



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE PSICOLOGIA

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO "APRENDIZAJE DE LAS ESTRATEGIAS REPETIR Y CAMBIAR BAJO UN PROCEDIMIENTO DE REFORZAMIENTO ALEATORIO"



ALUMNOS PROFESIONALES

T E S I S QUE PARA OBTENER EL TITULO DE: LICENCIADO EN PSICOLOGIA P R E S E N T A : DIANA HERRERA ESPINOSA

DIRECTOR DE TESIS: MTRO. GUSTAVO BACHA MENDEZ REVISOR DE TESIS: MTRO. RAUL AVILA SANTIBANEZ SINODALES: DR. FLORENTE LOPEZ RODRIGUEZ MTRO. JULIO ESPINOSA RODRIGUEZ MTRO. OSCAR VLADIMIR ORDURA TRUJILLO



MEXICO, D. F.,

2002

TESIS CON FALLA DE ORIGEN



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

PAGINACION DISCONTINUA

Sí, pero quién nos curará del fuego sordo, del fuego sin color que corre al anochecer por la rue de la Huchette, saliendo de los portales carcomidos, de los parvos zaguanes, del fuego sin imagen que lame las piedras y acecha en los vanos de las puertas, cómo haremos para lavarnos de su quemadura dulce que prosigue, que se aposenta para durar aliada al tiempo y al recuerdo, a las sustancias pegajosas que nos tienen de lado, y que nos arderá dulcemente hasta calcinarnos.

...Todo es escritura, es decir fábula. ¿Pero de qué nos sirve la verdad que tranquiliza al propietario honesto? Nuestra verdad posible tiene que ser invención, es decir escritura, literatura, pintura, escultura, agricultura, piscicultura, todas las turas del mundo. Los valores, turas, la santidad, una tura, la sociedad, una tura, el amor, pura tura, la belleza, tura de turas.

... del napolitano que se pasó años sentado a la puerta de su casa mirando un tornillo en el suelo. Por la noche lo juntaba y lo ponía debajo del colchón. El tornillo fue primero risa, tomada de pelo, irritación comunal, junta de vecinos, signo de violación de los deberes cívicos, finalmente encogimiento de hombros, la paz, el tornillo fue la paz, nadie podía pasar por la calle sin mirar de reojo el tornillo y sentir que era la paz... A lo mejor el napolitano era un idiota pero también pudo ser el inventor de un mundo. Del tornillo a un ojo, de un ojo a una estrella...

Nadie nos curará del fuego sordo, del fuego sin color que corre al anochecer por la rue de la Huchette. Incurables, perfectamente incurables, elegimos por tura el Gran Tornillo, nos inclinamos sobre él, entramos en él, volvemos a inventarlo cada día, ... inventamos nuestro incendio, quizá las palabras envuelvan esto como la servilleta al pan y dentro esté la fragancia, la harina esponjándose, el sí sin el no, o el no sin el sí, el día sin Manes, sin Ormuz o Arriman, de una vez por todas y en paz y basta.

Julio Cortázar,
RAYUELA

Quien practica el no-obrar
es efectivo en la inacción,
halla sabor en lo insípido,
ve lo grande en lo pequeño,
ve lo mucho en lo poco.
Responde al odio con la virtud,
trata lo difícil cuando aún es fácil,
trata lo grande cuando aún es pequeño
porque todas las dificultades en el mundo
comienzan con la facilidad
y todas las grandezas
por la pequeñez...

Lao Tse
Verso LXIII, Tao Te King

A

MAVEKA

mis tres acompañantes

A QUIEN PIDIÓ QUE CONFIARA

(. . . sí confío)

AGRADECIMIENTOS

ALOS QUE VIVEN CONMIGO LAS MÁS EXTRAÑAS AVENTURAS

Mamá: por tu apoyo incondicional, tus consejos, tu constante inventiva, los desvelos, por las canciones y la poesía, los secretos divvca, por compartirnos La Respuesta.

Vernon: por ser tú y contagiarnos, por tu música, tu literatura, tus bromas, pero sobretodo por las Turas !!

Karla: por lo que cada día me enseñas, por ser tan diferente y tan parecida, por tu inagotable creatividad, por TSLO; por compartir esa Historia Interminable con toques de magia.

ALOS AMIGOS

Martha: por todos los colores que inventamos, porque a pesar de los años y la distancia estás ahí.

Offir: por las travesuras en el Teatro, los múltiples delirios compartidos (Ω), por las conversaciones sin palabras, por intercambiar tu fisingeniemate con mi bioquimipisco.

Capeto: por las estrellas de neutrones, el guiñol, por tu visión del mundo, por todo lo que compartimos y aprendimos juntos.

ADN y FUS: por comprender, por escuchar, por compartir, por los jalones de orejas, por hacer más fáciles los momentos más difíciles . . .

Alice, Capeto, Javier, Bernardo: por el amor a la divulgación científica, porque hemos mirado el pasado al contemplar las estrellas.

Adriana, Ayleen, Cristina, Fusac, Jorge, Ranier y Tania: por todos los cursos, por compartir un incomprensible gusto por la investigación pero principalmente por su Amistad.

Mike, Héctor, Isaac, Victor, Ray: por compartir el dōjo y hacer menos duros los golpes de la vida

ALOS MAESTROS

- Los que sesgaron mi camino: Martha A, Sonia, Alma, Ricardo I, Marivi, Ma. Luz, Raúl, Julio, Ricardo M, Gustavo . . .
- Los presentes:

Gustavo: por adentrarme en el mundo de la investigación, por dejarme ser parte de tu Laboratorio, por los consejos, los regaños, las oportunidades, la confianza, por todo lo que me has enseñado, pero esencialmente por ser mi Maestro.

Raúl: por tus comentarios, tu amistad, por toda tu ayuda y porque a pesar de los desacuerdos revisaste conmigo c-a-d-a palabra de esta Tesis.

Julio: por tu apoyo, por compartir tu investigación y tus ideas, porque se puede "contar" contigo, porque eres un maestro pero sobretodo un amigo.

- A los maestros que sólo conozco a través de sus enseñanzas:

Baden Powell

Micael

Morihei Ueshiba

A LOS SINODALES

Mtro. Gustavo Bachá, Mtro. Raúl Ávila, Dr. Florente López, Mtro. Julio Espinosa y Mtro. Vladimir Orduña, por su tiempo, su paciencia, todos sus comentarios y aportaciones a este trabajo.

A LOS COMPAÑEROS DE LABORATORIO

Adriana, Alejandro, Dulce, Erika y Livia por compartir la investigación de cada día

A LOS ÁNGELES...

- Los de la Prepa (sin ellos no estaría aquí): Caro, Atzimba, Sonia, Offir...
- Los que me adoptaron en su hogar y me brindaron apoyo y compañía durante esos meses: Sra. Rosa, Sr. Arturo, Nayelli y Arturo (Familia Becerril)
- Los de la carrera, los Congresos y durante la realización de la Tesis: Capeto, Julio, ADN, FUS, Gustavo...
- Los que me brindaron su apoyo, su espacio y su paciencia para llegar hasta este momento: Cely, Ama y Sofi.
- Los que hicieron posible la impresión de esta Tesis: Mónica, Lilia - David.

A quienes dieron su vida en nombre de la ciencia

ratas DN1, DN2, DN3 y DN5

RESUMEN

En gran parte de la literatura se ha encontrado que, conforme a la Ley del Efecto, los animales tienden a repetir una respuesta cuando es seguida de la entrega de un reforzador (estrategia *ganar-repetir*) y que tienden a emitir una respuesta distinta después de la ausencia del mismo (estrategia *perder-cambiar*). En otros trabajos, también se ha mostrado que los animales aprenden a responder de acuerdo a las estrategias de respuesta opuestas, *ganar-cambiar* y *perder-repetir*. La evidencia de que los animales aprenden a responder con cualquiera de estas cuatro estrategias surge de trabajos que han sido realizados para estudiar diversos fenómenos conductuales, pero no para estudiar el posible mecanismo mediante el cual los animales aprenden estas estrategias de respuesta. El objetivo del presente trabajo fue desarrollar un procedimiento para estudiar la adquisición de las estrategias ganar-repetir y ganar-cambiar cuando se mantiene constante la conducta y la consecuencia que determinan la respuesta correcta en el siguiente ensayo. Se entrenó a cuatro ratas a responder en un procedimiento de ensayos discretos en el cual las secuencias de dos respuestas: izquierda-izquierda (II) y derecha-derecha (DD) eran reforzadas de manera aleatoria. Hubo dos condiciones, en una de ellas los sujetos eran reforzados por responder de acuerdo a la estrategia ganar-repetir y en la otra por responder de acuerdo a ganar-cambiar. De manera aleatoria se presentaron *ensayos de estrategia* en los que los sujetos podían ganarse un reforzador con una probabilidad de 1.0 si repetían (condición ganar-repetir) o si emitían la secuencia opuesta (condición ganar-cambiar) a la previamente reforzada. Para dos de los sujetos, estos ensayos se presentaron después de que la secuencia II era reforzada, y para los otros dos después de ser reforzada DD. Se utilizó un diseño tipo A-B-A-B y las condiciones se contrabalancearon entre los sujetos. Los resultados muestran que para la condición ganar-repetir, todos los sujetos tuvieron una ejecución superior al 80% de respuestas correctas, mientras que para la condición ganar-cambiar la mitad de los sujetos alcanzó este criterio de ejecución.

ÍNDICE

RESUMEN

INTRODUCCIÓN	1
MÉTODO	13
Sujetos	13
Aparatos	13
Procedimiento	14
Entrenamiento Preliminar	14
Entrenamiento	15
Fases Experimentales	16
RESULTADOS	23
DISCUSIÓN	30
REFERENCIAS	39
ANEXO	46

La conducta de los organismos sufre cambios constantemente. Desde un punto de vista molecular, se puede clasificar la modificación del flujo conductual en repeticiones o cambios de la respuesta que un animal emite de un momento determinado al siguiente. Es decir, cuando un organismo emite una conducta, lo que se puede observar en el siguiente instante es que el sujeto vuelve a emitir esa conducta (repetir), o que realiza una distinta (cambiar). Muchas variables pueden influir para que los animales repitan o cambien su comportamiento; algunas de estas variables pueden ser la fatiga, cambios químicos del ambiente, el estado motivacional del sujeto, etc. (Domjan, 1999). Sin embargo, en ocasiones las consecuencias del comportamiento determinan la conducta que emitirá un organismo en el futuro (Skinner, 1981).

Thorndike (1911) realizó una investigación sistemática acerca de cómo se modifica la conducta en relación a sus consecuencias. Como resultado de sus investigaciones postuló la Ley del Efecto, la cual indica que la conducta se “estampa” o se conecta más firmemente con el contexto en el que ocurrió cuando le sigue una condición *placentera*, de manera que ante las mismas circunstancias tiene una mayor probabilidad de ser emitida, y que tal conexión se debilita cuando le sigue un evento *displacentero*, de forma que ante las mismas condiciones es menos probable que ocurra (Thorndike, 1911, 1927).

En diversos experimentos (e.g., Evenden & Robbins, 1984; McElroy & Neuringer, 1990; Morgan, 1974; Neuringer, 1991; Schwartz, 1992;

“... si se repiten las circunstancias, es probable que se repitan los movimientos musculares que fueron seguidos por el éxito: lo que al principio era una combinación accidental de movimientos, ahora será una combinación que tendrá una probabilidad considerable.”

Hebert Spencer (1855)

Shimp, 1966, 1967, 1969; Silbelberg, Hamilton et al., 1978; Silbelberg & Williams, 1974) se ha encontrado que las consecuencias de la conducta tienen un efecto en la respuesta que se emite y que la manera en la que las consecuencias influyen en la conducta es congruente con la Ley del Efecto. Es decir, es más probable encontrar un patrón de repetición después de que una respuesta es reforzada y un patrón de cambio después de una no reforzada. Asimismo, se ha demostrado que la probabilidad de que un sujeto emita una respuesta particular incrementa cuando la presentación de una recompensa depende de la emisión de esa respuesta (Schwartz, 1989), y que la probabilidad de emisión de una conducta se modifica dependiendo de la manera en la que se programa la entrega de las consecuencias, por ejemplo, la magnitud de la recompensa, su calidad, demora con la que se presenta, etc. (Mackintosh, 1974; Schwartz, 1989; Williams, 1988). Cuando el cambio que se observa en la conducta, por ejemplo un incremento o decremento en su probabilidad de emisión, es duradero y resulta de la experiencia previa con estímulos y respuestas semejantes, entonces se dice que un organismo ha aprendido (Domjan, 1999).

Para que ocurra el aprendizaje son fundamentales dos condiciones. Una de ellas es la *contigüidad temporal* entre la conducta y la presentación de sus consecuencias (Blough, 1959; Staddon, 1988; Staddon & Zhang, 1989, 1991). La segunda condición es la existencia de una relación de *contingencia* entre la conducta y la consecuencia (Bloomfield, 1972; Staddon, 1988; Staddon & Zhang, 1989, 1991). Es decir, una relación de dependencia en la que existe

"El aprendizaje presupone el fracaso. El aprendizaje sólo supondrá una ventaja si existe una tendencia a hacer algo y la criatura falla en cualquier número de veces antes de tener éxito."
Grey Walter (1986)

una mayor probabilidad de que se presente la consecuencia después de que se emite una conducta particular y una menor probabilidad de que se presente la consecuencia después de la emisión de cualquier otra conducta.

La contigüidad temporal es necesaria para que ocurra el aprendizaje pero no es suficiente, es decir, cuando una recompensa se presenta de manera inmediata pero no contingente a la ejecución de una conducta específica, ésta no se aprende (Bloomfield, 1972; Staddon, 1988; Staddon & Zhang, 1989, 1991). Por otra parte, cuando existe una relación de contingencia entre la emisión de una conducta y la presentación de la recompensa, los animales aprenden dicha relación aún si la consecuencia se entrega de manera demorada (Lattal & Gleenson, 1990; Dickinson, Watt & Griffiths, 1992), siempre y cuando el momento de su presentación no sobrepase un intervalo que pueda afectar la asociación entre la ejecución de la conducta y su consecuencia (Fetterman & Stubbs, 1982; Kramer, 1982; Shimp, 1976a, 1981, 1982).

En relación a los patrones de repetición y cambio de la respuesta que se emite de un momento determinado al siguiente, cuando un organismo aprende que la emisión de una conducta lleva a la obtención de un reforzador el animal la repite. En cambio, cuando lo que aprende un organismo es que la emisión de una conducta lleva a la ausencia de reforzador, el sujeto deja de ejecutarla (Catania, 1997; Lieberman, 1993; Schwartz, 1989). Se ha atribuido la aparición de estos patrones conductuales al efecto que tiene el reforzador

"In order for a signal to transmit information about an environmental event, it is necessary not only for it to concur with that event, but also for its absence to indicate the absence of the event."

T. M. Bloomfield (1972)

sobre la conducta que se acaba de emitir. Algunos autores han sugerido que el reforzador fortalece la respuesta a la que es contingente (Nevin, 1974, 1979; Skinner, 1938) o la selecciona (Baum, 2001; Blue, 1979; Hull, Langman & Glenn, 2001; Skinner, 1981; Staddon, 1975, 1983), de tal forma que la presentación del reforzador produce la repetición o estereotipia de la respuesta; mientras que la omisión del reforzador produce un patrón de cambio o variabilidad conductual (Schwartz, 1982). Sin embargo, Neuringer (1992, 1993), Machado (1992) y Page y Neuringer (1985) mostraron que la variabilidad conductual puede caer bajo el control del reforzador. Por lo que se dice que el reforzador produce tanto la selección o fortalecimiento de instancias de conducta como variabilidad conductual (Neuringer, 1992, 1993). El hallazgo de que los animales pueden aprender a emitir una respuesta distinta a la que es seguida de la entrega de reforzador, hace interesante estudiar cómo se aprenden los patrones de respuesta que no se derivan directamente del efecto que se dice que tiene el reforzador sobre la conducta. Estos patrones son aprender a repetir una respuesta cuando no es seguida de la entrega de reforzador y aprender a realizar una respuesta distinta después de que una determinada respuesta es reforzada. Algunos autores (Neuringer, Kornell & Olufs, 2001) mencionan que el reforzador parece tener dos tipos de efecto, uno de contigüidad y otro de contingencia. Se dice que el efecto de contigüidad genera una tendencia a repetir aquella respuesta seguida de la entrega de una recompensa; mientras que el de contingencia provee de información acerca de cuál es la respuesta que lleva a la obtención del reforzador (Bloomfield, 1972). Por lo que se ha propuesto

“...reinforcement serves the dual functions of selecting instances
and concurrently engendering variability”
Allen Neuringer (1993)

que patrones conductuales como cambiar después de que una respuesta es reforzada se pueden adquirir por el efecto de contingencia.

Un procedimiento que se puede usar para estudiar la ejecución de los diferentes patrones de respuesta mencionados (repetir la respuesta previa o cambiarla ya sea después de la entrega de reforzador o de la ausencia del mismo), puede ser aquel en el que se entrene a los animales a emitir dos respuestas – la respuesta A y la respuesta B – y que después de la emisión de una de ellas (por ejemplo, B) la entrega del reforzador se haga contingente a la ejecución de la misma conducta que en el ensayo previo (*repetir* B), o a la emisión de la otra respuesta (*cambiar* a la respuesta A). El estímulo discriminativo que indica en qué ensayos se debe repetir y en cuáles cambiar, es la consecuencia de la conducta emitida en el ensayo previo: la entrega de reforzador o la ausencia del mismo. De esta manera, la entrega de recompensa por la ejecución de una respuesta en un ensayo particular (ensayo n), dependerá de la conducta junto con sus consecuencias ocurridas en el ensayo previo (ensayo $n - 1$).

En la Figura 1 se muestra el ejemplo de un procedimiento en el que se refuerza un patrón de cambio después de que una respuesta es seguida de la entrega de reforzador. Como se puede observar, en el ensayo n el reforzador no se entrega por la emisión de una respuesta particular (A o B). La entrega del reforzador en el ensayo n depende de que en ese ensayo los sujetos emitan la respuesta opuesta a la que fue reforzada en el ensayo $n - 1$. De esta manera, sólo se refuerzan los patrones: *respuesta A–reforzador–respuesta B* y *respuesta B–reforzador–respuesta A*.

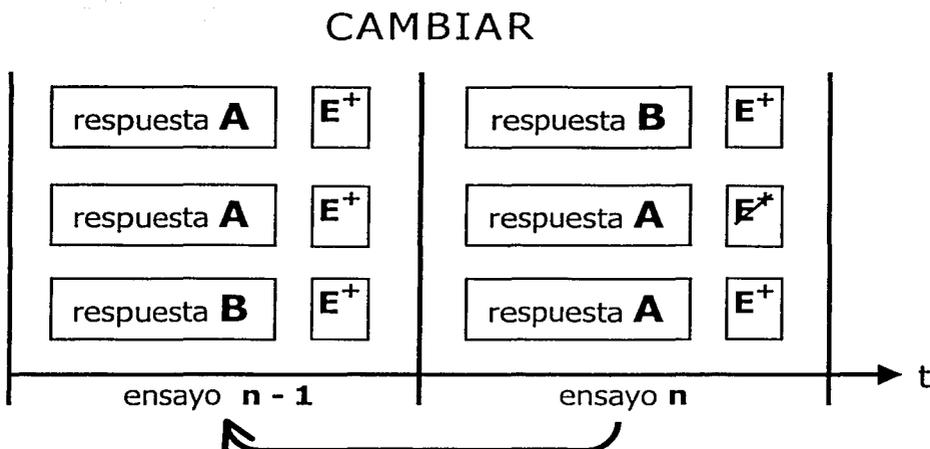


Figura 1. Procedimiento en el que la entrega del reforzador se hace contingente a la emisión de la respuesta opuesta a la emitida en el ensayo previo (cambiar). E^+ = reforzador, \bar{E}^+ = ausencia de reforzador, t = tiempo, n = número de ensayo.

Existen varios estudios que emplean procedimientos de ensayos discretos, en los que los sujetos tienen que responder con un patrón de repetición o cambio y estos patrones son reforzados dependiendo de las consecuencias de la conducta que ocurrió en el ensayo previo. Se ha empleado este tipo de procedimiento para estudiar maximización molecular (Hiraoka, 1984; Shimp, 1976b; Williams, 1972, 1991), recuerdo de la propia conducta (Shimp, 1976b; Randall & Zentall, 1997) y estrategias de forrajeo (Evenden & Robbins, 1984). En los diferentes trabajos en los que se utiliza este tipo de procedimiento, se les conoce como estrategias de: *ganar-repetir*, *ganar-cambiar*, *perder-repetir* y *perder-cambiar* a los patrones de repetición o cambio de respuesta dependientes de las consecuencias. Estas investigaciones pueden considerarse variantes de los procedimientos empleados por Shimp

(1976b) y Williams (1972). Por lo tanto, a continuación se describen con detalle los procedimientos empleados por estos autores.

Williams (1972) estudió las condiciones bajo las cuales se establecen las estrategias ganar-repetir y perder-cambiar en un programa en el cual la entrega del reforzador se programó de manera probabilística. Williams entrenó a 6 palomas para responder en la tecla izquierda (siempre iluminada en rojo) y en la tecla derecha (siempre iluminada en verde). La *probabilidad molar* de reforzamiento fue de 0.50 por responder en cada tecla. Las variables que manipuló fueron: a) el requisito de respuesta en las teclas, el cual podía ser de 1, 5 ó 10 respuestas (un requisito diferente para cada uno de 3 grupos de sujetos); y b) una *probabilidad local* de reforzamiento que se programó con dos valores dependiendo de si la respuesta en el ensayo previo era o no reforzada: 1) la probabilidad de reforzamiento era de 0.65 (0.80 en una segunda condición) si las palomas repetían la respuesta reforzada en el ensayo previo (estrategia ganar-repetir) y 2) la probabilidad fue de 1.0 si después de no obtener el reforzador en el ensayo previo, en el siguiente ensayo los sujetos respondían en la otra opción (estrategia perder-cambiar). La probabilidad molar de reforzamiento diferente a la probabilidad local implica que la tarea puede resolverse de dos formas: a) los sujetos pueden distribuir sus respuestas de manera que al final de la sesión igualen la proporción respuestas-reforzadores para cada una, sin que exista un patrón molecular específico por el cual se llega a dicha distribución; y b) que los sujetos distribuyan sus respuestas entre las teclas de acuerdo a las probabilidades locales de reforzamiento que implican ajustar las respuestas a los patrones ganar-repetir y perder-cambiar.

En este procedimiento, al inicio de cada ensayo los sujetos debían responder en las teclas laterales hasta que completaran la razón requerida

(1, 5 ó 10 respuestas) en alguna de las opciones. Si su elección (tecla en la que completaban la razón) era reforzada, en el siguiente ensayo recibían un reforzador con una probabilidad local de 0.65 (0.80) por responder de acuerdo a la estrategia ganar-repetir; mientras que si tal elección no era seguida de la entrega de recompensa, en el siguiente ensayo se les reforzaba con una probabilidad local de 1.0 si respondían conforme a la estrategia perder-cambiar. Williams encontró que los sujetos del grupo al que se le pedía una razón de 15 respuestas fueron los que mejor ajustaron la distribución de sus respuestas de acuerdo a las estrategias. Un segundo hallazgo fue que todos los sujetos ejecutaron la estrategia perder-cambiar cuando era necesaria, sin embargo, la estrategia ganar-repetir sólo la emitieron cuando la probabilidad local de reforzamiento fue de 0.80.

Años más tarde, Shimp (1976b) utilizó una variante del procedimiento de Williams (1972) con el propósito de investigar hasta qué punto es posible que las contingencias locales de reforzamiento establezcan los patrones de comportamiento: ganar-repetir y ganar-cambiar. Shimp entrenó a 18 palomas a emitir un picotazo en las teclas laterales Es decir, las opciones de respuesta fueron una respuesta en la tecla izquierda (I) o una respuesta en la tecla derecha (D). La probabilidad molar de reforzamiento para ambas alternativas fue de 0.50. Shimp manipuló las siguientes variables: a) intervalos entre ensayos de 0.5, 2.0 y 4.0 s después de que una respuesta era reforzada. La selección del valor para el intervalo la hizo de manera aleatoria; y b) para 9 palomas, la probabilidad local de reforzamiento fue de 0.80 si repetían la respuesta reforzada en el ensayo previo (estrategia ganar-repetir); mientras que para otras 9 palomas esta probabilidad fue de 1.0 si respondían en la alternativa que no se había reforzado en el ensayo anterior (estrategia ganar-cambiar). Shimp encontró que todas las palomas respondieron de acuerdo a

las estrategias reforzadas y que la exactitud de su ejecución disminuyó conforme se incrementó el valor del intervalo entre ensayos. A partir de estos resultados Shimp sugirió que la ejecución de los sujetos de acuerdo a las estrategias dependió del recuerdo de la respuesta previa.

A partir de los estudios realizados por Shimp (1976b) y Williams (1972), junto con los realizados por Hiraoka (1984) y Randall y Zentall (1997), que emplearon procedimientos semejantes, se puede decir que las variables que parecen importantes para que los animales respondan de acuerdo a las diferentes estrategias de respuesta son: un requisito de respuestas relativamente grande (por ejemplo, una RF10 en Randall & Zentall, 1997; y una RF15 en Williams, 1972), un intervalo entre ensayos pequeño (2s – 4 s en Randall & Zentall, 1997; 0.5 s – 2 s en Shimp, 1976b), una probabilidad local de reforzamiento mayor o igual a 0.80 (Hiraoka, 1984; Shimp, 1976b; Williams, 1972, 1991) y emplear un estímulo discriminativo (adicional a la posición: izquierda o derecha) para facilitar la distinción entre las alternativas (por ejemplo, colores rojo y verde; Randall & Zentall, 1997; Williams, 1972, 1991). En estos trabajos se ha encontrado de manera consistente una mejor ejecución para las estrategias ganar-repetir (Hiraoka, 1984; Randall & Zentall, 1997; Shimp, 1976b; Williams, 1991; c.f., Williams, 1972) y perder-cambiar (Williams, 1972) que para las otras dos opciones; ganar-cambiar y perder-repetir.

La ejecución de las diferentes estrategias de respuesta (ganar-repetir, ganar-cambiar, perder-repetir y perder-cambiar) incluye un patrón conductual conformado básicamente por tres elementos: 1) la conducta que se emite en el ensayo n , 2) la consecuencia de esa conducta, y 3) la conducta que se emite en el ensayo $n + 1$. En los trabajos mencionados anteriormente, se

han realizado diferentes combinaciones de estos elementos, de tal manera que se ha estudiado la adquisición de las estrategias ganar-repetir y perder-cambiar (Hiraoka, 1984; Williams, 1972, 1991), ganar-repetir y ganar-cambiar (Shimp, 1976), y la de ganar-repetir/perder-cambiar y ganar-cambiar/perder-repetir (Randall & Zentall, 1997). Debido a que la mayoría de estos trabajos se han realizado dentro del área de conducta de elección, la ejecución del patrón de repetición o cambio siempre se ha estudiado conforme a la *elección* de una opción de respuesta (por ejemplo, la tecla izquierda vs. la derecha) y no en relación a la repetición (o cambio) de una *conducta particular*. Es decir, en estos trabajos se ha definido la elección de una alternativa ya sea como el cumplimiento de un requisito de respuestas para esa opción (por ejemplo, 10 respuestas en la tecla izquierda; Williams, 1972, 1991) o como la primera respuesta que se emite en la alternativa para la que está asignado el reforzador (Hiraoka, 1984; Shimp, 1976b). Del tal forma que no se presta atención al patrón de respuestas mediante el cual los sujetos cumplen con el requisito para la elección. Por ejemplo, cuando la elección de una alternativa se define como un requisito de 10 respuestas, cualquiera de los siguientes patrones: IIDIIIDDIIIDII, DDDDIIIIIIIII, IIIIDIDIDIDII, o IIIIIIIII, se puede considerar una “repetición” de la elección de la opción izquierda (o un “cambio” a izquierda cuando la elección en el ensayo previo fue derecha).

En una parte de la literatura presentada se ha encontrado que los animales tienden a repetir una respuesta cuando es seguida de la entrega de reforzador y que tienden a emitir una respuesta distinta después de la ausencia del mismo. Otra porción de esa literatura ha mostrado que los

“Los animales tienden a repetir las acciones que los llevan al placer y desisten de aquellas que los llevan al sufrimiento” Herbert Spencer (1855)

animales aprenden a responder de acuerdo a los patrones opuestos, es decir, repetir después de que no se entrega el reforzador y cambiar después de que una respuesta es reforzada. En resumen, se ha demostrado que los animales aprenden a responder con cualquiera de los cuatro diferentes patrones de respuesta mencionados: ganar-repetir, ganar-cambiar, perder-repetir y perder-cambiar.

Debe aclararse que algunos de esos trabajos han sido realizados para estudiar conducta de elección (Hiraoka, 1984; Shimp, 1976b; Williams, 1972, 1991), otros para observar patrones de variabilidad conductual, (Neuringer, 1992, 1993; Machado, 1992; Page & Neuringer, 1985), para estudiar memoria (Shimp, 1976b; Randall & Zentall, 1997) o estrategias de forrajeo (Evenden & Robbins, 1984). Y aunque todos ellos coinciden en procedimientos en los que la entrega del reforzador en un ensayo depende de la respuesta y sus consecuencias ocurridas en el ensayo previo; lo cierto es que no se llevaron a cabo para estudiar el posible mecanismo mediante el cual los animales aprenden estas estrategias de respuesta.

Estudiar este mecanismo es importante porque, si bien es cierto que a partir de la Ley del Efecto se predice la aparición de los patrones ganar-repetir y perder-cambiar, de ella no se deriva directamente la aparición de los patrones opuestos. El hecho de que los animales aprendan tanto a repetir una respuesta como a cambiarla después de la entrega de una recompensa sugiere que el reforzador puede tener dos funciones, la de seleccionar una respuesta y la de generar variabilidad conductual.

Una explicación de cómo un reforzador puede producir ambos patrones de respuesta es asumiendo que el reforzador tiene dos efectos sobre la

"What do reinforcers do to behavior?
I don't know why reinforcers affect behaviour, but I do know that some behavioural
consequences change the probability of behaviours that they happen to follow..."

Michael Davison (2001)

conducta, un efecto de fortalecimiento que produce la repetición de la respuesta inmediata previa (efecto por contigüidad temporal con el reforzador) y otro que provee "información" acerca de la contingencia de reforzamiento (efecto de contingencia).

En el presente trabajo se propone desarrollar un procedimiento que permita estudiar la forma en la que una conducta y su consecuencia pueden tomar el control sobre lo que hace un organismo en el siguiente momento. Para lograr esto, el procedimiento debe permitir el control de los tres componentes del patrón de las estrategias (la conducta en el ensayo n , su consecuencia y la conducta en el ensayo $n + 1$). El procedimiento que se propone consiste en un programa de reforzamiento aleatorio para entrenar dos alternativas de respuesta, bajo el cual se presentan al *azar* ensayos en los que la obtención de un reforzador depende de la ejecución de alguna de las estrategias de respuesta. En el experimento que se presenta en este trabajo se mantiene constante tanto la conducta que se debe repetir o cambiar en el siguiente ensayo como la consecuencia de la misma y sólo se manipula el patrón de repetición o cambio al que es contingente la entrega del reforzador en cada condición experimental: *ganar-repetir* o *ganar-cambiar*.

"¿Cómo pueden tener efecto los resultados de una acción sobre la misma acción cuando la acción está terminada y hecha antes de que se pongan de manifiesto los resultados?, ¿Qué clase de relación causa-efecto es ésta, en la que el efecto tiene que hacer de causa?"

Keller & Schoenfeld (1950)

MÉTODO

Sujetos

Se emplearon cuatro ratas hembra de la cepa Wistar de tres meses de edad al inicio del experimento y sin experiencia experimental. Todos los sujetos se mantuvieron al 85% de su peso *ad libitum* y con acceso libre a agua. Al final de cada sesión experimental se les alimentó para mantenerlos en su peso.

Aparatos

Se utilizaron dos cámaras experimentales de condicionamiento operante para ratas (MED Associates Mod. ENV-007). Cada cámara se colocó dentro de una caja sono-amortiguadora de 60 x 90 x 80 cm que contenía un extractor de aire que funcionó como ruido blanco. En el panel posterior de las cámaras operantes había un foco, colocado en la parte central y a una distancia de 2 cm del techo, que se utilizó como luz general. En el panel frontal, a una altura de 7 cm, había dos palancas: una del lado izquierdo y otra del lado derecho. Cada palanca requería para su funcionamiento de una fuerza de 0.5 N. Colocados a 7 cm sobre cada una de las palancas había un foco; la distancia horizontal entre los focos y las palancas era de 17 cm. En la parte central del panel frontal y situado a una altura de 2 cm se encontraba una abertura cuadrada de 4 cm por lado que funcionó como receptáculo de pellets. Se utilizaron como reforzadores pellets de 45 mg de la compañía Noyes (Fórmula F). Las cajas estaban conectadas a una interfase (MED Associates Mod. 715) y ésta a una computadora Pentium a través de la cual, mediante un programa elaborado en Medstate Notation, se controlaron las sesiones experimentales y se registraron las respuestas en tiempo real.

Procedimiento

El procedimiento se dividió en tres condiciones generales: 1) *Entrenamiento Preliminar*, 2) *Entrenamiento*, y 3) *Fases Experimentales*. Sin embargo, existen aspectos comunes a todo el experimento que se mencionan a continuación.

La operante con la que se trabajó fueron secuencias de dos respuestas en las palancas: izquierda-izquierda (II), izquierda-derecha (ID), derecha-izquierda (DI) y derecha-derecha (DD). Se utilizó un procedimiento de ensayos discretos que ocurrieron de la siguiente manera: al inicio de cada ensayo se encendía la luz general y las luces que se encontraban sobre las palancas. Después de que los sujetos emitían una secuencia de dos respuestas se apagaban todas las luces y, conforme a cada condición experimental, si la secuencia era correcta se entregaba el reforzador. Las sesiones se condujeron 6 días a la semana y cada sesión terminó después de 50 ensayos.

Entrenamiento Preliminar

Esta condición consistió de dos partes. Primero se entrenó a las ratas para que presionaran dos veces las palancas. Las dos presiones podían ser sobre la misma palanca (secuencias homogéneas: II y DD) o una presión en cada operando (secuencias heterogéneas: ID y DI), de modo que hubo cuatro posibles secuencias de respuestas reforzadas. La sesión terminaba después de 50 ensayos o después de transcurridos 30 minutos. Esta condición se mantuvo vigente hasta que los sujetos obtuvieron un mínimo de 45 reforzadores en una sola sesión.

El objetivo de la segunda parte del Entrenamiento Preliminar fue evitar una preferencia a responder sólo en una de las palancas, por lo tanto se reforzaron únicamente las secuencias heterogéneas (ID y DI). Es decir, los

sujetos tenían que responder en ambas palancas para obtener un reforzador. Se expuso a los sujetos a esta segunda etapa durante 4 días.

Entrenamiento

Se utilizó un programa de reforzamiento aleatorio con reemplazo para reforzar sólo las secuencias homogéneas (II y DD). El propósito de utilizar este tipo de programa fue entrenar a los sujetos a emitir las dos secuencias con las que se trabajó durante las Fases Experimentales de manera que no se reforzara diferencialmente la estrategia *ganar-repetir*: ejecutar la secuencia previamente reforzada, ni la estrategia *ganar-cambiar*: emitir la secuencia homogénea opuesta a la previamente reforzada. Este programa consistió en lo siguiente.

Al inicio de cada ensayo se encendían todas las luces y el programa elegía, con una probabilidad de 0.50, la secuencia II o la secuencia DD. Si la secuencia que emitía el sujeto coincidía con la elegida por el programa entonces se apagaban todas las luces, se entregaba el reforzador y después de 4 s se iniciaba un nuevo ensayo. En cambio si la secuencia de respuestas que ejecutaba el sujeto no coincidía con la elegida por el programa se apagaban todas las luces y había un tiempo entre ensayos de 8 s (véase la Figura 2).

Al inicio de cada nuevo ensayo el programa volvía a elegir una de las secuencias independientemente de la elección y las consecuencias del ensayo anterior; es decir, en todos los ensayos ambas secuencias tenían la misma probabilidad de ser seleccionadas para reforzamiento. El criterio de cambio a las Fases Experimentales fue que en 3 sesiones consecutivas la frecuencia de secuencias homogéneas emitidas fuera mayor a 40 por sesión y un mínimo de 10 sesiones en el Entrenamiento.

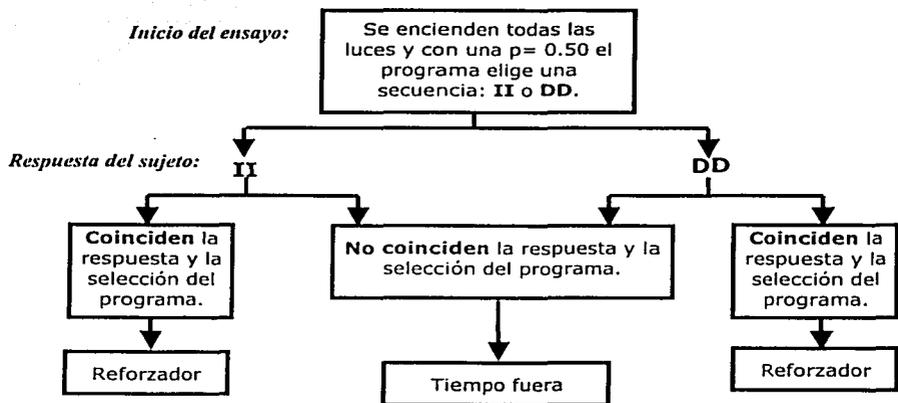


Figura 2. Diagrama en donde se ejemplifica el procedimiento para el Entrenamiento.

Fases Experimentales

Se utilizó un procedimiento en el cual había dos tipos de ensayo: a) *ensayos aleatorios*, que se presentaron 50 veces por sesión y b) *ensayos de estrategia* que se presentaban sólo después de algunos ensayos aleatorios que terminaban con la entrega de reforzador (véase la Figura 3). Para dos de los sujetos, los ensayos de estrategia se presentaron después de que en un ensayo aleatorio era reforzada la secuencia II, mientras que para los otros dos se presentaron después de que en un ensayo aleatorio era reforzada la secuencia DD.

CONDICIÓN:		II - GANAR - CAMBIAR					
NÚMERO DE ENSAYO		1	2	3	4	5	6
E N S A Y O A L E A T O R I O	SELECCIÓN DEL PROGRAMA	II	II	II	DD	II	DD
	RESPUESTA DE LA RATA	II	DD	II	DD	II	II
	CONSECUENCIA	E+	E+	E+	E+	E+	E+
	LUCES SOBRE LAS PALANCAS	encendidas	apagadas	apagadas	apagadas	encendidas	apagadas
E N S A Y O D E E S T R A T E G I A	RESPUESTA DE LA RATA (ESTRATEGIA)	II (REPITE)				DD (CAMBIA)	
	CONSECUENCIA	E+				E+	

Figura 3. Ejemplo de 6 ensayos para la condición II - Ganar - Cambiar.

Nota. II = izquierda-izquierda; DD = derecha-derecha; E+ = reforzador; ~~E+~~ = no reforzador; = ensayos de estrategia en los que se requería que el sujeto cambiara (emitir DD) después de que II era reforzada en los ensayos aleatorios.

La variable que se manipuló fue el tipo de estrategia necesaria para obtener un reforzador en los ensayos de estrategia de cada Fase Experimental: *ganar-repetir* o *ganar-cambiar*. Esta variable definió la condición experimental de cada Fase. Se expuso a los sujetos a cuatro Fases

Experimentales. Es decir, se sometieron en dos ocasiones a cada condición de tal manera que se pudieran comparar para un mismo sujeto los efectos de las condiciones sobre la ejecución de las estrategias. Se contrabalanceó entre los sujetos el orden en el que se presentaron las dos condiciones y la secuencia de respuestas (II o DD) que antecedía a la presentación de los ensayos de estrategia. Las Fases 1 y 2 tuvieron una duración de 10 días cada una mientras que la Fase 3 y la Fase 4 duraron 20 días (véase la Tabla 1).

Tabla 1

Estrategia requerida para cada uno de los sujetos en las diferentes Fases Experimentales, secuencia reforzada que antecedió la presentación de los ensayos de estrategia y duración de cada Fase.

Sujeto	Secuencia	Estrategia			
		Fase 1	Fase 2	Fase 3	Fase 4
DN1	II	Rep	Cam	Rep	Cam
DN2	II	Cam	Rep	Cam	Rep
DN3	DD	Rep	Cam	Rep	Cam
DN5	DD	Cam	Rep	Cam	Rep
Duración	(días):	10	10	20	20

Nota. II = izquierda-izquierda; DD = derecha-derecha; Rep = condición ganar-repetir; Cam = condición ganar-cambiar.

A continuación se explica el procedimiento de cada tipo de ensayo.

Ensayos aleatorios. En estos ensayos, las secuencias II y DD se reforzaron con una probabilidad de 0.25 conforme al siguiente procedimiento.

Al inicio de este tipo de ensayo se encendían la luz general y luces sobre las palancas y con una probabilidad de 0.50 el programa elegía una de las dos secuencias homogéneas. Si la secuencia con la que respondía el sujeto *no coincidía* con la secuencia seleccionada por el programa, se apagaban todas las luces y después de 4 s iniciaba otro ensayo aleatorio. Cuando la secuencia seleccionada por el programa *coincidía* con la que emitía el sujeto, el programa determinaba con una probabilidad de 0.50 si la secuencia emitida sería reforzada; por lo tanto, la probabilidad total de reforzamiento programada para cada una de las secuencias de respuestas homogéneas fue de 0.25. Cuando no se reforzaba la secuencia emitida se apagaban todas las luces y después de 8 s de tiempo fuera iniciaba un nuevo ensayo aleatorio (véase la Figura 4).

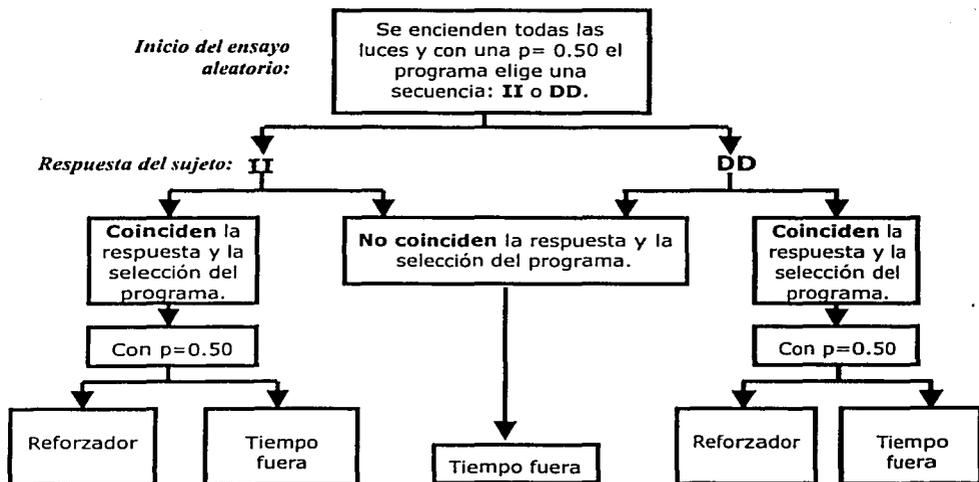


Figura 4. Diagrama en donde se ejemplifica el procedimiento de reforzamiento aleatorio.

Ensayos de Estrategia. Estos ensayos se presentaron un segundo después de que finalizaban los ensayos aleatorios en los cuales era reforzada la secuencia II para los sujetos DN1 y DN2 y la secuencia DD para los sujetos DN3 y DN5. En los ensayos de estrategia la entrega del reforzador dependía de que la respuesta del sujeto correspondiera, de acuerdo a la condición experimental, a una de dos estrategias de respuesta: a) ganar-repetir, que consistía en repetir la secuencia reforzada en el ensayo aleatorio previo, o b) ganar-cambiar, que consistía en emitir la secuencia opuesta a la reforzada en el ensayo aleatorio anterior. El procedimiento en este tipo de ensayos fue el siguiente.

Al final de los ensayos aleatorios que eran seguidos por un ensayo de estrategia, se entregaba el reforzador al mismo tiempo que se apagaba la luz general y quedaban encendidas las luces sobre las palancas. De esta manera, los estímulos discriminativos indicadores de que el siguiente ensayo sería un ensayo de estrategia eran las luces encendidas durante el tiempo entre ensayos y la secuencia reforzada en el ensayo aleatorio. Después de 1 s se volvía a encender la luz general. En ese momento, los sujetos que se encontraban en la condición ganar-repetir tenían que emitir la secuencia que se les acababa de reforzar mientras que los sujetos que se encontraban en la condición ganar-cambiar tenían que emitir la secuencia homogénea que no se había reforzado. Una vez que el sujeto ejecutaba una secuencia de dos respuestas en el ensayo de estrategia, se apagaban todas las luces, sonaba un tono y sólo si el sujeto había respondido con la secuencia que correspondía a la estrategia correcta se entregaba el reforzador (probabilidad de reforzamiento = 1.0). El tiempo entre ensayos fue de 8 s (véase la Figura 6).

DD - GANAR-CAMBIAR

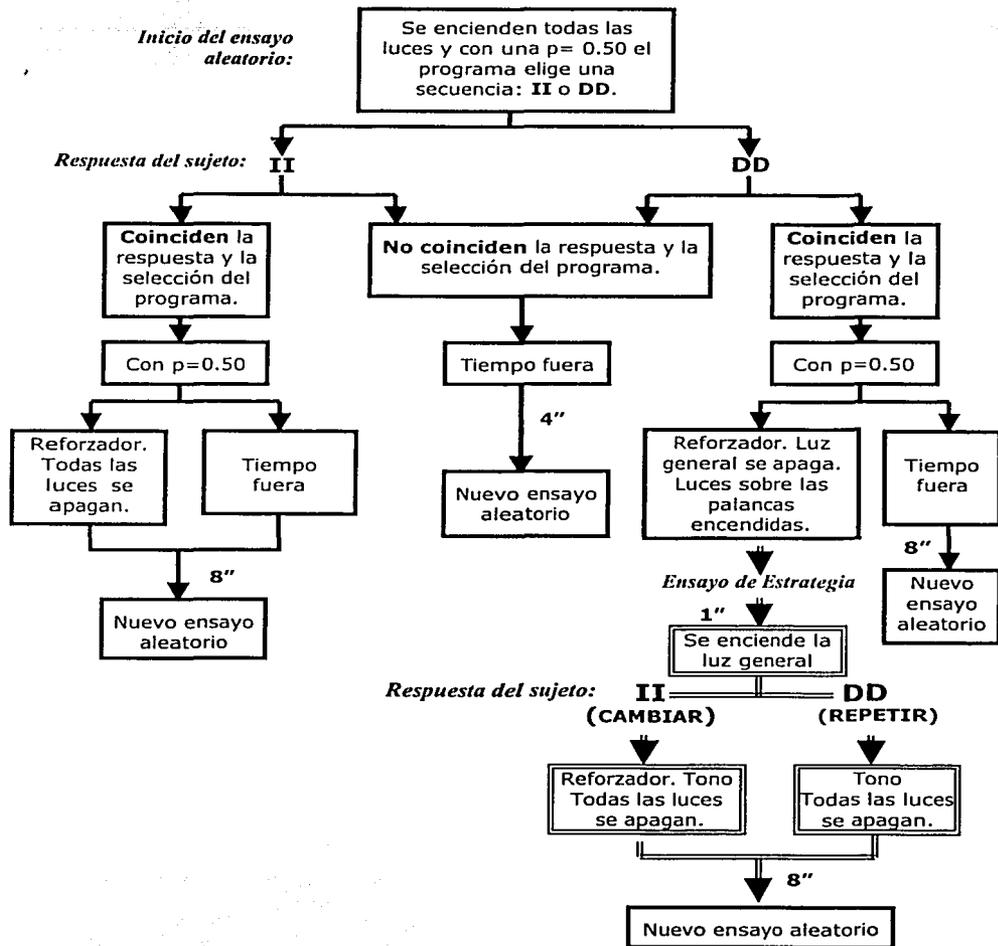


Figura 6. Diagrama en donde se ejemplifica el procedimiento cuando los ensayos de estrategia (líneas dobles) se presentan después de que la secuencia DD es reforzada y la estrategia requerida es ganar-cambiar. Durante los tiempos fuera todas las luces se mantenían apagadas.

RESULTADOS

Para determinar si durante el Entrenamiento los sujetos obtuvieron la misma cantidad de reforzadores por emitir la secuencia II y la secuencia DD se calculó el promedio de la proporción de reforzadores ganados por ejecutar cada una de las secuencias (número de reforzadores obtenidos por emitir una secuencia entre el total de reforzadores) durante los últimos cinco días del entrenamiento. En la Figura 7 se muestra que esta proporción fue muy semejante para ambas secuencias. Para los cuatro sujetos, el promedio de la proporción de reforzadores obtenidos por emitir II fue de 0.51 ($DE = 0.05$) y por ejecutar DD fue de 0.49 ($DE = 0.05$).

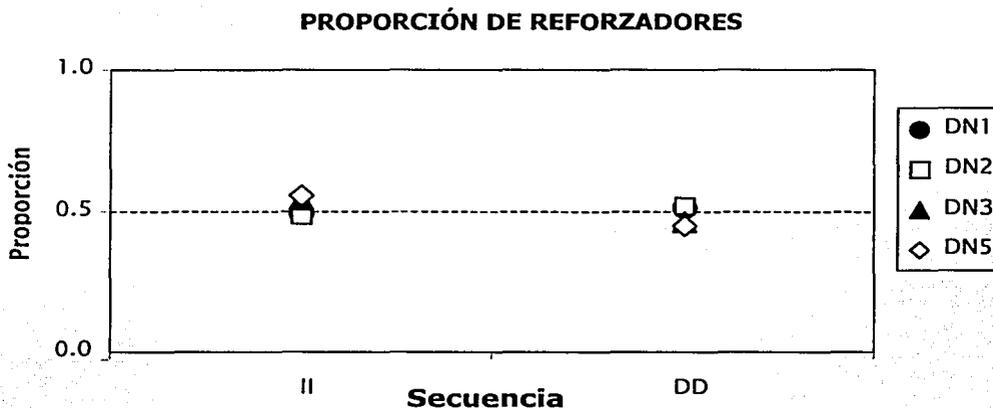
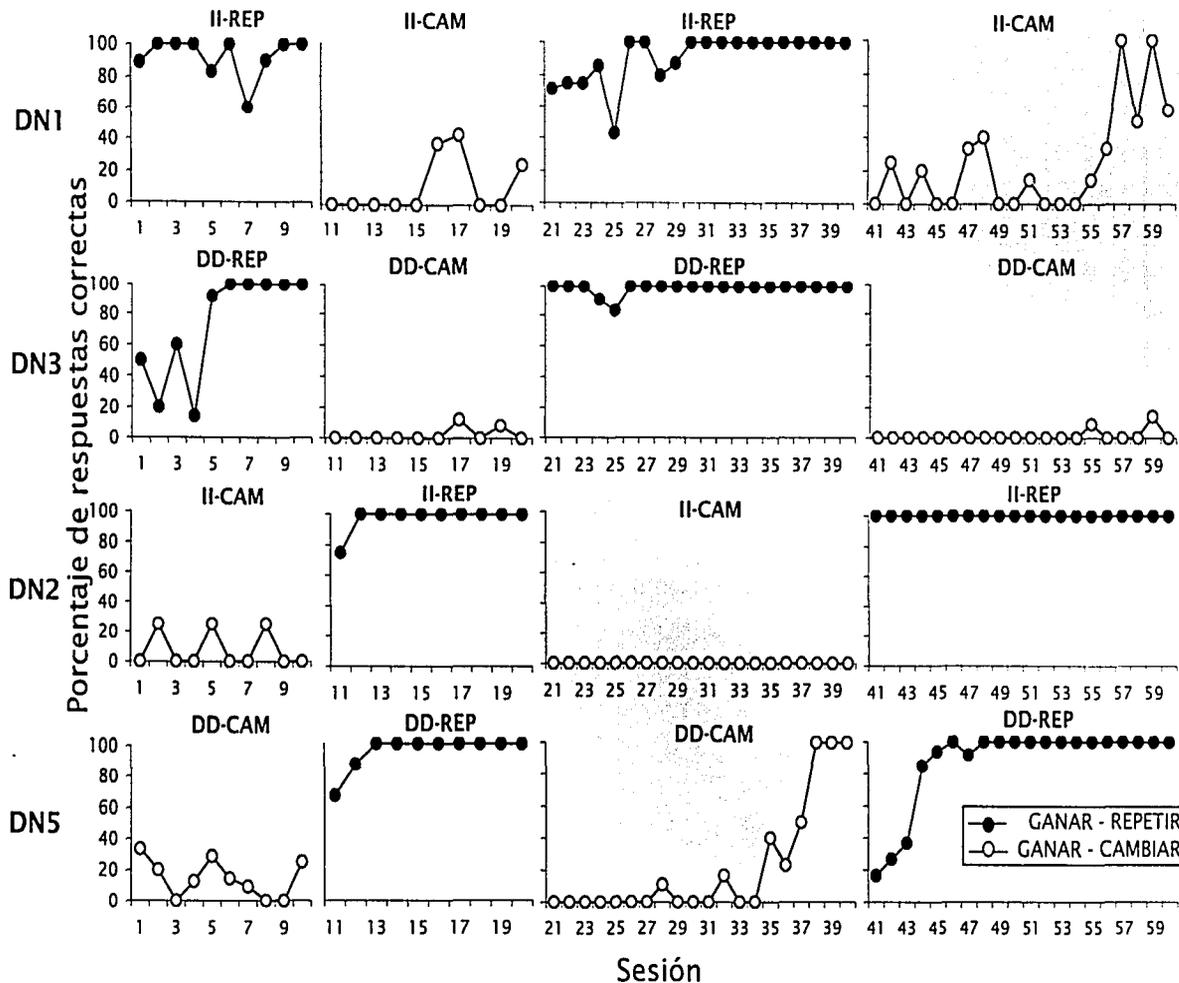


Figura 7. Promedio para cada sujeto de la proporción de reforzadores obtenidos por emitir la secuencia II y la secuencia DD durante los últimos 5 días del Entrenamiento.

En las Fases Experimentales el promedio de ensayos de estrategia que se presentaron por sesión fue de 5 ($DE = 2$) para el sujeto DN1, de 5 ($DE = 4$) para el sujeto DN2, de 9 ($DE = 3$) ensayos para DN3 y de 10 ($DE = 4$) para DN5. Se calculó el porcentaje de respuestas correctas en los ensayos de estrategia en las diferentes Fases Experimentales. Se consideró como una respuesta correcta cuando: a) los sujetos se encontraban en la condición ganar-repetir y repetían la secuencia previamente reforzada y, b) cuando los sujetos se encontraban en la condición ganar-cambiar y emitían la secuencia homogénea que no les había sido reforzada en el ensayo anterior. En la Figura 8 se presentan los porcentajes diarios para cada una de las Fases a las que fueron expuestos los sujetos. Los porcentajes muestran que existe una mayor tendencia a repetir (círculos negros) que a cambiar (círculos blancos) después de una secuencia reforzada. Durante la condición ganar-repetir, el porcentaje de veces que los sujetos repitieron la secuencia previamente reforzada se encuentra por arriba del 80% de respuestas correctas y su ejecución se mantuvo relativamente estable. Cuando la condición fue ganar-cambiar, el porcentaje de veces que los sujetos cambiaron después de una secuencia reforzada, en general, se mantuvo por debajo del 40% de respuestas correctas; en el extremo se encuentra el sujeto DN2 que durante la Fase 3 no tuvo respuestas correctas. Sin embargo, la segunda ocasión en que los sujetos DN1 (Fase 4) y DN5 (Fase 3) fueron expuestos a la condición ganar-cambiar, el porcentaje de respuestas correctas incrementó durante las últimas 4 sesiones para el sujeto DN1 quien tuvo un promedio de 77% de respuestas correctas y durante las últimas 3 sesiones para el sujeto DN5 con el 100% de respuestas correctas.



TESIS CON FALLA DE ORIGEN

Figura 8. Porcentaje de respuestas correctas en los ensayos de estrategia para cada sujeto en las cuatro Fases Experimentales. *II-REP*: condición ganar-repetir después de ser reforzada II; *II-CAM*: condición ganar-cambiar después de ser reforzada II; *DD-REP*: condición ganar-repetir después de ser reforzada DD; *DD-CAM*: condición ganar-cambiar después de ser reforzada DD.

A partir de estos resultados, se averiguó si durante el Entrenamiento ya existía en el patrón de respuestas de los sujetos una tendencia por emitir las secuencias de acuerdo a alguna de las estrategias. Con esta finalidad, se obtuvo la probabilidad condicional de emitir la misma secuencia (repetir) o ejecutar la contraria (cambiar) en el siguiente ensayo dado que en el ensayo previo una de las secuencias homogéneas había sido reforzada. La probabilidad condicional de que los sujetos emitieran la secuencia II cuando en el ensayo previo era reforzada II se abrevió como $P(II | II+)$, la probabilidad condicional de que ejecutaran la secuencia II después de ser reforzada DD se abrevió como $P(II | DD+)$ y de manera similar se hizo la notación de las demás probabilidades condicionales. Para calcular la $P(DD | II+)$, se dividió la frecuencia de veces que un sujeto fue reforzado por emitir la secuencia II y en el siguiente ensayo ejecutó DD, entre la frecuencia total de reforzadores obtenidos por emitir II. Las otras probabilidades condicionales se calcularon de manera semejante. En la Tabla 2 se muestra el promedio de los últimos 5 días de las probabilidades condicionales de ganar-repetir y ganar-cambiar para cada uno de los sujetos.

Los datos de la Tabla 2 muestran que para los sujetos DN3 y DN5 existió una mayor probabilidad de repetir la secuencia homogénea reforzada en el ensayo previo. Para el sujeto DN2 también existió una mayor tendencia a repetir pero sólo después de que II era reforzada; cuando se entregaba el reforzador después de DD la tendencia fue a cambiar. Mientras que, para el sujeto DN1 se encontró una mayor probabilidad de repetir después de ser reforzada DD y de cambiar después de II.

Tabla 2

Probabilidad condicional de emitir las estrategias ganar-repetir [$P(II | II+)$ y $P(DD | DD+)$] y ganar-cambiar [$P(DD | II+)$ y $P(II | DD+)$]. Promedio de los 5 últimos días en el Entrenamiento.

ESTRATEGIA	PROBABILIDAD CONDICIONAL	SUJETO			
		DN1	DN2	DN3	DN5
GANAR REPETIR	$P(II II+)$	0.3097	0.9085	0.6525	0.6598
	$P(DD DD+)$	0.5351	0.3397	0.5044	0.5956
GANAR CAMBIAR	$P(DD II+)$	0.5538	0.0167	0.2026	0.1171
	$P(II DD+)$	0.3370	0.5617	0.3736	0.1036

Nota. $P(II | II+)$ = probabilidad condicional de emitir II dado que II fue reforzada;
 $P(DD | DD+)$ = probabilidad condicional de emitir DD dado que DD fue reforzada;
 $P(DD | II+)$ = probabilidad condicional de emitir DD dado que II fue reforzada;
 $P(II | DD+)$ = probabilidad condicional de emitir II dado que DD fue reforzada.

Para los datos de las Fases Experimentales también se calcularon las probabilidades condicionales de cambiar o repetir una secuencia después de ser reforzada para observar si se modificaron, con respecto a los valores que se obtuvieron en el Entrenamiento, de acuerdo a la condición de cada Fase. Se esperaba que durante la condición ganar-repetir incrementara la probabilidad de repetir la secuencia reforzada para la que se presentaron los ensayos de estrategia [$P(II | II+)$ o $P(DD | DD+)$] y durante la condición ganar-cambiar se esperaba que cuando era reforzada la secuencia para la que se presentaban los ensayos de estrategia incrementara la probabilidad de emitir la secuencia opuesta [$P(DD | II+)$ o $P(II | DD+)$]. En la Figura 9 se muestra para cada sujeto el promedio de los últimos 5 días de la probabilidad

condicional de cambiar (barras blancas) o de repetir (barras negras) la secuencia previamente reforzada en los ensayos de estrategia de cada una de las Fases Experimentales; los datos que se muestran del Entrenamiento corresponden al promedio de los últimos 5 días en esta condición únicamente para la secuencia en la cual se presentaron los ensayos de estrategia en las Fases Experimentales (los datos de las probabilidades condicionales para la otra secuencia se muestran en la Tabla 2).

Los datos de la Figura 9 muestran que para los sujetos DN2 y DN3, sin importar la Fase en la que se encontraban, la probabilidad de repetir la secuencia previamente reforzada en los ensayos de estrategia fue muy cercana a 1.0. En el caso del sujeto DN2, la tendencia a repetir la secuencia II después de que era reforzada ya existía desde el Entrenamiento, sin embargo, para el sujeto DN3 el promedio de la probabilidad de repetir la secuencia DD incrementó de 0.50 en el Entrenamiento a 1.00 en la primera Fase. Por otra parte, para el sujeto DN1 el promedio de la probabilidad de repetir la secuencia II después de ser reforzada incrementó de 0.30 en el Entrenamiento a 0.88 en la primera Fase (condición ganar-repetir), esta tendencia a repetir la secuencia previamente reforzada se mantuvo durante las tres primeras Fases Experimentales. Sin embargo, en la Fase 4 (condición ganar-cambiar) el promedio de la probabilidad de cambiar incrementó de cero en la Fase anterior a 0.68 y el promedio de la probabilidad de repetir decrementó de 1.00 a 0.31. Los valores de las probabilidades condicionales también se modificaron para el sujeto DN5. Para este sujeto, la tendencia durante el Entrenamiento y las dos primeras Fases Experimentales fue repetir la secuencia DD después de que era reforzada. No obstante, en la Fase 3 (condición ganar-cambiar), el promedio de la probabilidad de emitir la

PROBABILIDAD CONDICIONAL
EN LOS ENSAYOS DE ESTRATEGIA

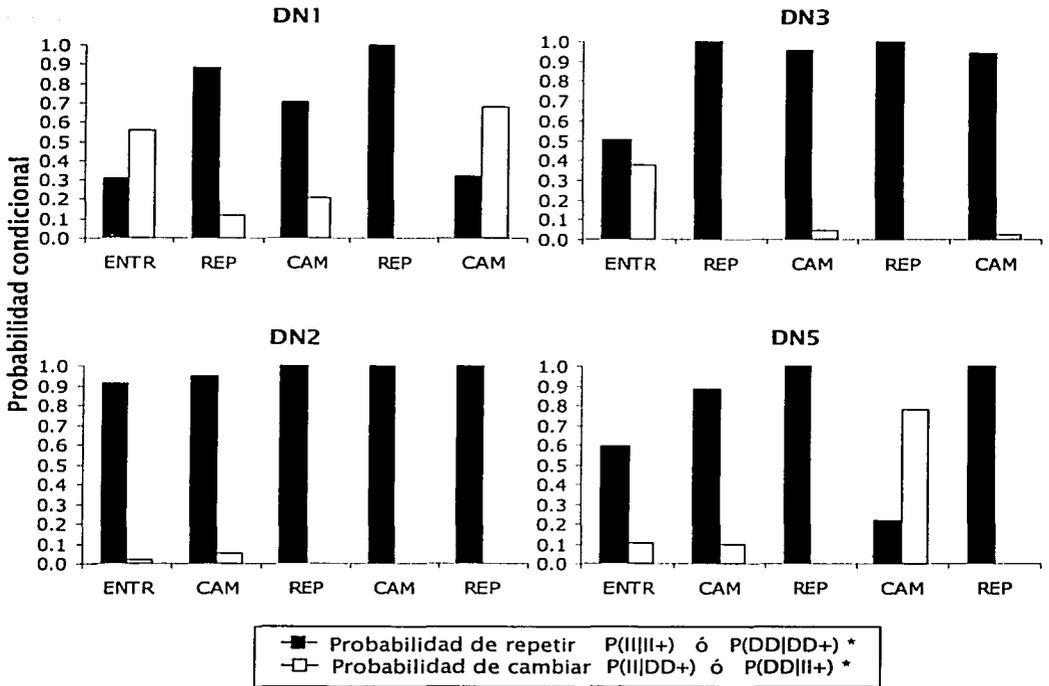


Figura 9. Promedio de los últimos 5 días en el Entrenamiento y en cada Fase Experimental, de la probabilidad condicional de cambiar (barras blancas) o repetir (barras negras) la secuencia previamente reforzada.

Nota. ENTR= Entrenamiento; REP= condición ganar-repetir; CAM= condición ganar-cambiar

* para los sujetos DN1 y DN2 se calcularon las probabilidades después de ser reforzada II y para los sujetos DN3 y DN5 después de ser reforzada DD (véase la Tabla 1).

secuencia que no había sido reforzada en el ensayo previo fue de 0.78 y el de repetir fue de 0.22 y , para la Fase 4 (condición ganar-repetir) el promedio de la probabilidad de repetir la secuencia previamente reforzada incrementó de 0.22 a 1.0 y la de cambiar decrementó de 0.78 a cero.

DISCUSIÓN

En el presente trabajo se propone el desarrollo de un procedimiento que permita tener una primera aproximación al estudio de cómo una conducta y sus consecuencias pueden tener el control sobre la conducta que se emite en el siguiente instante. Bajo este procedimiento, la oportunidad de obtener un reforzador por responder de acuerdo a las diferentes estrategias (ganar-repetir, ganar-cambiar, perder-repetir y perder-cambiar) se presenta de manera aleatoria. En el experimento que se realizó se estudió la adquisición de las estrategias de respuesta ganar-repetir y ganar-cambiar, cuando la conducta y la consecuencia que determinan la respuesta correcta en el siguiente ensayo se mantienen constantes.

El criterio para determinar si los animales ajustaron su conducta de acuerdo a la estrategia de respuesta reforzada en cada condición fue un nivel de ejecución por arriba del 80% de respuestas correctas en los ensayos de estrategia. El número de ensayos de estrategia que hubo en cada sesión fue variable debido a que su presentación dependía de que los sujetos emitieran la secuencia para la cual estaba programada la presentación de estos ensayos y que esta secuencia fuera seleccionada por el programa. Se consideró una respuesta correcta en estos ensayos cuando: a) los sujetos se encontraban en la condición ganar-repetir y repetían la secuencia previamente reforzada y b) cuando los sujetos se encontraban en la condición ganar-cambiar y emitían la secuencia que no había sido reforzada en el ensayo previo. Los resultados mostraron que todos los animales rebasaron el criterio establecido cuando la condición experimental fue ganar-repetir. Sin embargo, cuando la condición fue ganar-cambiar todos los sujetos mantuvieron su nivel de ejecución por debajo del 40% de respuestas correctas y fue sólo hasta la segunda ocasión en que se impuso esta condición, que el sujeto DN5 cumplió el criterio al tener el

100% de respuestas correctas en las últimas 3 sesiones y el sujeto DN1 alcanzó el 100% de respuestas correctas en 2 de las últimas sesiones. En experimentos en los cuales se estudia la ejecución de los patrones ganar-repetir y ganar-cambiar también se ha encontrado una mayor tendencia a repetir la respuesta previamente reforzada y una menor tendencia a cambiar después de la misma (e.g., Hiraoka, 1984; Randall & Zentall, 1997; Shimp, 1976), aunque la diferencia en la ejecución no ha sido tan marcada como en el presente experimento.

Los valores de la probabilidad condicional de repetir o cambiar la secuencia reforzada en el ensayo previo sugieren que, para la condición ganar-repetir la conducta-consecuencia del ensayo $n - 1$ sí tomó el control de la secuencia emitida en el ensayo n . Mientras que para la condición ganar-cambiar, la conducta-consecuencia sólo tomó el control de la secuencia emitida en el siguiente ensayo para la mitad de los sujetos durante las últimas sesiones de dicha condición.

El dato que muestra que dos de los sujetos responden de acuerdo al patrón ganar-cambiar en las últimas sesiones de la segunda Fase en la que operó dicha condición, sugiere que es necesaria una mayor cantidad de ensayos para que los sujetos cambien después de que una respuesta es reforzada. Morgan (1974) y Schwartz (1992) mencionan que bajo procedimientos en los que la recompensa se entrega de manera aleatoria o en los que se refuerza variabilidad, se puede observar una tendencia a repetir la respuesta que es seguida de una recompensa. Sin embargo, a lo largo del entrenamiento, se va perdiendo dicha tendencia y en su lugar aparecen los patrones de respuesta que sí mantienen una relación de contingencia con la entrega del reforzador. En relación con los hallazgos de Morgan (1974) y Schwartz (1992) se podría esperar que al prolongar el número de sesiones

para la condición ganar-cambiar, los sujetos DN2 y DN3 empezaron a responder de acuerdo a la estrategia de respuesta ganar-cambiar y para los sujetos DN1 y DN5 dicho patrón se veía más estable. Sin embargo, la tendencia tan pronunciada de repetir después de ganar para los sujetos DN2 ($p = 1.00$) y DN3 ($p = 0.94$), pudo haber dificultado la adquisición del patrón opuesto, a pesar de que el número de sesiones para las Fases 3 y 4 se duplicó en relación al que tuvieron las Fases 1 y 2.

El análisis de la proporción de respuestas-reforzadores obtenidos durante el Entrenamiento, mostró que los animales recibieron casi la misma cantidad de reforzadores por responder con la secuencia izquierda-izquierda (II) que por responder con derecha-derecha (DD). A partir de este dato se puede decir que no existió preferencia por responder en alguna de las alternativas. Este resultado podría parecer obvio debido a que el programa asignaba los reforzadores con la misma probabilidad para ambas opciones de respuesta. No obstante, debido a que esta probabilidad fue independiente para cada ensayo, el hecho de haber encontrado una proporción de respuestas-reforzadores cercana al 0.5 es un dato importante para el presente estudio, puesto que bajo procedimientos semejantes, generalmente se encuentra que los sujetos responden de manera exclusiva o preferencial en una de las alternativas (Morgan, 1974; Shimp, 1975).

A pesar de que durante el Entrenamiento las estrategias ganar-repetir y ganar-cambiar no se reforzaron de manera diferencial, tampoco existieron restricciones impuestas por el programa para la ejecución de las secuencias de acuerdo a alguna de dichas estrategias. El análisis de la probabilidad condicional mostró que para dos de los animales (DN3 y DN5) existió una tendencia por repetir la secuencia previamente reforzada y para los otros dos (DN1 y DN2) se encontró una tendencia a repetir después de una de las

secuencias y a cambiar después de la otra. Para el sujeto DN2, la tendencia a repetir pudo haber sesgado la ejecución durante las Fases Experimentales ya que sin importar en qué condición se encontraba este sujeto repetía, con una probabilidad muy cercana a 1.0, la secuencia II después de que era reforzada.

El hecho de que se hayan establecido diferentes tendencias durante el Entrenamiento puede tener dos explicaciones. La primera de ellas es por lo que se conoce como un efecto de contigüidad del reforzador (véase Neuringer, Kornell & Olufs, 2001) o lo que Morgan (1974) llama un efecto de "fortalecimiento simple" (*simple strengthening*) de la respuesta por parte de la recompensa. Este efecto consiste en que los sujetos tienden a repetir la respuesta que fue seguida de recompensa a pesar de que no exista una contingencia de reforzamiento diferencial para dicho patrón de repetición. Esta tendencia ha sido encontrada en varios experimentos en los que la entrega del reforzador no depende de la repetición de la respuesta previa (McElroy & Neuringer, 1990; Morgan, 1974; Neuringer, 1991; Schwartz, 1992; Shimp, 1966, 1967, 1969; Silbelberg, Hamilton et al., 1978; Silbelberg & Williams, 1974). El efecto de fortalecimiento simple o de contigüidad con el reforzador predice la tendencia a repetir la respuesta previamente reforzada que se encontró durante el Entrenamiento, pero no puede dar cuenta de la tendencia a *cambiar* después de la entrega de reforzador que se encontró para el sujeto DN1 cuando II era reforzada y para el sujeto DN2 después de que era reforzada DD. Una explicación alternativa es que los *patrones* que aparecieron con más frecuencia durante el Entrenamiento, fueron los que eran seguidos de la entrega de recompensa con una mayor probabilidad. Es decir, la interacción de las respuestas de los sujetos con el programa de asignación aleatoria de los reforzadores, generó una secuencia de entrega de la recompensa diferente para cada sujeto; la cual podía incluir el

reforzamiento por la emisión de las respuestas de acuerdo a alguna de las estrategias. Ahora bien, un análisis de la manera en la que se modificó el programa de asignación de los reforzadores para cada sujeto (probabilidad condicional del programa de reforzar cada una de las posibles estrategias de respuesta), mostró que los patrones de repetición o cambio efectuados con mayor frecuencia por cada rata, fueron los que eran seguidos por la entrega de alimento con una mayor probabilidad (véase el Anexo 1).

Los resultados encontrados en el presente experimento: una superioridad en la ejecución del patrón ganar-repetir sobre el de ganar-cambiar, no resultan sorprendentes. Por ejemplo, a partir de la Ley del Efecto (Thorndike, 1911) se espera que exista una mayor probabilidad de repetir la respuesta previamente reforzada. De manera semejante, la propuesta de que el reforzador fortalece la respuesta emitida previamente (Nevin, 1974, 1979), sugiere que, cuando una conducta es seguida de la entrega de reforzador existe una "resistencia" a emitir una respuesta diferente, lo que se observa como un incremento en la probabilidad de repetir esa conducta.

Desde esta perspectiva podría parecer contradictorio esperar que un animal repita la conducta que no le fue reforzada y que después de recibir reforzador por emitir una conducta, exista una tendencia a emitir una distinta. Sin embargo, patrones como el de ganar-cambiar y perder-repetir, pueden considerarse acordes con la Ley del Efecto si se asume que la unidad conductual sobre la que tiene efecto la entrega de reforzador no es la conducta inmediata previa (por ejemplo, la secuencia II o DD), sino un *patrón conductual* que incluye: a) la conducta emitida en el ensayo previo (II o DD), b) su consecuencia (entrega o ausencia de reforzador), y c) la conducta ejecutada en el ensayo actual (II o DD). De esta manera, cuando la entrega del reforzador se hace contingente a la ejecución de un patrón como

ganar-cambiar, existe una tendencia a "repetir" *todo* el patrón. Por ejemplo, en el ensayo $n - I$ un sujeto es reforzado por emitir la secuencia II y en el ensayo n es reforzado por responder con la secuencia DD (cambiar). Si para este ejemplo, se considera que el reforzador tiene un efecto sólo sobre la secuencia emitida en el instante previo, entonces parece contradictorio que un sujeto responda con la secuencia DD, ya que se espera que vuelva a emitir II. Por el contrario, si se considera que el reforzador tiene un efecto sobre el patrón conductual *II- reforzador-DD*, entonces se puede decir que los animales repiten la conducta que los lleva a la obtención del segundo reforzador (el que es contingente a la emisión de la estrategia ganar-cambiar) al volver a emitir el patrón: *II- reforzador-DD*.

Si un reforzador contingente a la emisión de una estrategia afecta un patrón de respuestas, entonces se sugiere que el reforzador no sólo tiene efecto sobre la conducta inmediata previa. En varios trabajos (Catania, 1971; Catania, Sagvolden & Keller, 1989) se ha mostrado que el reforzador también tiene un efecto sobre las respuestas previas a la reforzada. El problema consiste en explicar el mecanismo por el cual el reforzador produce la selección de la conducta a la que es contingente, ya sea una respuesta simple o un patrón conductual (Wyne, 1994).

Algunos autores (Neuringer, Kornell & Olufs, 2001) sugieren que el reforzador parece tener dos efectos sobre la conducta, un efecto de contigüidad que produce una tendencia a repetir la respuesta inmediata previa a la entrega del reforzador (Morgan, 1974; Neuringer, Kornell & Olufs, 2001) y un efecto de contingencia que provee información acerca de la conducta que lleva a la obtención de la recompensa (Bloomfield, 1972; Neuringer, Kornell & Olufs, 2001; Shimp, 1975, 1981, 1982, 1983).

En el presente trabajo, en el caso de la adquisición de la estrategia ganar-repetir, resulta difícil separar un efecto de contigüidad del de contingencia debido a que ambos producen el mismo resultado: la repetición de la conducta. El dato que podría indicar una diferencia entre estos efectos es la comparación de la tendencia a repetir después de la entrega del reforzador durante la Fase de Entrenamiento (cuando este patrón era reforzado de manera *aleatoria*), en relación a la misma tendencia durante las Fases Experimentales en las que la estrategia ganar-repetir se reforzó con una probabilidad de 1.0. La comparación entre la ejecución durante el Entrenamiento y las Fases Experimentales, muestra que para tres de los sujetos (DN2, DN3 y DN5), hubo un incremento en la probabilidad de repetir la respuesta previamente reforzada, durante los tres primeros días de la Fase Experimental ganar-repetir (véase la Figura 8).

Con relación a la adquisición de la estrategia ganar-cambiar, el efecto de contigüidad y el de contingencia producen patrones distintos. Por contigüidad con el reforzador, se espera que el sujeto repita la secuencia previamente reforzada. Mientras que, por un efecto de contingencia se espera que los sujetos ejecuten la secuencia opuesta a la reforzada en el ensayo previo. Durante las Fases Experimentales, todos los sujetos tendieron a repetir la secuencia previamente reforzada y sólo los sujetos DN5 y DN1 emitieron la secuencia opuesta a la previamente reforzada durante las últimas 4 sesiones de las Fases 3 y 4 respectivamente.

El experimento realizado en el presente trabajo consiste en una simplificación de las variables que se pueden manipular con el procedimiento que se propone. La característica básica de este procedimiento consiste en un programa de asignación aleatoria de los reforzadores a las diferentes opciones de respuesta. El empleo de este programa cumple tres funciones: 1) la de

mantener una ejecución semejante para las alternativas de respuesta, 2) la de no reforzar diferencialmente las estrategias de respuesta, de tal forma que se pueden observar las tendencias "naturales" de los sujetos a repetir o a cambiar; y 3) la de presentar al azar los ensayos en los que se la entrega del reforzador depende de la ejecución de alguna de las estrategias; de manera que ante la imposibilidad de predecir en qué momento se presentará uno de estos ensayos, se puede observar si la conducta-consecuencia del ensayo previo toma el control sobre la respuesta que se emite en el siguiente ensayo.

En investigaciones futuras, manipular otros elementos como el tipo de estrategia que se pide (e.g., ganar-repetir/perder-cambiar o ganar-cambiar/perder-repetir), las características del reforzador contingente a la emisión de las estrategias (e.g., su probabilidad, demora, calidad o magnitud), o la operante con la que se trabaja (e.g., secuencias heterogéneas de dos respuestas con las que se ha encontrado que después de que una de estas secuencias es reforzada, existe una mayor tendencia a cambiar; Wasserman, et al. 1980; Wasserman, et al. 1983), puede aportar datos acerca de cómo una conducta-consecuencia toma el control de la respuesta que se emite en el siguiente ensayo y de observar los efectos de contingencia y contigüidad involucrados en el aprendizaje de las diferentes estrategias de respuesta.

“To those outside the field it may seem that reinforcement is a fairly simple thing. You give a subject something good for doing something, and it does it more; give a subject something unpleasant and it will stop what it was doing. Thorndike (1898) already had got that far. The problem is, in nearly one hundred years, we have not come much further than Thorndike’s functional relation. The problem (. . .) – how does an individual know what it is being rewarded for (the “assignment of credit” problem) – has remained largely untouched until rediscovered in artificial intelligence by Minsky (1961).”

Wynne, C., 1984

REFERENCIAS

Baum, W. (2001). Behavior and the general evolutionary process. *Behavioral and Brain Sciences*.

Bloomfield, T. M. (1972). Reinforcement schedules: contingency or contiguity?. R. M., Gilbert & J. R., Millenson, (Eds). *Reinforcement. Behavioral analyses*. New York: Academic Press.

Blough, D.S. (1959). Delayed matching in the pigeon. *Journal of the Experimental Analysis of Behavior*, 2, 157 – 160.

Blue, M. (1979). Learning theory and the evolutionary analogy. Manuscrito no publicado, Erindale College, Universidad de Toronto, EE. UU.

Catania, A. C. (1971). Reinforcement schedules: The role of the responses preceding the one that produces the reinforcer. *Journal of the Experimental Analysis of Behavior*, 15, 271-289.

Catania, A. C. (1997). *Learning*. New York, NY, EE. UU.: Prentice Hall.

Catania A. C., Sagvolden, T. & S  ller, K. J. (1988). Reinforcement schedules: Retroactive and proactive effects of reinforcers inserted into fixed-interval performances. *Journal of the Experimental Analysis of Behavior*, 49, 49 – 73.

Davison, M. (2001). What do reinforcers do to behavior?. [En red]. *Society for the Quantitative Analyses of Behavior. Abstracts of 2001 Conference Papers*. Disponible en: <http://sqab.psychology.org/abstracts-2001.html>.

Dickinson, A., Watt, A., & Griffiths, W. J. H. (1992). Free-operant acquisition with delayed reinforcement. *The Quarterly Journal of Experimental Psychology*, 45B, 3, 241-258.

Domjan, M. (1999). *Principios de aprendizaje y conducta* (Palos, E. & Núñez, J. L., Trads.). México, D.F.: International Thomson Editores, S. A. de C. V. (Trabajo original publicado en 1998, 4a. ed.).

Evenden, J. L. & Robbins, T.W. (1984). Win-stay behaviour in the rat. *Quarterly Journal of Experimental Psychology*, 36B, 1 - 26.

Fetterman, J. G. & Stubbs, D. A. (1982). Matching, maximizing and the behavioral unit: concurrent reinforcement of response sequences. *Journal of the Experimental Analysis of Behavior*, 37, 97-114.

Hiraoka, K. (1984). Discrete-trial probability learning in rats: Effects of local contingencies of reinforcement. *Animal Learning & Behavior*, 12, 343 - 349.

Hull, Langman & Gleen (2001). A general account of selection: Biology, Immunology and Behavior. *Behavioral and Brain Sciences*, 24, (2).

Keller, F. S. & Shoenfeld, W.N. (1975). *Fundamentos de psicología*. (Daurella N., Trad.). Barcelona, España: Editorial Fontanella, S. A. (Trabajo original publicado en 1950, 1a. ed.)

Kramer, S. P. (1982). Memory for recent behavior in the pigeon. *Journal of the Experimental Analysis of Behavior*, 38, 71-85.

Lattal, K. A. & Gleenson, S. (1990). Response acquisition with delayed reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes*, 16, 27 - 39.

Lieberman, D. A. (1993). *Learning. Behavior and Cognition*. (2a ed.). California, EE. UU.: Brooks/Cole Publishing Company.

Machado, A. (1992). Behavioral variability and frequency-dependent selection. *Journal of the Experimental Analysis of Behavior*, *58*, 241 – 263.

Mackintosh, N. J. (1974). *The psychology of animal learning*. London, G.B. : Academic Press Inc.

McElroy, E., & Neuringer, A. (1990). Effects of alcohol on reinforced repetitions and reinforced variations in rats. *Psychopharmacology*, *102*, 49-55.

Morgan, M.J. (1974). Effects of random reinforcement sequences. *Journal of the Experimental Analysis of Behavior*, *22*, 301-310.

Neuringer, A. (1991). Operant variability and repetitions as functions of interresponse time. *Journal of Experimental Psychology: Animal Behavior Processes*, *17*, 3-12.

Neuringer, A. (1992). Choosing to vary and repeat. *Psychological Science*, *3*, 246 – 250.

Neuringer, A. (1993). Reinforced variation and selection. *Animal Learning and Behavior*, *12*, (2), 83-91.

Neuringer, A., Kornell, N. & Olufs, M. (2001). Stability and variability in extinction. *Journal of Experimental Psychology: Animal Behavior Processes*, *27*, 79-94.

Nevin, J. A. (1974). Response strength in multiple schedules. *Journal of the Experimental Analysis of Behavior*, *21*, 389 – 369.

Nevin, J. A. (1979). Overall matching versus momentary maximizing: Nevin (1969) Revisited. *Journal of Experimental Psychology: Animal Behavior Processes*, *5*, (3), 300-306.

Page & Neuringer (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*, 429-452.

Randall, C. K. & Zentall, T.R. (1997). Win-stay / lose-shift and win-shift / lose-stay learning by pigeons in the absence of overt response mediation. *Behavioural Processes*, *41*, 227 - 236.

Schwartz, B. (1982). Interval and ratio reinforcement of a complex sequential operant in pigeons. *Journal of the Experimental Analysis of Behavior*, *37*, 349-357.

Schwartz, B. (1989). *Psychology of learning and behavior*. NY, EEUU: W. W. Norton & Company.

Schwartz, B. (1992). Failure to produce response variability with reinforcement. *Journal of the Experimental Analysis of Behavior*, *37*, 171 - 181.

Shimp, C. P. (1966). Probabilistically reinforced choice behavior in pigeons. *Journal of the Experimental Analysis of Behavior*, *9*, 443 - 455.

Shimp, C. P. (1967). Reinforcement of least-frequent sequences of choices. *Journal of the Experimental Analysis of Behavior*, *10*, 57 - 65.

Shimp, C. P. (1969). Optimal behavior in free-operant experiments. *Psychological Review*, *76*, (2), 97-112.

Shimp, C. P. (1975). Perspectives on the behavioral unit: Choice behavior in animals. Estes, W. K. (Ed). *Handbook of Learning and Behavior Processes. Conditioning and Behavior Theory*. Vol 2.

Shimp, C. P. (1976a). Short-term memory in the pigeon: relative recency. *Journal of the Experimental Analysis of Behavior*, *25*, 55 - 61.

Shimp, C. P. (1976b). Short-term memory in the pigeon: the previously reinforced response. *Journal of the Experimental Analysis of Behavior*, *26*, 487-493.

Shimp, C. P. (1981). The local organization of behavior: Discrimination of and memory for simple behavioral patterns. *Journal of the Experimental Analysis of Behavior*, *36*, 303-315.

Shimp, C. P. (1982). On metaknowledge in the pigeon: An organism's knowledge about its own behavior. *Animal Learning & Behavior*, *10* (3), 358-364.

Shimp, C. P. (1983). The local organization of behavior: Dissociations between a pigeon's behavior and self-reports of that behavior. *Journal of the Experimental Analysis of Behavior*, *39*, 61-68.

Silbelberg, A., Hamilton, B., Zirriax, J. M., & Casey, J. (1978). The structure of choice. *Journal of Experimental Psychology: Animal Behavior Processes*, *4*, (4), 368-398.

Silbelberg, A. & Williams, D. R. (1974). Choice behavior on discrete trials: A demonstration of the occurrence of a response strategy. *Journal of the Experimental Analysis of Behavior*, *21*, 315-322.

Skinner, B. F. (1938). *La conducta de los organismos*. Barcelona, España: Editorial Fontanella, S. A.

Skinner, B. F. (1981). Selection by consequences. *Science*, *213*.

Staddon, J. E. R. (1975). Learning as adaptation. *Handbook of Learning and Cognitive Processes: Vol. 2*.

Staddon, J. E. R. (1983). *Adaptive Behavior and Learning*. Cambridge University Press.

Staddon, J. E. R. (1988). Learning as inference. En R. C. Bolles & M. D. Beecher (Eds.). *Evolution and Learning* (pp. 59 – 77). Hillsdale, N J EE. UU.: Lawrence Erlbaum Associates.

Staddon, J. E. R. & Zhang, Y. (1989). Response selection in operant learning. *Behavioural Processes*, 20, 189 – 197.

Staddon, J. E. R. & Zhang (1991). On the assignment-of-credit problem in operant learning. En Commons, Grossberg & Staddon (Eds.). *Neural network models of conditioning action*.

Thorndike (1911). *Animal Intelligence: An Experimental study of the associative processes in animals*. Monograph supplement No. 8 of the Psychological Review, 551-553.

Thorndike, E. L. (1927). *The law of effect*. Proceedings of the National Academy of Sciences, 212-222.

Walter, G. (1986). *El cerebro viviente*. D.F., México: Fondo de Cultura Económica.

Wasserman, E. A., Deich, J. D., & Cox, K. E. (1983). The learning and memory of response sequences. En Commons, Herrnstein & Wagner (Ed.), *Quantitative Analyses of Behavior. Discrimination Processes. Vol. IV: Discrimination Processes* (pp. 99 – 113). Cambridge, MA: Ballinger.

Wasserman, E. A., Keith R. N., & Larew, M. D. (1980). Memory for sequences of stimuli and responses. *Journal of the Experimental Analysis of Behavior*, 17, 49 – 59.

Williams, B. A. (1972). Probability learning as a function of momentary reinforcement probability. *Journal of the Experimental Analysis of Behavior*, 17, 363-368.

Williams, B. A. (1988). Reinforcement, choice, and response strength. En Atkinson, R. et al. (Eds). *Steven's Handbook of Experimental Psychology*. Vol. 2 Learning and Cognition.

Williams, B. A. (1991). Choice as a function of local versus molar reinforcement contingencies. *Journal of the Experimental Analysis of Behavior*, 56, 455 – 473.

Wyne, C. (1994). The return of the reinforcement theorists. En P. R. Killeen, (Ed). Mathematical principles of reinforcement. *Behavioral and brain sciences*, 17, 105 – 172.

Anexo 1

Se calcularon las probabilidades condicionales de que el programa reforzara el patrón ganar-repetir [$P(II+ | II+)$ y $P(DD+ | DD+)$] o el patrón ganar-cambiar [$P(DD+ | II+)$ y $P(II+ | DD+)$] para cada uno de los sujetos. En la Tabla 3 se muestra el promedio de la probabilidad condicional de los 5 últimos días del Entrenamiento.

Tabla 3

ESTRATEGIA	PROBABILIDAD CONDICIONAL	SUJETO			
		DN1	DN2	DN3	DN5
GANAR REPETIR	$P(II+ II+)$	0.2012	0.4924	0.3738	0.3188
	$P(DD+ DD+)$	0.2827	0.2622	0.2999	0.2805
GANAR CAMBIAR	$P(DD+ II+)$	0.2344	0.0000	0.0582	0.0487
	$P(II+ DD+)$	0.1603	0.3005	0.1821	0.1036

Nota. $P(II+ | II+)$ = probabilidad condicional de reforzar II dado que II fue reforzada;
 $P(DD+ | DD+)$ = probabilidad condicional de reforzar DD dado que DD fue reforzada;
 $P(DD+ | II+)$ = probabilidad condicional de reforzar DD dado que II fue reforzada;
 $P(II+ | DD+)$ = probabilidad condicional de reforzar II dado que DD fue reforzada.

