

Adaptación del Comportamiento a Cambios en las Propiedades Estadísticas del Entorno

Arturo Bouzas y Carlos Velázquez

distribuciones
elección
restricciones
evolución
incentivos
información
modelos
tiempo
cooperación
reforzamiento
ajuste
dinámica
incertidumbre
creencias
cuantitativa
inferencia
comportamiento
adaptable
Lab25
separación
preferencias
contingencias
coordinación
modelamiento
riesgo
juegos

Colaboradores

- Manuel Villarreal
- Alejandro Segura
- Luis Baroja

- Priming







Conclusiones

- 1. Selección natural opera sobre mecanismos que permiten comportamiento adaptable a propiedades estadísticas del entorno.
- 2. Análisis de soluciones óptimas dadas restricciones evolutivas.
- 3. Mecanismos que funcionan como algoritmos que implementan la soluciones y que están compuestos de componentes bien definidos y generales.
- 4. Como creamos un robot. wall-E

- Selección natural opera sobre el éxito reproductivo diferencial, pero este a su vez depende de las acciones del organismo y todas estas acciones consumen y agotan el suministro de energía.

- Los sistemas biológicos gastan energía y requieren de un constante reabastecimiento.
- Los procesos evolutivos van moldeando mecanismos que generan distintas formas exitosas de reabastecimiento y reproducción (comportamiento adaptable) a través del éxito reproductivo diferencial de la instancia metabólica (el organismo).

Propiedades Estadísticas del Entorno

- Los sucesos biológicamente importantes (SBI) para los organismos se distribuyen de acuerdo a una regla que especifica su disponibilidad. Si no hubiese variabilidad en la disponibilidad de SBI no habría selección de comportamientos que pudiesen cambiar y ajustarse

- Los retos de adaptación surgen de las propiedades estadísticas que describen la disponibilidad de sucesos biológicamente importantes (SBI) para un organismo.
- La primera tarea es caracterizar estas propiedades, acerca de las cuales los organismos tienen incertidumbre.

- Es claro que si las consecuencias relevantes en el entorno se distribuyen en ciertos tiempos, lugares, asociadas con ciertas señales y con el comportamiento, un organismo que pueda detectar estas propiedades estadísticas, podrá asignar óptimamente su comportamiento a las metas en competencia.

- Cinco características de la disponibilidad de las SBI acerca de las cuales puede haber incertidumbre:
- Tiempo de ocurrencia
- Lugar de ocurrencia
- Su covarianza con otras características del entorno y
- Su covarianza con el comportamiento de un organismo.
- Tasa de ocurrencia

- Adicionalmente las propiedades estadísticas del entorno incluyen la incertidumbre acerca del tiempo, lugar, covarianzas, y tasas, la cual puede ser esperada o inesperada.
- Si es esperada hay una distribución de probabilidad que la genera, si es inesperada los parámetros de la distribución cambian de acuerdo a otra distribución de probabilidad.

- Un ejemplo de la primera es la probabilidad de águila al lanzar una moneda al aire, un ejemplo de la segunda es una moneda que cambia la probabilidad de águila en el tiempo de acuerdo a otra distribución de probabilidad.

Modelo de Refuerzo

- Un algoritmo de ascenso de colina
- Don componentes:
 - 1. Una regla que integra las experiencias previas (memoria)
 - 2. Una regla que compara la experiencia integrada con una nueva experiencia.

Modelo de refuerzo

- Los organismos están inmersos en entornos en los que SBI se despliegan en el tiempo.
- Los modelos de refuerzo especifican la regla que describe como se integra la experiencia con el flujo de un SBI para poder predecir la ocurrencia del siguiente suceso.
- La predicción le sirve al organismo para distribuir óptimamente su comportamiento.

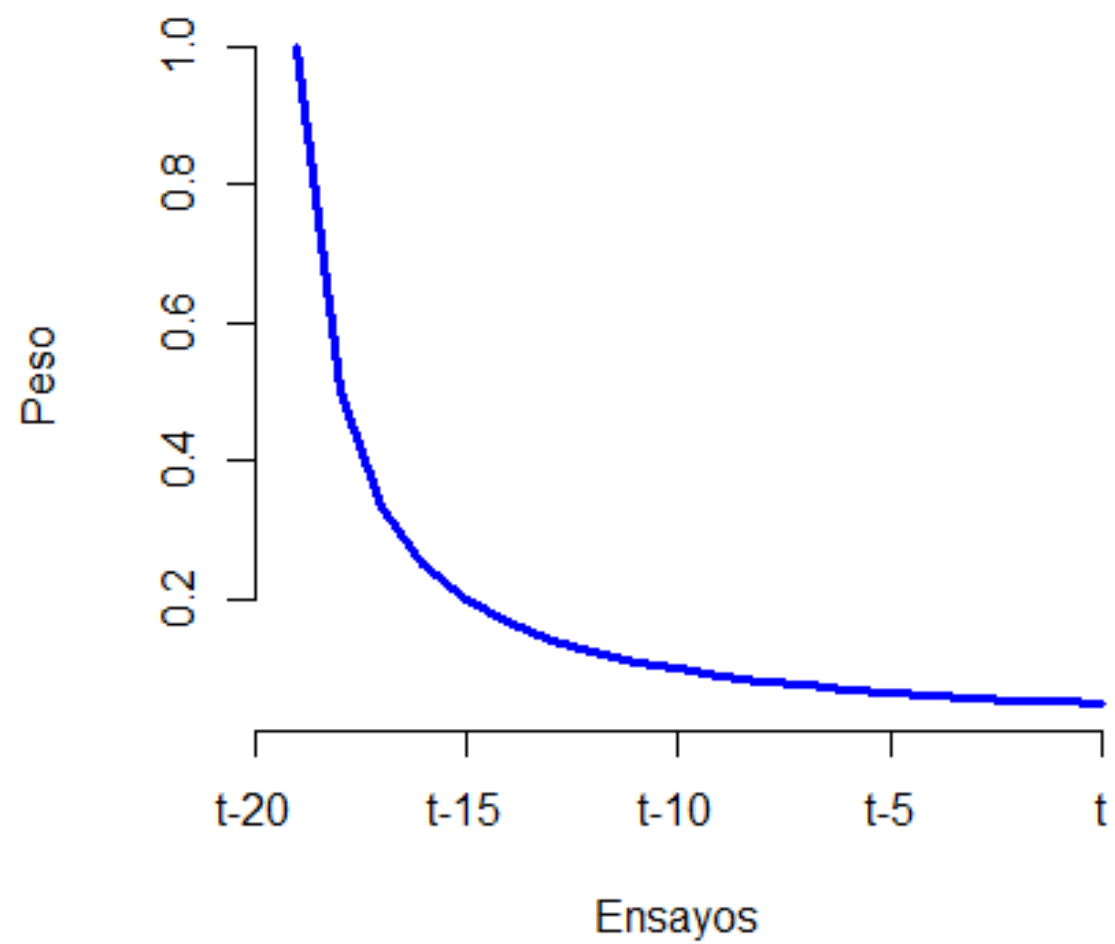
- Considere el caso más sencillo, una serie temporal que describe si ocurrió o no un refuerzo.
- Dos factores determinan la predicción del siguiente suceso:
 1. la experiencia con los SBI hasta el momento t y
 2. la ocurrencia o no del SBI en el momento t

- Necesario especificar:
 1. Como se integra la experiencia acumulada y
 2. El impacto que tiene una nueva experiencia.

Computar una media corrida.

Si la serie no cambia en el tiempo, lo mas sencillo es una media aritmética simple.

Poco sensible al cambio



- Una solución que toma en cuenta el que la serie de experiencias no es constante consiste en permitir que la predicción o valor predictivo de un suceso (Estímulo o Respuesta) sea una suma *ponderada* de las dos variables:
 1. la experiencia con los SBI hasta el momento t y
 2. la ocurrencia o no del SBI en el momento t

- El factor de ponderación determina la importancia de la historia con los sucesos (SBI) relativa a la importancia de la experiencia presente con la ocurrencia o no de ese SBI.

$$VP_{t+1} = (1 - \alpha)VP_t + \alpha R_t$$

donde: $0 < \alpha < 1$

- $$VP_{t+1} = (1 - \alpha)VP_t + \alpha R_t$$
- El parámetro α determina la importancia de la experiencia acumulada hasta el momento n , relativa a si ocurre un SBI.
- Valores cercanos a cero sugieren que la experiencia acumulada es más importante que una nueva experiencia,

.

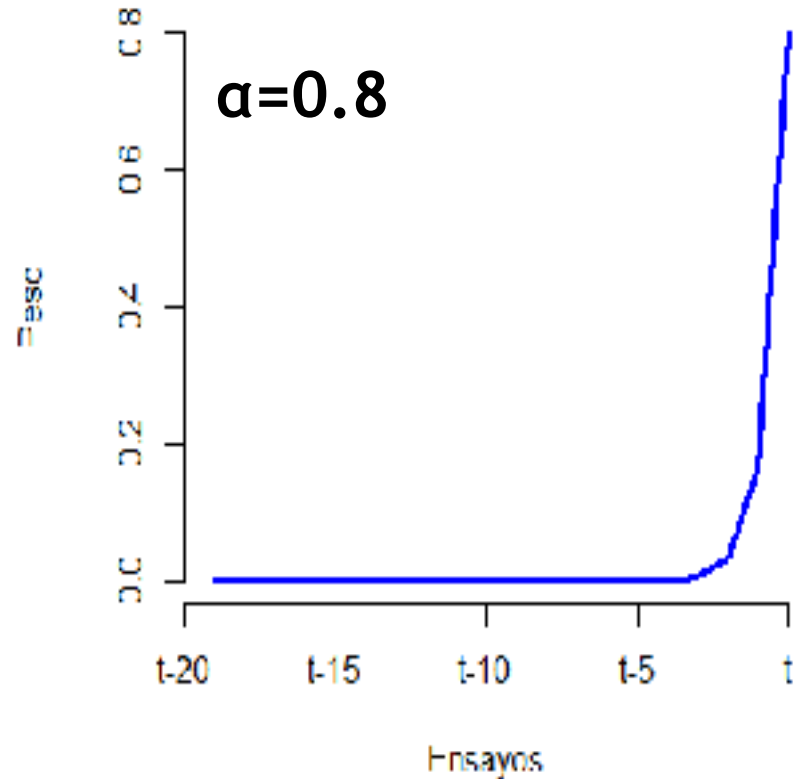
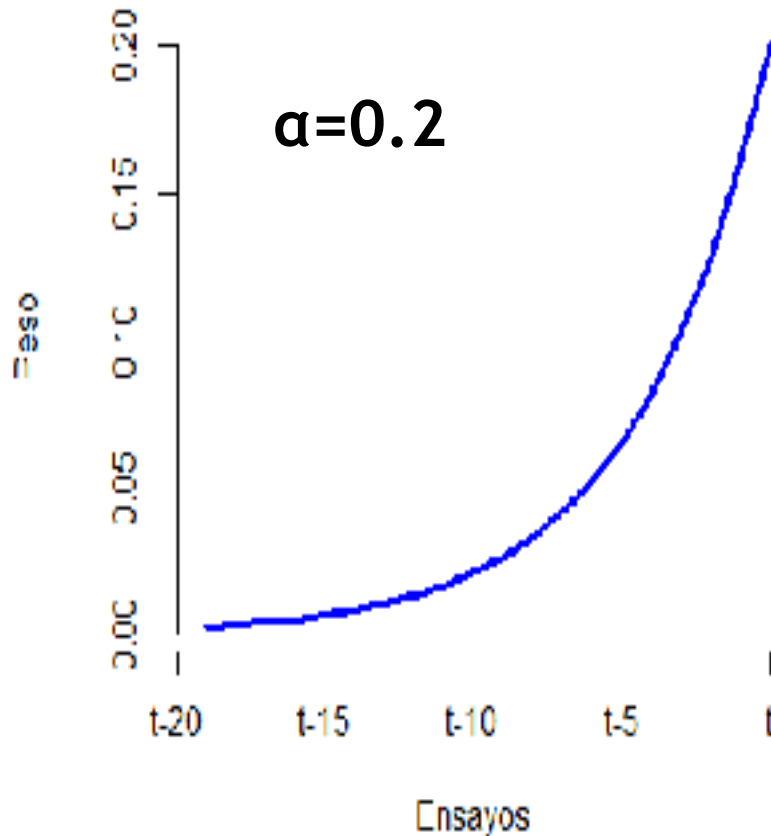
$$VP_{t+1} = (1 - \alpha)VP_t + \alpha R_t$$

- Valores cercanos a uno sugieren que la presentación de un refuerzo, minimiza el impacto de la experiencia acumulada hasta ese momento.
- El parámetro α puede interpretarse como una ventana temporal que determina que tan atrás contempla un organismo su experiencia para predecir el futuro.

- Esta media ponderada es una media corrida exponencial que le da más peso a los datos recientes y permite una adaptación más rápida a cambios en la serie de refuerzos.

Importancia de los reforzadores pasados

-Mayor peso a eventos recientes.



- Una pequeña manipulación algebraica de la ecuación anterior genera una forma equivalente del modelo de refuerzo que es el más común hoy en día.

$$VP_{t+1} = (1 - \alpha)VP_t + \alpha R_t$$

$$VP_{t+1} = VP_t - \alpha VP_t + \alpha R_t$$

$$\mathbf{VP}_{t+1} = \mathbf{VP}_t + \alpha(\mathbf{R}_t - \mathbf{VP}_t)$$

- Si dejamos que: $\delta = (R_n - VP_n)$ sea el error de predicción entonces tenemos la expresión que más comúnmente se usa en la literatura:

$$VP_{n+1} = VP_n + \alpha \delta$$

- A esta forma equivalente de la ecuación se le conoce como *regla delta*
- Las dos versiones (regla delta y suma ponderada) implican una media corrida exponencial

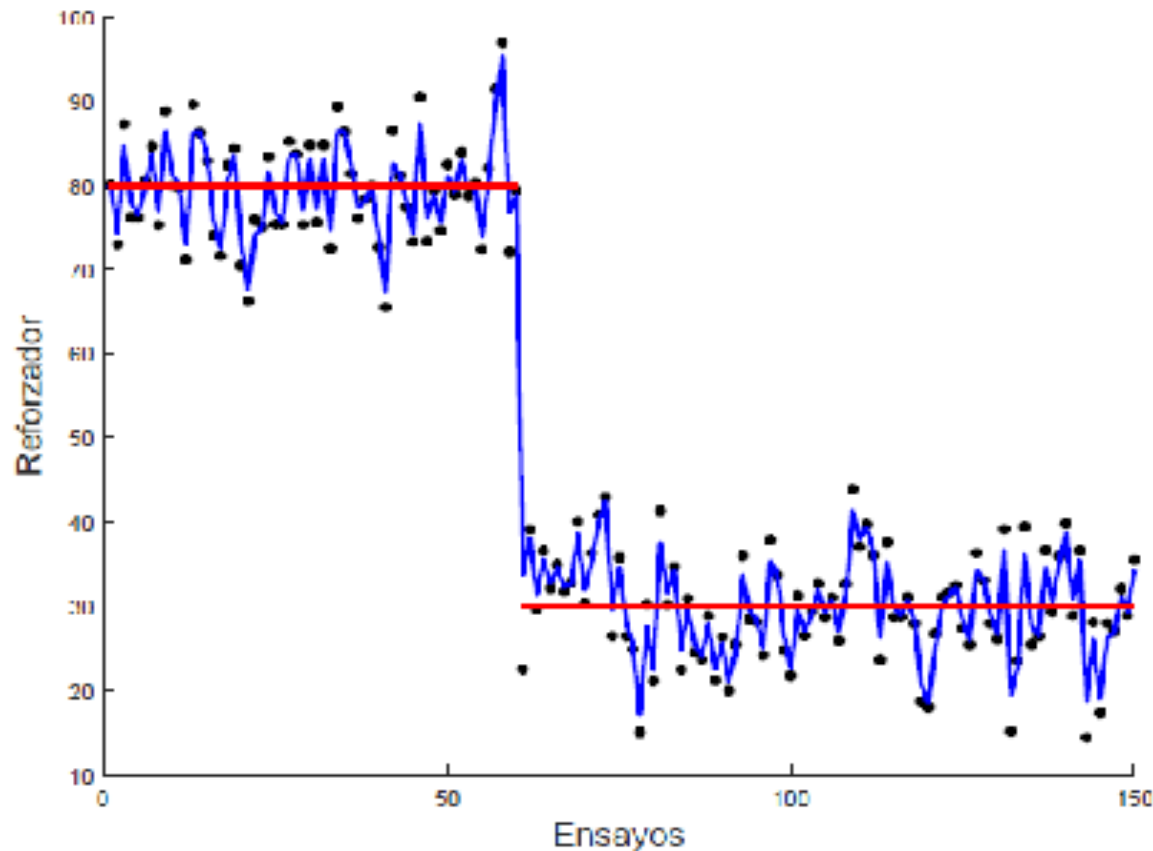
- En esta forma es más claro que el valor predictivo va cambiando como una función de la magnitud del error de predicción y que el motor del aprendizaje es la reducción de este error.
- En ambas interpretaciones el objetivo computacional es buscar la mejor predicción posible, en un caso vista como una serie temporal, en la otra como un algoritmo que reduce error en la predicción

- Como se comporta el modelo de refuerzo cuando las condiciones no son estables y el organismo enfrenta cambios en las condiciones de refuerzo?

- Considere dos casos extremos, en uno alfa tiene un valor cercano a 1, en este caso un cambio abrupto generaría un error de predicción grande que tendría un efecto importante en el nuevo valor actualizado . Con este valor de α los organismos serían sensibles a cambios bruscos después de la exposición a una secuencia con poca variabilidad

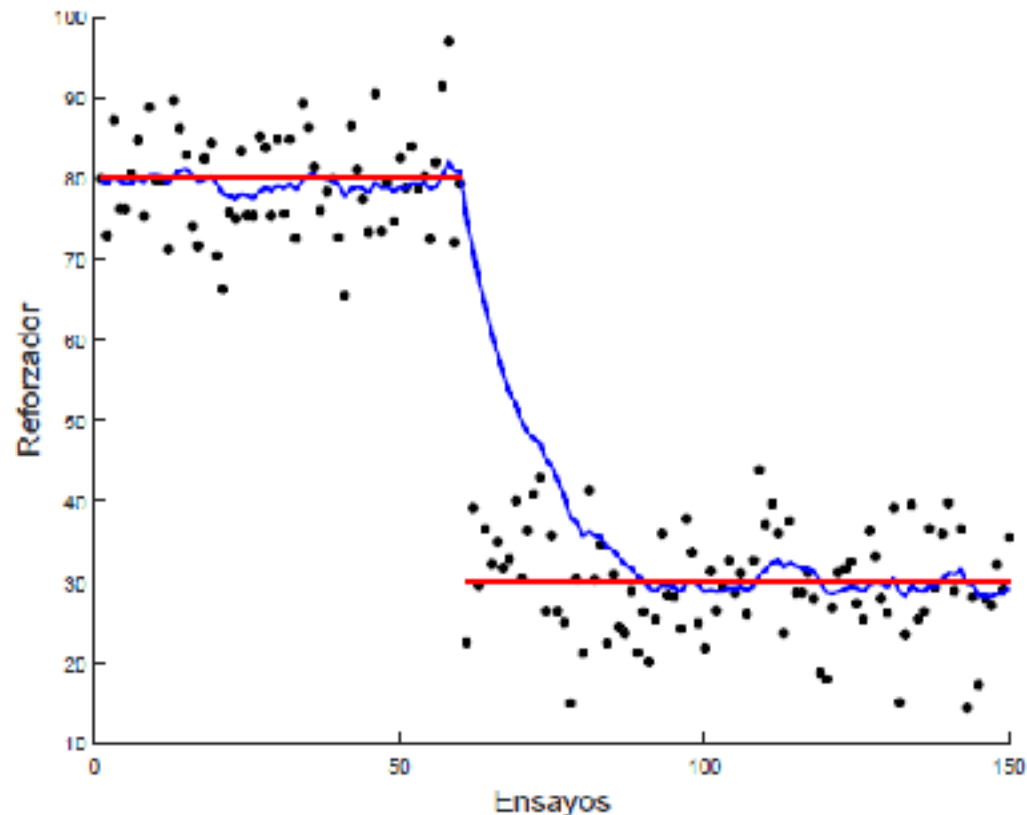
- Sin embargo, si la serie es volátil un valor de alfa cercano a 1 haría que los organismos respondieran a ruido como si fueran cambios verdaderos en los valores que generan la serie

- Un α elevada “libera” el peso de reforzadores pasados pero hace al modelo inestable



- En el caso que α sea cercana a 0, el pasado es más importante y el organismo es poco sensible a los errores en predicción. Este valor de α es muy ventajoso cuando la serie es muy volátil pues no se ajusta ante ruido, pero es desventajosa ante un cambio abrupto en una serie con poca variabilidad en la que le tomará mucho tiempo ajustar su comportamiento al nuevo valor.

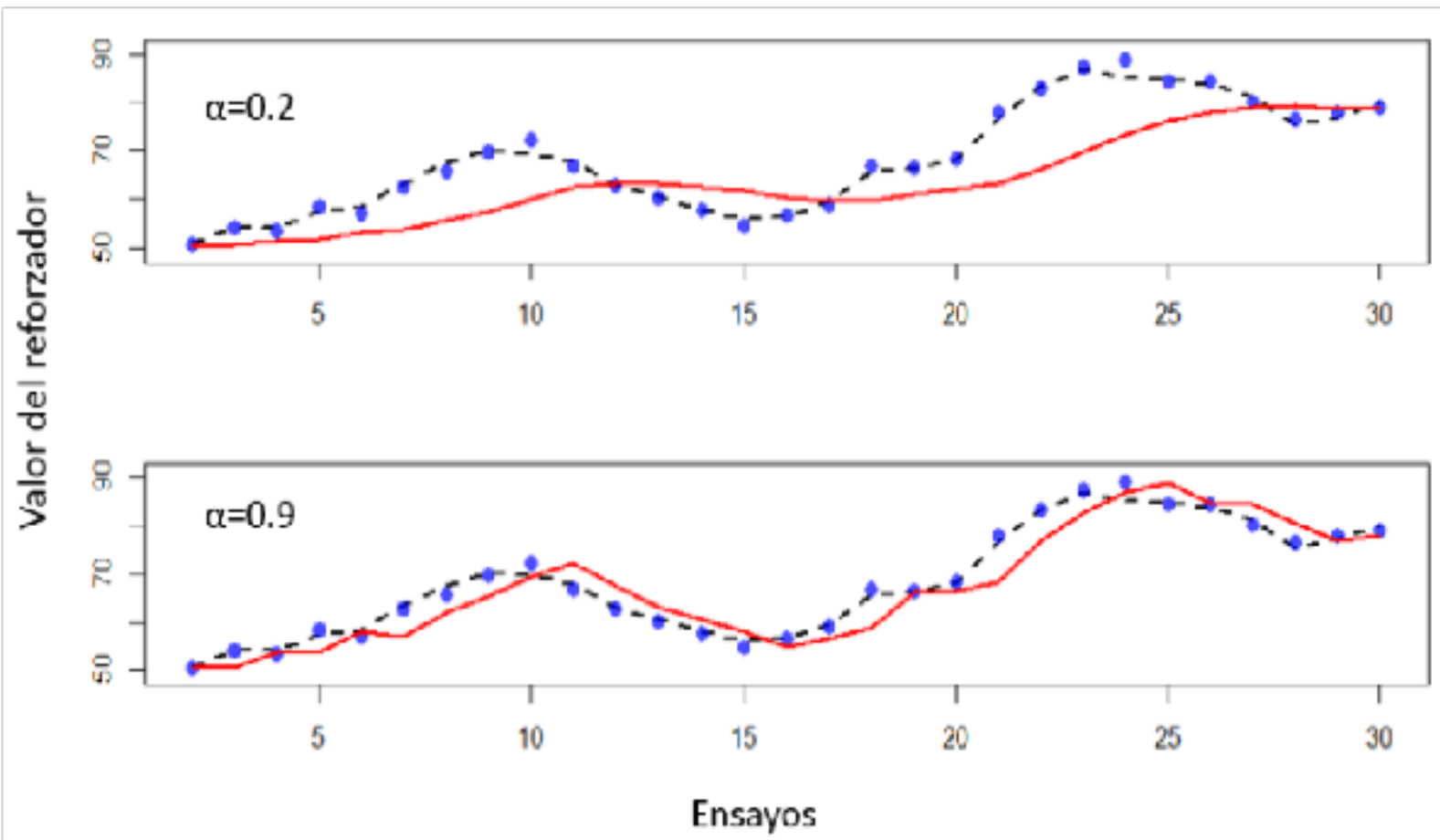
- Si α es baja, el peso de reforzadores pasados retrasa el ajuste cuando la condición cambia.



Adaptación a entornos que varían momento a momento.

- Los entornos que varían momento a momento se pueden representar como carreras aleatorias (random walk).

- Las siguientes dos figuras muestran las predicciones de un modelo de reforzamiento estándar en entornos de esta naturaleza, para dos valores del parámetro de aprendizaje α .
- Para ambos valores las predicciones del modelo de refuerzo se encuentran desfasadas de los valores verdaderos generados por el proceso de carrera aleatoria.



Experimento (Carlos Velázquez)

Se recolectaron datos de 72 sujetos.

La tarea consistió en estimar la posición de una nave espacial que orbita alrededor de la tierra.



- En cada ensayo la nave se muestra por un segundo en algún punto de su trayectoria y después desaparece. El participante debe indicar con su cursor el lugar donde creé que se moverá en el siguiente ensayo. Esta secuencia se repite por un total de cuatro condiciones de trescientos ensayos cada uno.

Un entorno que varía momento a momento

Velocidad de
cambio

$$\longrightarrow v_{t+1} = v_t + \sigma_v$$

Valor

$$\longrightarrow \mu_{t+1} = \mu_t + v_{t+1}$$





esperado

Observación

$$\longrightarrow r_{t+1} = \mu_{t+1} + \sigma_r$$

Condición	Razón		
1	0.05	0.0049	0.098
2	0.5	0.0049	0.0098
3	1	0.0049	0.0049
4	2	0.0049	0.00245

- En todas las condiciones la variante de la velocidad (señal) fue constante y las condiciones diferían en el tamaño de la variante de la observación.

Condición	$\frac{\text{señal}}{\text{ruido}}$	Representación
1	0.05	
2	0.5	
3	1	
4	2	

- Intuitivamente, la razón señal-ruido nos indica el grado en que es posible discriminar el proceso generador. Si es menor que 1, existe un mayor nivel de ruido que de señal, y por lo tanto es más difícil de discriminar; si es mayor que 1, la cantidad de señal es mayor que la de ruido, y por lo tanto, es más sencillo de discriminar.

Regla de integración

Diagram illustrating the integration rule with annotations:

Top equation: $\mu_{t+1} = \mu_t + v_{t+1} + \alpha(r_t - \mu_t)$

Annotations for the top equation:

- Valor en el ensayo t (points to μ_t)
- Tasa de aprendizaje para posición (points to α)
- Error de predicción (points to $r_t - \mu_t$)
- Valor en el ensayo t+1 (points to μ_{t+1})
- Velocidad en el ensayo t+1 (points to v_{t+1})

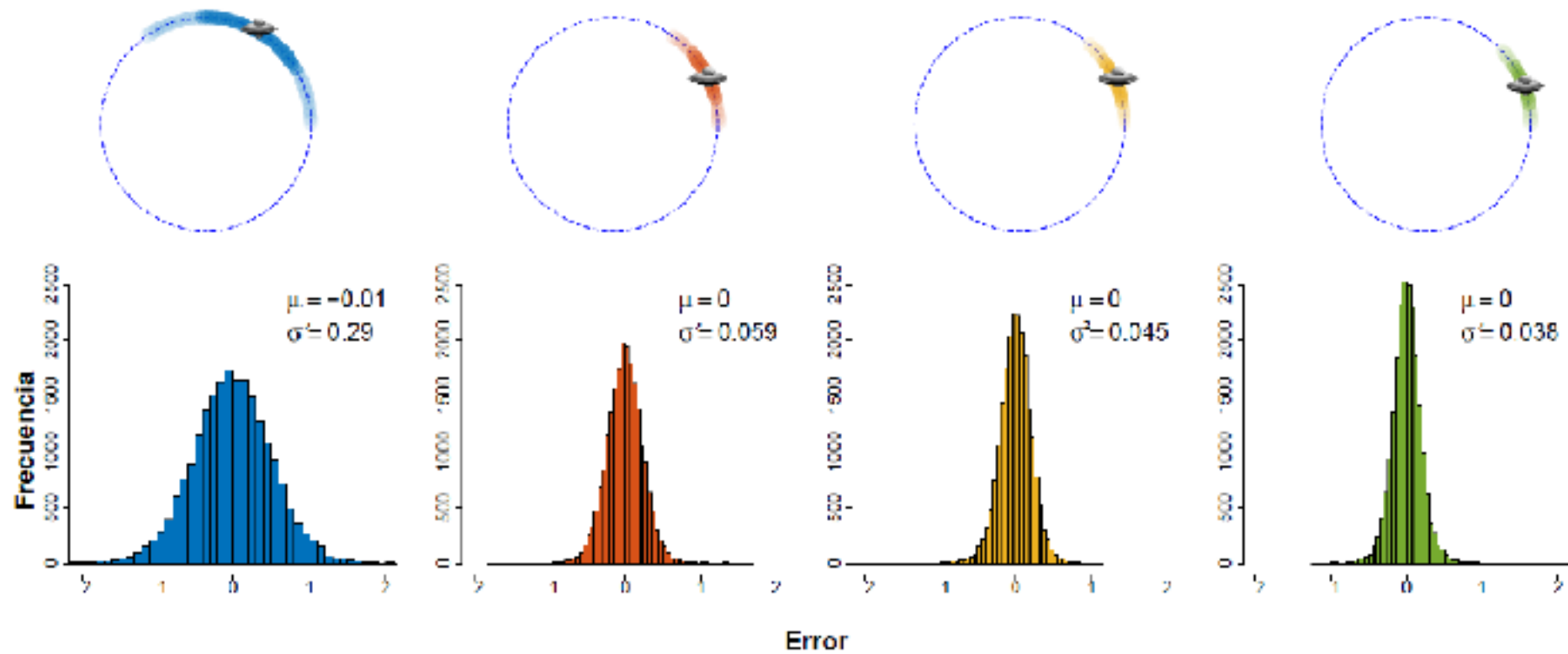
Bottom equation: $v_{t+1} = v_t + \beta(r_t - \mu_t)$

Annotation for the bottom equation:

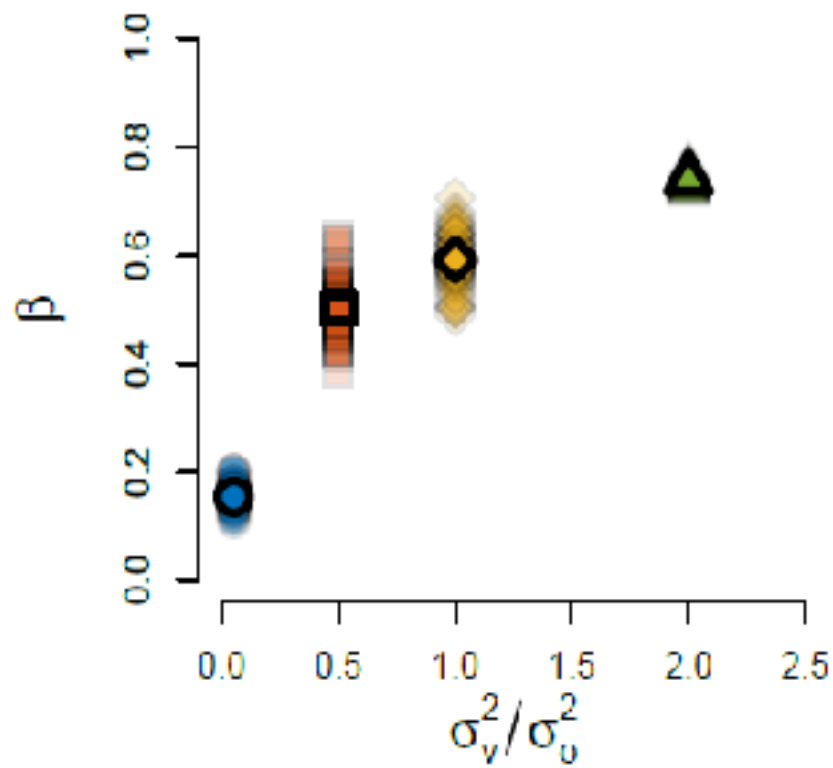
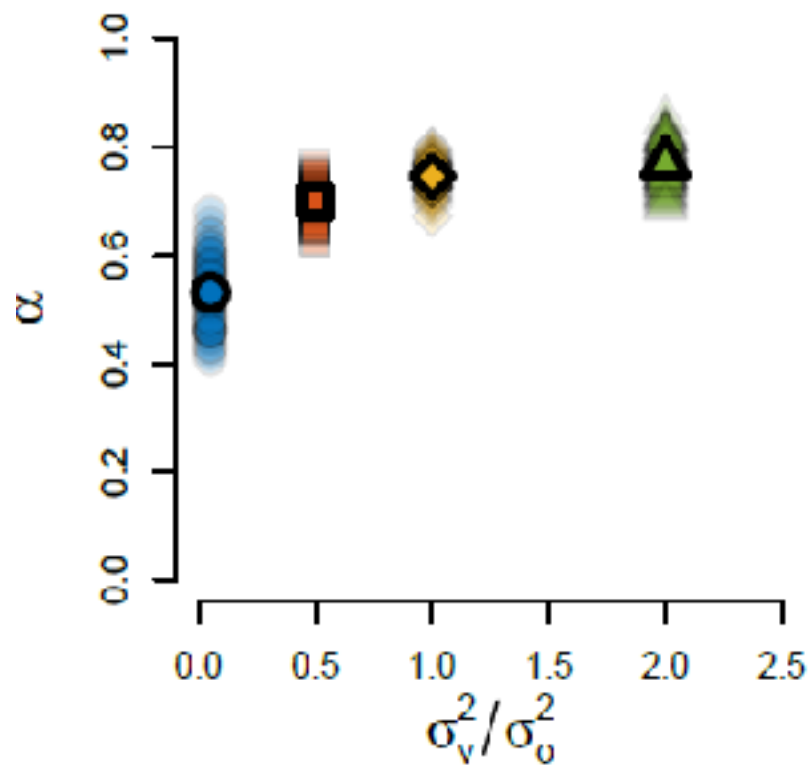
- Tasa de aprendizaje para velocidad (points to β)

A bracket connects the error term $(r_t - \mu_t)$ from the top equation to the bottom equation, indicating its shared use.

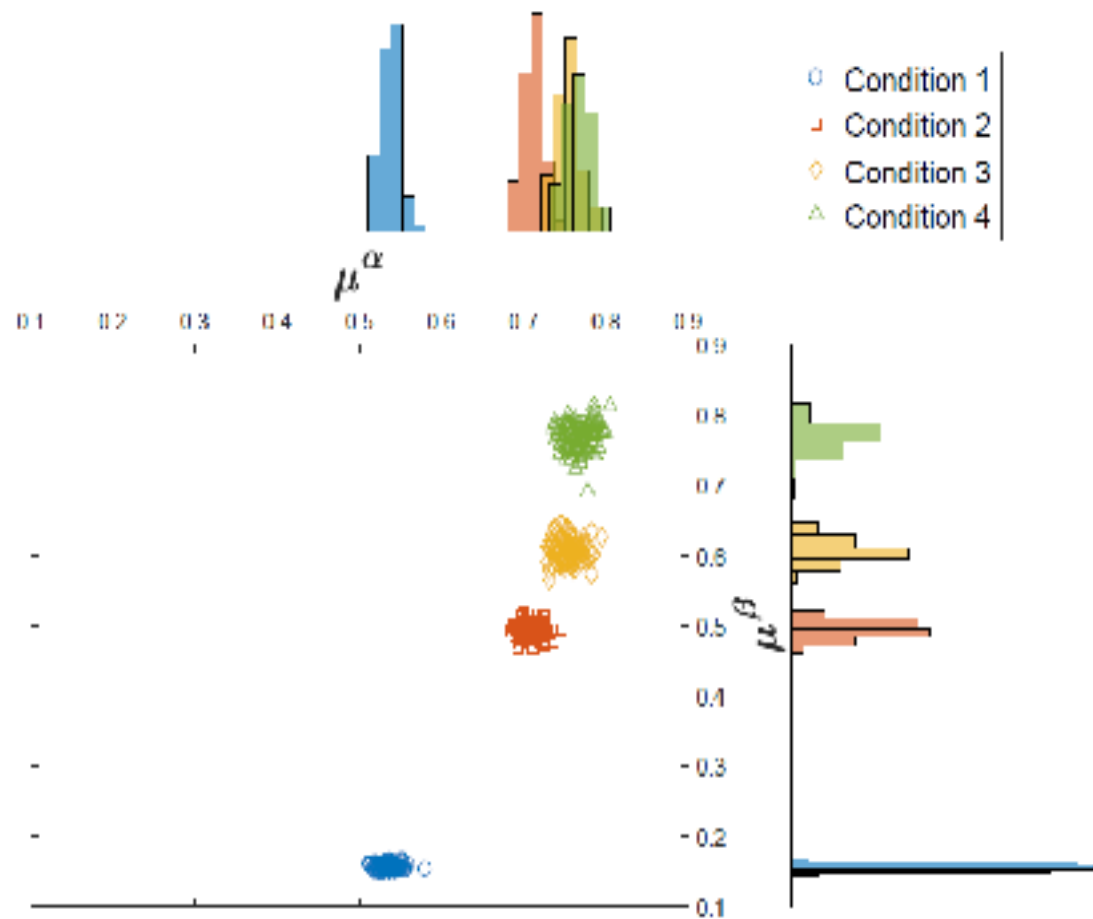
Resultados

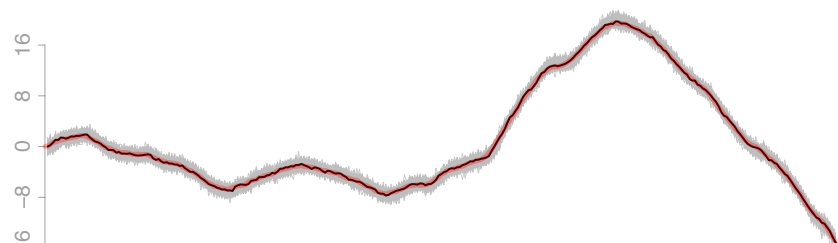
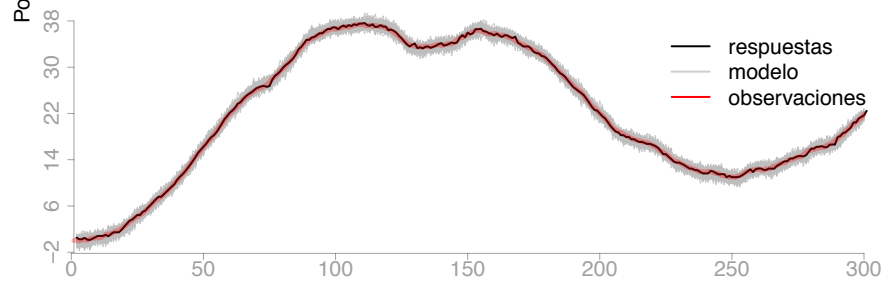
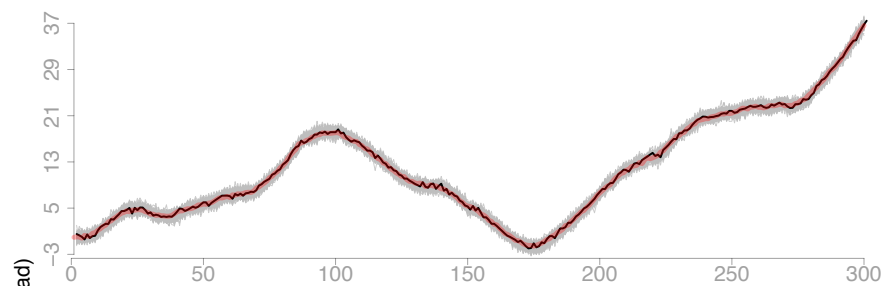
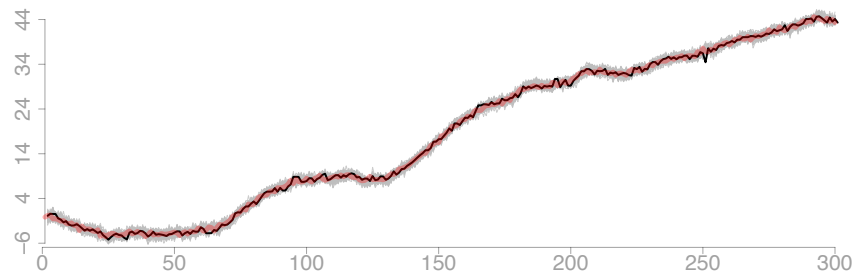


Tasas de aprendizaje

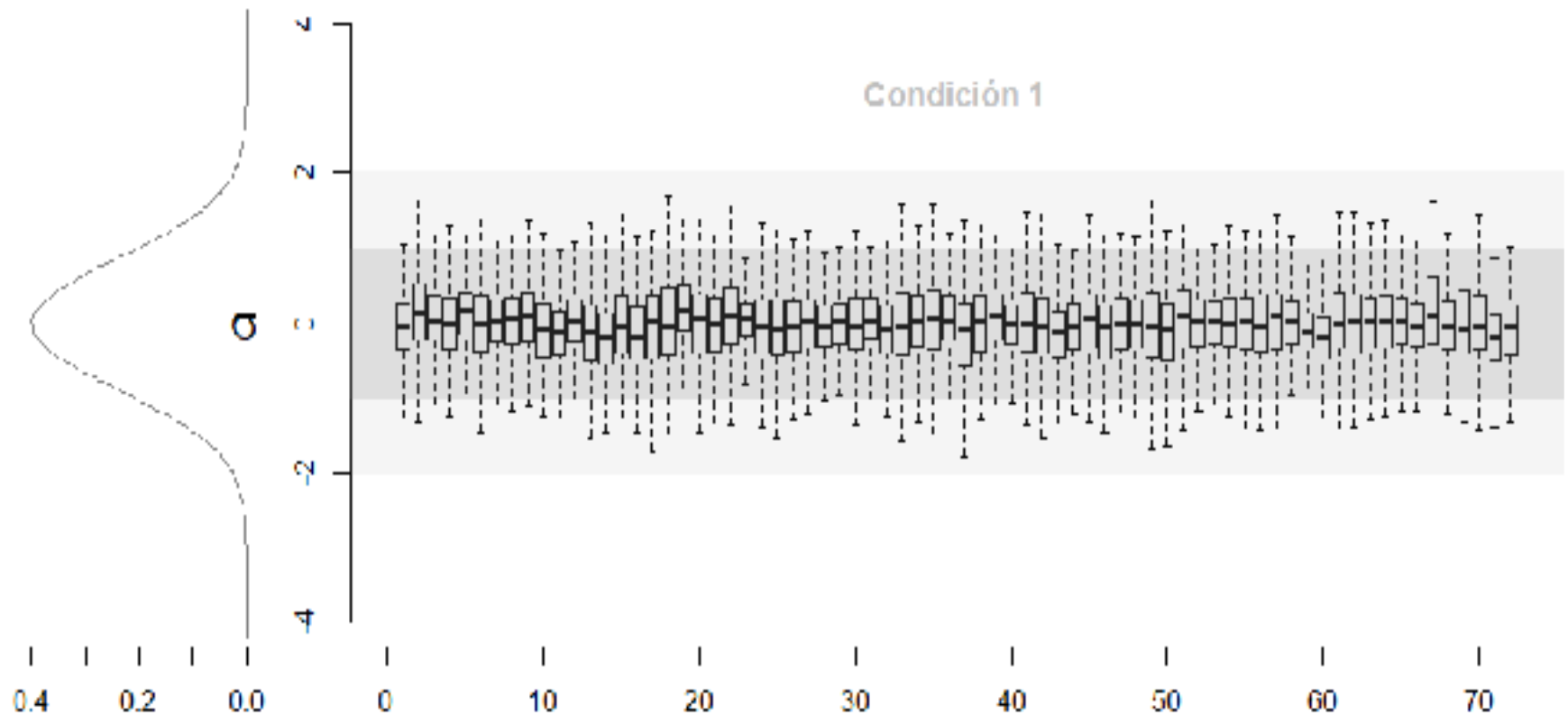


Distribuciones conjuntas de las tasas de aprendizaje



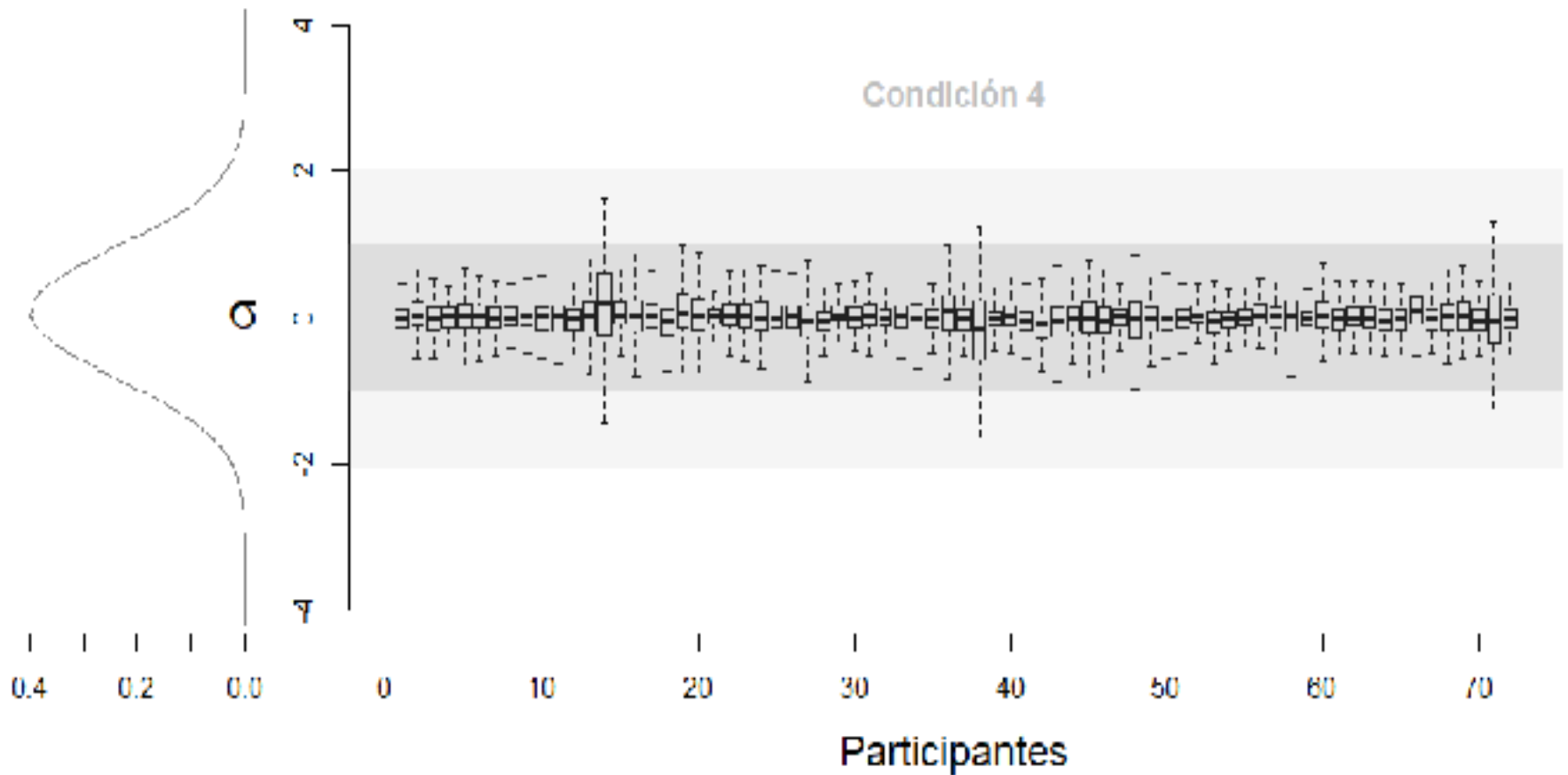


Predicciones del modelo



- Errores estandarizados entre las predicciones del modelo y las respuestas de los sujetos.

Predicciones del modelo



- Los organismos detectan propiedades de segundo orden como lo es los cambios en velocidad (tendencias)

Adaptación a Cambios en Condiciones de Refuerzo con Diferentes Incertidumbres

- 12 palomas
- Programa Mult RV 30 – Ext, componentes 30 segs. Sesiones de 60 componentes.
- Condición 1 Componente rico alterna día a día
- Condición 2 Componente rico alterna a mitad de sesión
- Condición 3 Componente rico asignado probabilísticamente día a día ($p=.5$)

- Condición 4 Componente rico siempre el mismo al inicio de la sesión y cambia aleatoriamente ($p = .5$) a mitad de la sesión
- Condición 5 Componente rico asignado aleatoriamente al inicio de cada sesión y cambia aleatoriamente sin señal a mitad de la sesión. Condición con mayor incertidumbre.

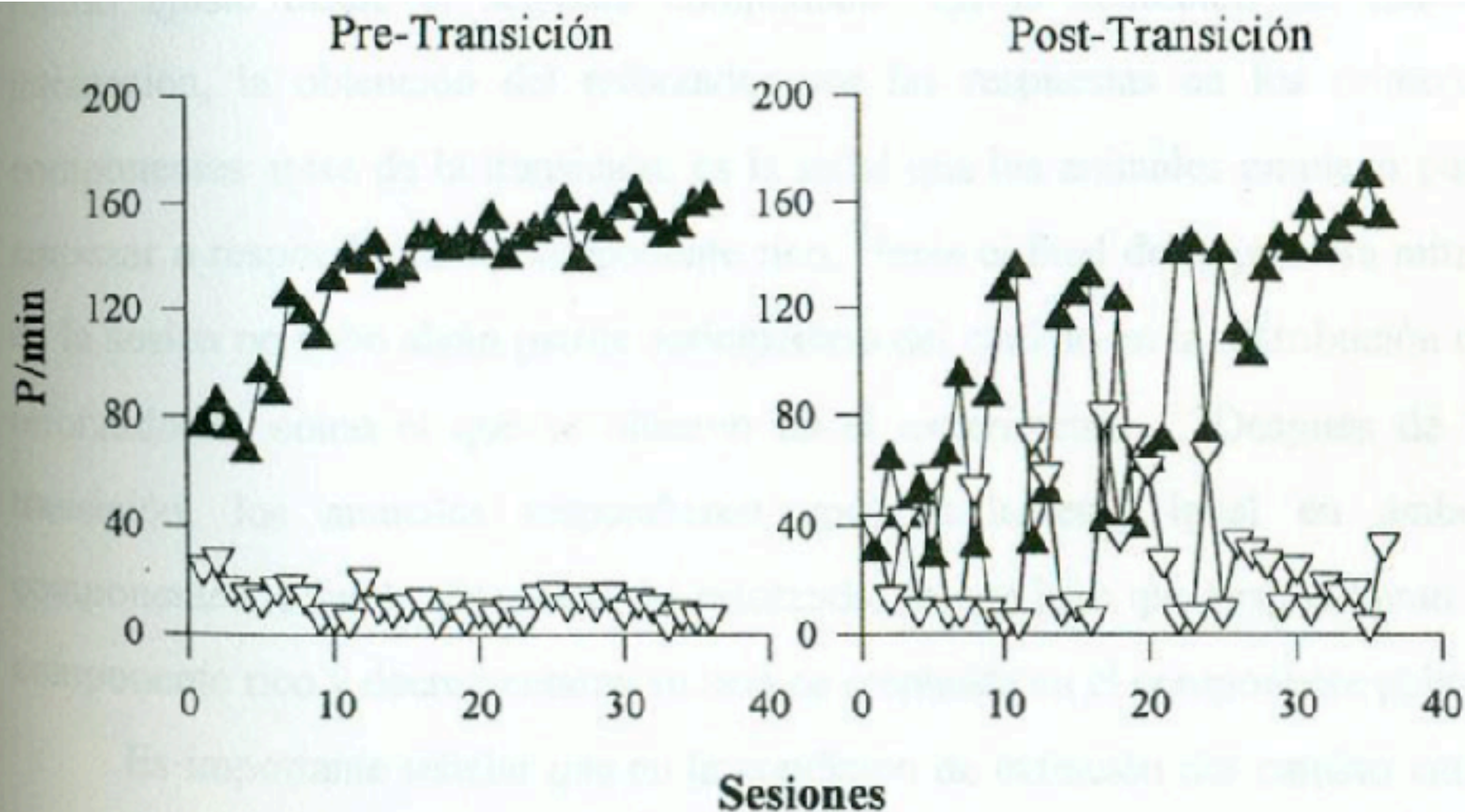


Fig. 5b. Ejecución asintótica promedio en cada componente del programa múltiple, durante la condición de cambio aleatorio en la segunda mitad de la sesión experimental. Los símbolos cerrados representan las respuestas al componente rico, y los símbolos abiertos las respuestas al componente pobre.

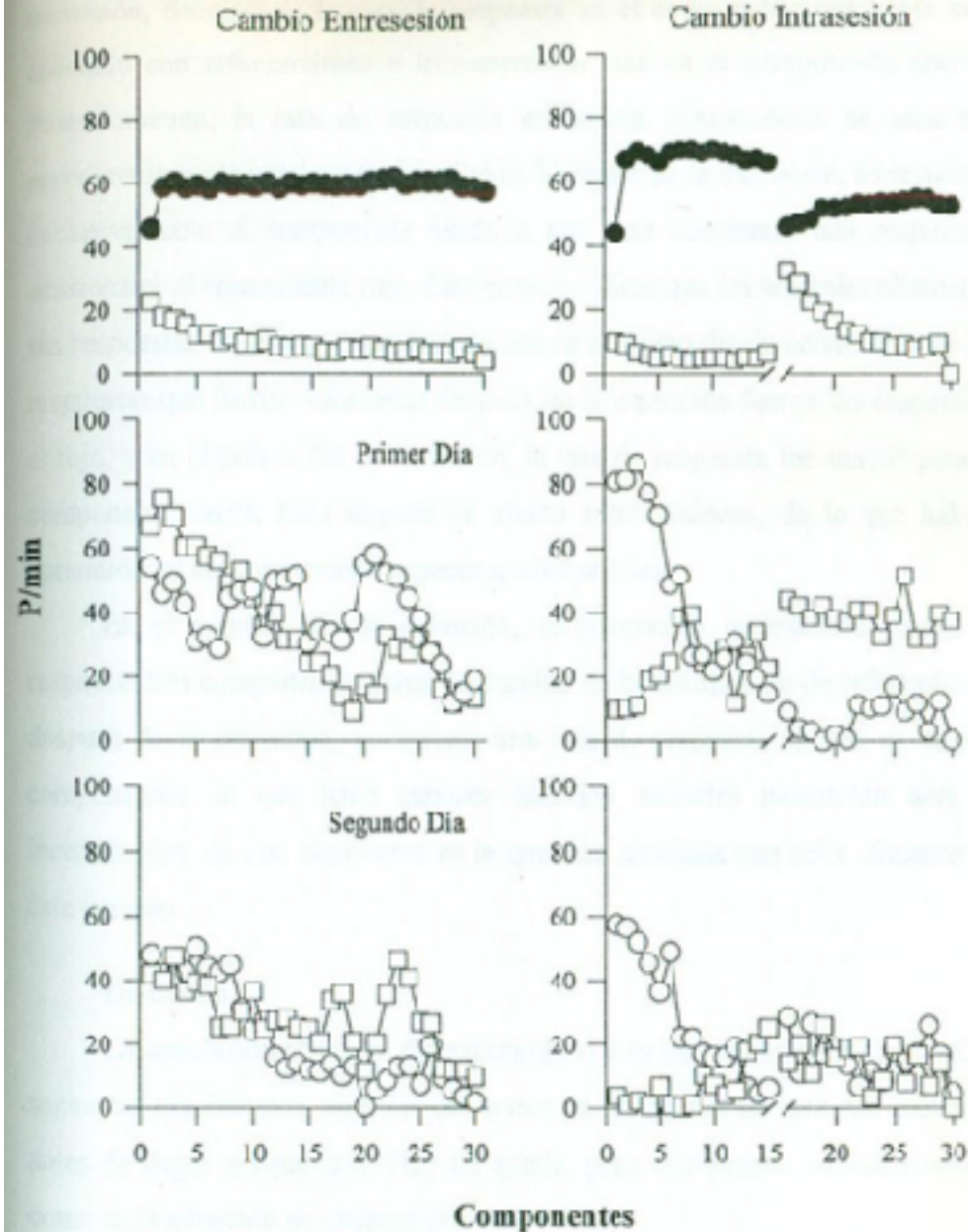


Fig. 6. Tasas de respuesta promedio al final del entrenamiento (panel superior), y en los dos primeros días de extinción (paneles inferiores), durante el cambio entre sesión (panel izquierdo) y cambio intrasesión (panel derecho) en la condición aleatoria diaria y en la segunda mitad de la sesión experimental. Ver detalles en el texto.

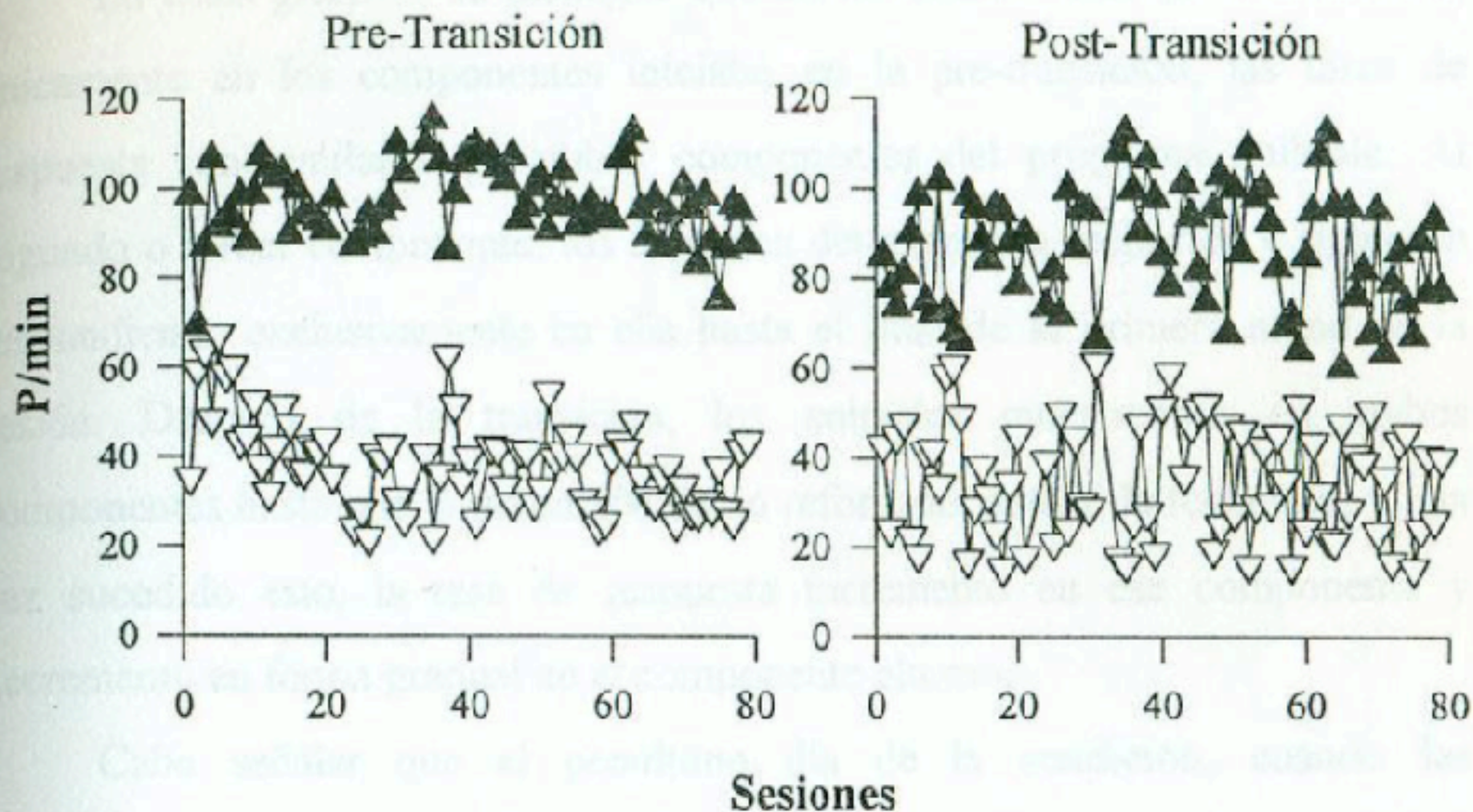


Fig.7. Tasas absolutas promedio para cada componente del programa múltiple, durante la condición transición probabilística intrasesión. Las respuestas al componente rico están representadas por los símbolos cerrados y las respuestas al componente pobre, por símbolos abiertos.

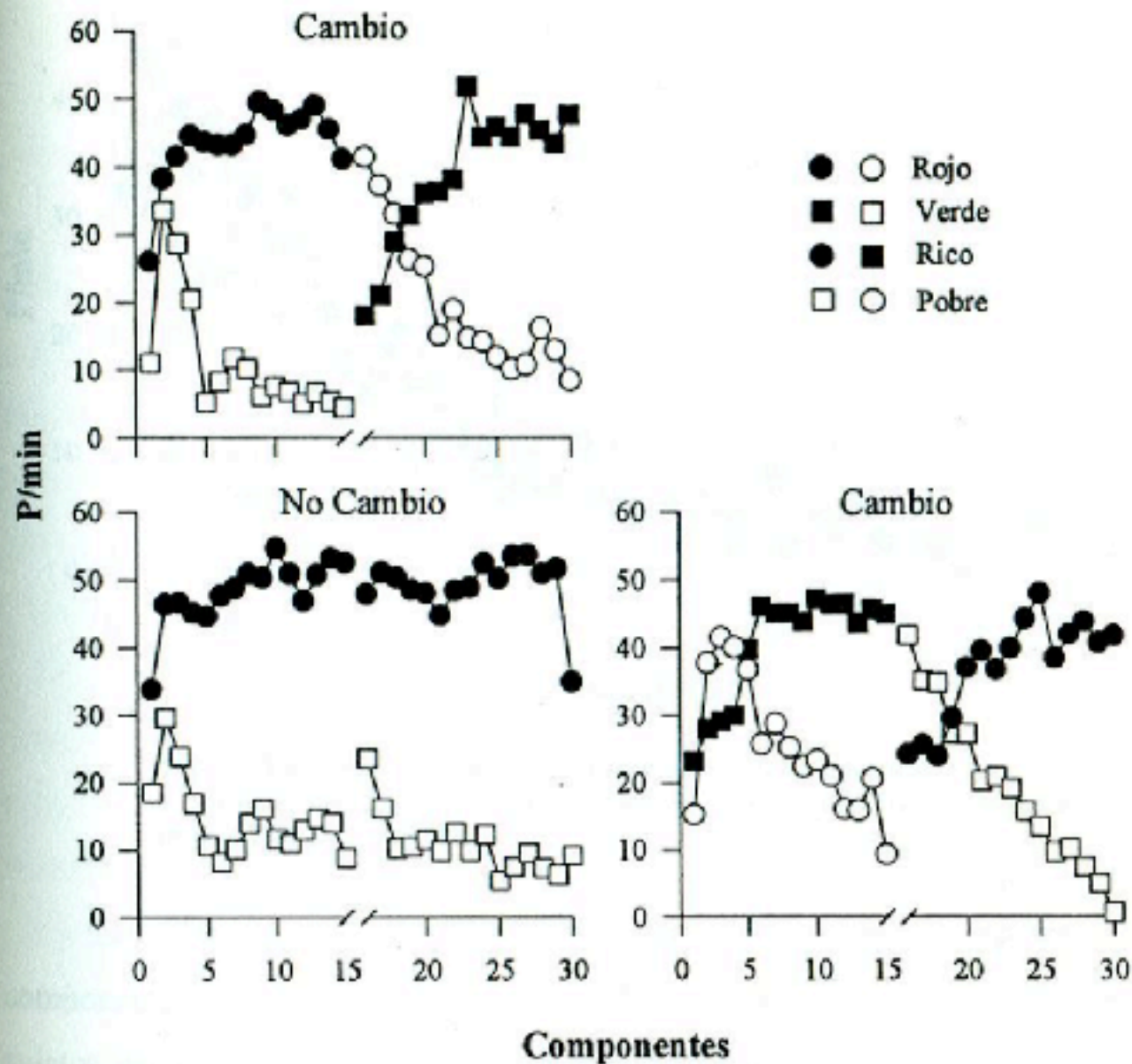
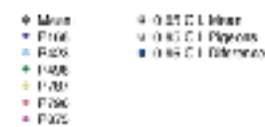
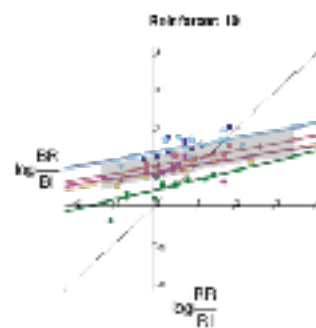
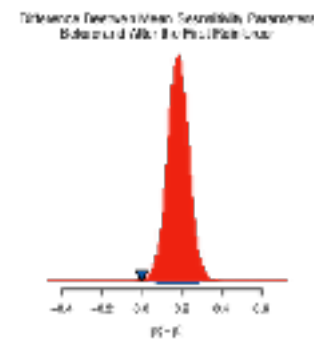
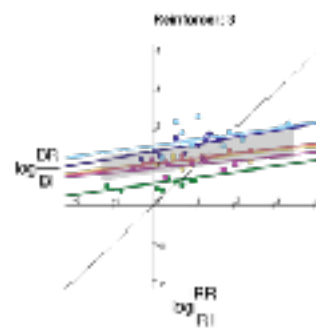
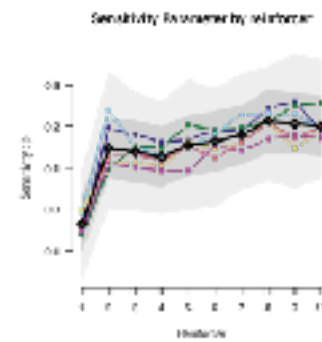
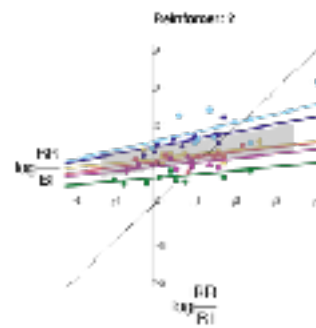
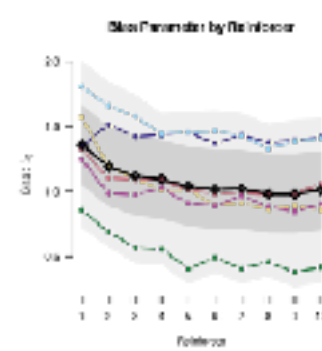
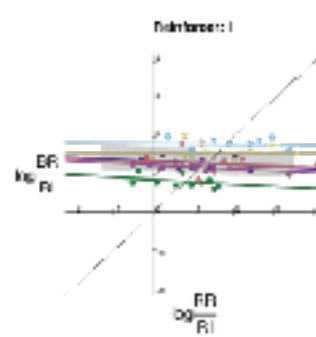
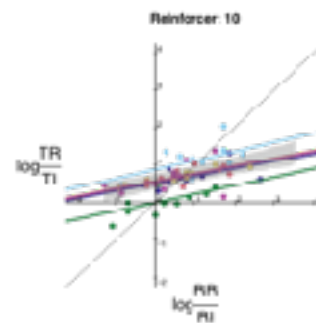
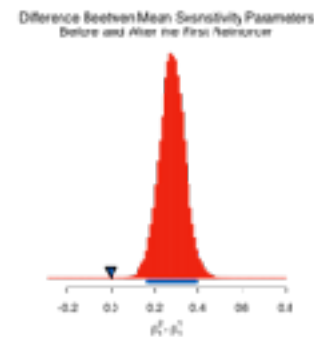
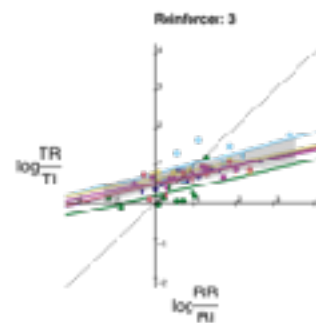
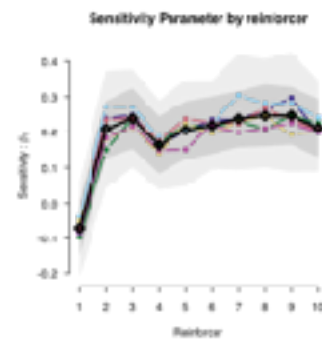
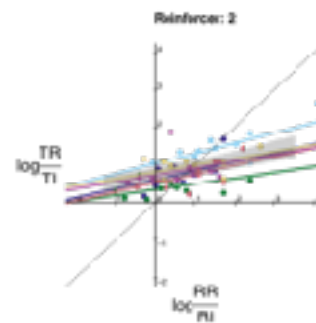
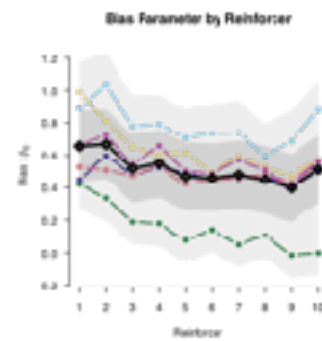
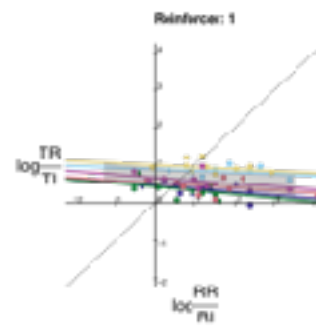


Fig. 8. Tasas de respuesta promedio al final del entrenamiento durante los tres últimos días de la condición transición probabilística intrasesión, cuando hubo cambios o no cambios en la distribución de reforzamiento de la tecla rica.

- Adaptación a Incertidumbre en programas conc $RRx - Rly$ (Findley con tecla de cambio)
- 6 pichones
- 10 pares de valores presentados cada sesión presentes aleatoriamente por 10 refuerzos

	programas originales			programas aleatorios	
	BLOQUE 1			BLOQUE 1	
PAR	VR	VI	PAR	FR	FI
0	30	7,5	0	0,03333333	0,13333333
1		15	1		0,06666667
2		30	2		0,03333333
3		60	3		0,01666667
4		120	4		0,00833333
	BLOQUE 2			BLOQUE 2	
PAR	VR	VI	PAR	FR	FI
5	15	60	5	0,06666667	0,01666667
6	30		6	0,03333333	
7	45		7	0,02222222	
8	60		8	0,01666667	
9	120		9	0,00833333	





abouzasr@gmail.com

<http://bouzaslab25.github.io>

@abouzasr