares. La T^a del mejoramiento asume que q los sujetos cambian d una alternativa a otra para mejorar **la tasa local de reforzamiento**.

Eligen la alternativa q en cada momento presenta mejor tasa local d reforzamiento.

(nº de Resp o Reforz / Tiempo dedicado a responder).

Por tanto, intrínseco a la T^a del mejoramiento es **la igualación temporal:**

- Al elegir la alternativa q proporciona un mejor reforzamiento: darán una mayor tasa de respuesta, obtendrán una mayor tasa de reforzamiento y permanecerán más tiempo en esa alternativa.
- Responderán menos a la alternativa menos favorable, se obtendrán menos reforzadores pero también permanecerán menos tiempo en esta alternativa.



Por tanto

Las tasas locales de respuesta y de reforzamiento serán iguales en las diferentes alternativas de respuesta.

CONDUCTA DE ELECCIÓN

ELECCIÓN CON COMPROMISO:

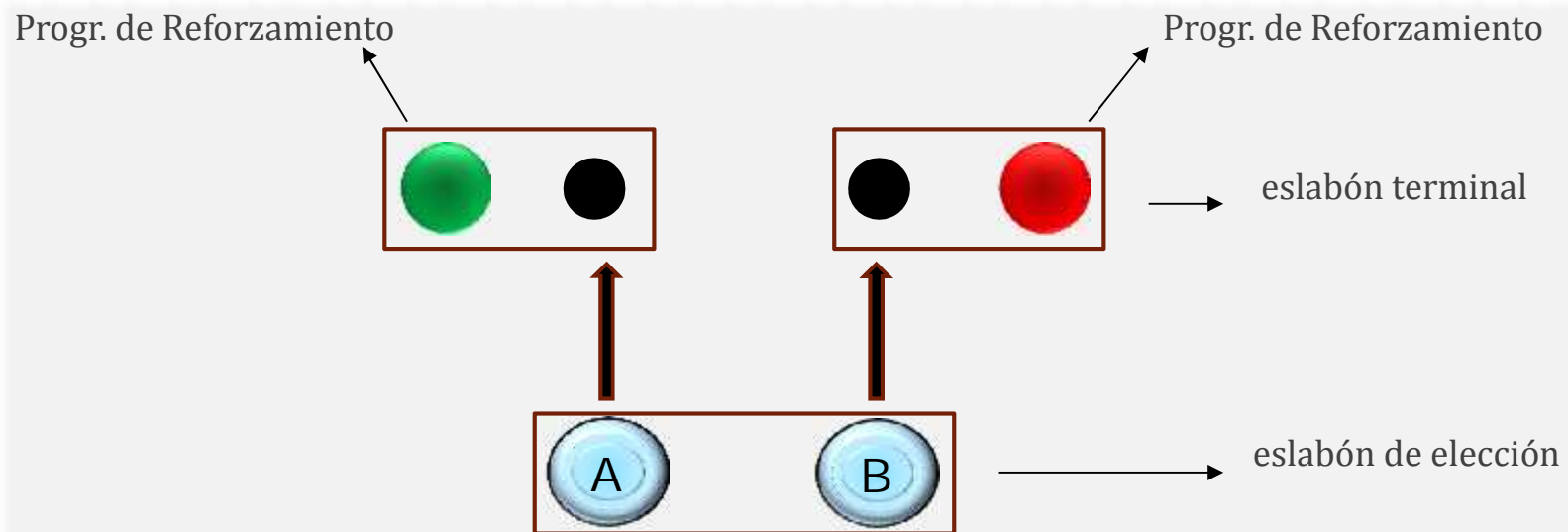
Programas concurrentes encadenados

Procedimiento de reforzamiento complejo, en el que se permite al participante elegir cuál de varios programas simples de reforzamiento funcionará.

Una vez realizada la elección, las alternativas rechazadas no estarán disponibles durante algún tiempo. Estos programas tiene al menos dos fases.

-1ª fase: **eslabón de elección**, el sujeto elige entre dos opciones d resp. idénticas

-2ª fase: **eslabón terminal**, programa de reforzamiento.



CONDUCTA DE ELECCIÓN

Programas concurrentes encadenados

Características:

- La oportunidad de obtener el reforzador se da **sólo** en el eslabón terminal.
- Los programas concurrentes encadenados implican una **elección con compromiso**. Una vez realizada la elección el sujeto está comprometido con esa elección hasta el final del eslabón terminal.
- La consecuencia de responder en el eslabón inicial no es un reforzador primario como la comida, sino un **estímulo** asociado con el eslabón terminal que se convierte en un **reforzador condicionado**.
- Los sujetos igualan la tasa relativa de respuesta con la tasa relativa de reforzamiento en los eslabones terminales.
- Además, igualan la tasa relativa de respuesta en el eslabón inicial con la tasa relativa de reforzamiento en el eslabón terminal

CONDUCTA DE ELECCIÓN

ESTUDIOS DE AUTOCONTROL

El **autocontrol** se trata de elegir una recompensa grande demorada sobre una recompensa pequeña pero inmediata (**impulsividad**).

Los procedimientos utilizados en los estudios de autocontrol han sido:

o **Procedimiento de elección directa** (p. concurrente simple):

Se prefiere la recompensa pequeña inmediata. No se muestra autocontrol.

La inmediatez del reforzador influye más en la elección que a magnitud

o **Programa concurrente encadenado:**

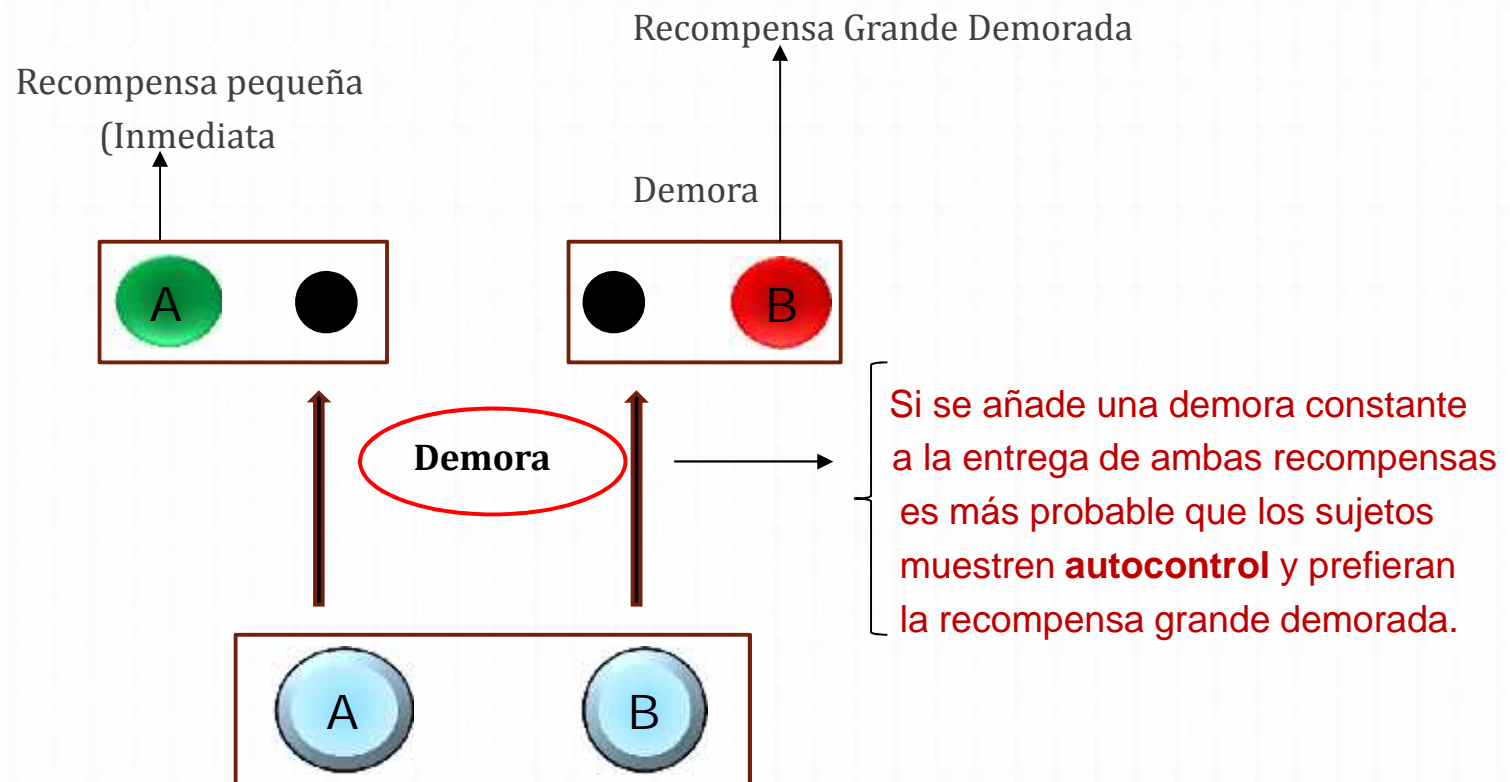
Si se impone un tiempo de demora suficiente antes del componente terminal se muestra autocontrol:

Se prefiere la recompensa grande demorada.

“Si se añade una demora constante a la entrega de ambas recompensas, es más probable que los sujetos muestren autocontrol y prefieran la recompensa grande demorada”

CONDUCTA DE ELECCIÓN

PROGRAMA CONCURRENTENTE ENCADENADO EN AUTOCONTROL (Rachiln y Green)



CONDUCTA DE ELECCIÓN

- o La clave para la explicación del autocontrol es que **el valor de un reforzador se reduce en función del tiempo** que se tiene que esperar para obtenerlo.
- o La función matemática que describe esta disminución del valor del reforzador es:

$$Va = \frac{Ma}{(1 + kDa)}$$

El valor de un reforzador (**V**) está directamente relacionado con la magnitud de la recompensa (**Ma**) e inversamente relacionado con la demora de la recompensa (**D**).

k indica con que rapidez disminuye el valor de la recompensa en función de la demora.

Esta función explica que al **incrementar la demora** de las recompensas grande y pequeña hace más fácil que se muestre **autocontrol** ya que aumenta el valor de la recompensa grande.

TEORÍAS DEL APRENDIZAJE INSTRUMENTAL



TEORIAS DEL APRENDIZAJE INSTRUMENTAL

TEORIAS ASOCIACIONISTAS (la fuerza de la respuesta se adquiere por asociaciones)

➤ Teorías E-R:

Thorndike

- Su aportación fue teórica: formula la **ley del efecto** .
- Postula q la asociación entre la situación ambiental (E) y la respuesta (R) depende de la propia consecuencia d la acción, por tanto , el papel del reforzador es crear la asociación E-R

Guthrie

- La función del reforzador es facilitar la formación de una asociación entre E-R.
- Fue el primero en distinguir entre **actos y movimientos**.

Hull

- La consecución del reforzador refuerza la formación d una asociación E-R debido a una reducción del impulso (necesidad biológica), al tiempo que incita a la acción (incentivo).
- La ejecución conductual va a depender de la fuerza del **hábito**, del nivel del **impulso** y del valor del **incentivo**.

➤ Teorías E-E:

Tolman

- Para Tolman los reforzadores no actúan directamente sobre la asociación E-R sino que la conducta está mediada por eventos internos como las **expectativas de recompensa**.
- Distinguió entre **aprendizaje y ejecución**. El reforzador no es necesario para el aprendizaje, pero sí para la ejecución.

TEORIAS DEL APRENDIZAJE INSTRUMENTAL

TEORIAS SELECCIÓN de RESPUESTAS

➤ Tª DE LA PROBABILIDAD DIFERENCIAL (Premack)

“Dadas dos repuestas en un procedimiento de C. Instrumental, la respuesta más probable reforzará a la menos probable y la respuesta menos probable no reforzará a la más probable”

El principio de Premack

El efecto del reforzamiento se basa en la **probabilidad diferencial** de 2 actividades y tiene un **gran poder predictivo**:

- Primero:** se observa la probabilidad o preferencia de 2 actividades cuando el sujeto no está limitado.(establecimiento de la **línea base**)
- Segundo:** se establece una **relación de reforzamiento** en la que la **actividad preferida (A)** se limita, y debe ocurrir **sólo de forma contingente** con la realización de la actividad **menos preferida (B)**.
(Actividad preferida: reforzador / Actividad menos preferida: instrumental)
- Resultado:** la actividad reforzadora ocurre a menor probabilidad que en condiciones de libre acceso y la actividad instrumental aumenta su probabilidad de ocurrencia.

(Ej: A: podrás ver la tele cuando, B: hayas recogido los platos)

TEORIAS DEL APRENDIZAJE INSTRUMENTAL

TEORIAS SELECCIÓN de RESPUESTAS

➤ Tª DE LA PRIVACIÓN DE RESPUESTA (Einsenberger, Karpman, Trattner)

La restricción de la respuesta reforzadora es el factor crítico para el reforzamiento instrumental.

- Los programas de reforzamiento incrementarán la ocurrencia de una respuesta operante por encima de su línea base, **si y sólo si** dichos programas **privan** al individuo de realizar la respuesta reforzadora.
- Para que una actividad pueda funcionar como reforzador hay que restringir la realización de dicha actividad en relación con su **punto de bienestar** (línea base), pero **no es necesario que tenga que ser la más preferida**.
- Predice que la tasa de respuesta y la tasa de reforzamiento deben relacionarse en forma de **U invertida** en vez de forma lineal:

A medida q disminuye la frecuencia d reforzamiento, la tasa de respuesta instrumental aumenta (xq el beneficio d obtener el reforzador compensa el hecho d responder más) pero si posteriormente aumenta el requisito del programa y baja más la frecuencia de reforzamiento, el coste de responder no compensa, por tanto disminuyen las respuestas.

Ej: Actividades de un niño en su punto de bienestar (sin limitaciones)

- ver la televisión: 2horas (actividad preferida)
- dibujar: 1h
- leer: ½ hora