

Los Clickers Ardientes¹

Steve Martin
Natural Encounters Inc.

Susan G. Friedman
Utah State University

Marian Kruse y Keller Breland fueron de los primeros entrenadores en usar los clickers hace cerca de 70 años. Desde entonces, este creador de ruido se ha convertido en una herramienta popular con los entrenadores de animales alrededor del mundo. Usado correctamente, el clicker es un preciso marcador de eventos que conduce a una comunicación clara acerca de la contingencia entre un comportamiento y sus consecuencias, fortaleciendo así la conducta que le sigue. Sin embargo, algunos entrenadores mal usan los clickers al hacerlos sonar repetidamente al azar sin una entrega contigua de un refuerzo primario o de respaldo después de cada click. Nosotros llamamos a esta práctica Los Clickers Ardientes. Sin el necesario y fuerte antecedente click-recompensa y sin un plan sistemático, ésta práctica común de solo hacer click a menudo conduce a la extinción de la fuerza del refuerzo del clicker, a un progreso pobre de aprendizaje y a la agresión.

De forma clara, la comunicación bidireccional es el pilar del éxito en el entrenamiento animal. A través de una clara comunicación entrenadores expertos moldean de forma fluida las respuestas de una aproximación a otra, resultando en un comportamiento nuevo y complejo en minutos en vez de semanas. Una de las más importantes herramientas de comunicación es el refuerzo condicionado, también conocido como refuerzo secundario, marcador de eventos, marcador, estímulo de puente y puente. Los refuerzos condicionados mejoran la comunicación bidireccional porque pueden ser entregados en el instante en el que la conducta correcta ocurre. Esta estrecha asociación temporal entre la conducta y el refuerzo es una característica esencial de refuerzo efectivo conocido como contigüidad.

Marian Kruse y Keller Breland fueron de los primeros entrenadores en usar los clickers para mejorar los resultados del entrenamiento hace más de 70 años. Junto con Bob Bailey y otros, exploraron una gran variedad de otros reforzadores condicionados también, como silbatos, luces, tacto y palabras. Éstos reforzadores condicionados son ahora comunes en nuestros kits de herramientas de entrenamiento. Curiosamente, la función precisa de los reforzadores condicionados todavía se está investigando (para una buena discusión de los diferentes reportes, consulte Pierce y Cheney, 2008). Sin embargo, en un nivel práctico tenemos la información que más necesitamos saber acerca de los reforzadores condicionados:

- 1) Como conseguirlos – asociar estrecha y repetidamente un estímulo neutro con un reforzador bien establecido, es decir, respuesta (clásica).condicionada
- 2) Como romperlos – dejar de asociar el refuerzo condicionado con un refuerzo de apoyo, es decir, extinción de la respuesta.

Al consultar en zoológicos alrededor del mundo, a menudo vemos entrenadores inadvertidamente romper o debilitar, sus reforzadores condicionados por no respaldarlos con un reforzador bien establecido. Llamamos a este enfoque de entrenamiento Los Clickers Ardientes.

Los Clickers ardientes se definen como hacer un click rápido y asistemático en cada respuesta correcta en una serie de respuestas correctas, sin seguir cada click con un reforzador bien establecido, es decir, hacer click, sin premio.

Durante los muchos debates que hemos tenido con entrenadores que arden sus clickers, hemos llegado a creer que este enfoque resulta de varios malentendidos acerca de los procesos básicos del comportamiento relacionados con los reforzadores condicionados. El propósito de este trabajo es mejorar la efectividad de tu entrenamiento, abordando los cinco malentendidos más comunes asociados con los clickers ardientes y sumar nuestras voces a aquellos entrenadores que recomiendan asociar cada, o casi todos los clicks con un reforzador de apoyo bien establecido (véase, por ejemplo, Fernández, 2001; Ramírez, 1999; B. Bailey, comunicación personal, 17 de abril de 2011, K. Pryor, comunicación personal, 16 de abril de 2011).

TERMINOS

Para facilitar la comunicación a lo largo de este documento utilizamos la siguiente manera abreviada:

1. La palabra **click** se refiere a cualquier reforzador condicionado utilizado en el entrenamiento para reforzar un comportamiento con súper contigüidad. Se utiliza como sinónimo de reforzador condicionado o secundario, estímulo de puente, puente, marcador de eventos y el marcador.
2. La palabra **premio** se refiere a cualquier reforzador bien establecido, condicionado o incondicionado, utilizado para condicionar y mantener la fuerza del refuerzo del clic. Premio se utiliza como sinónimo de **reforzador de apoyo** (más a menudo el reforzador de apoyo en entrenamiento de animales es comida).
3. El término **clicker ardientes** se refiere a la práctica de hacer click repetidamente sin entregar sistemáticamente el refuerzo de apoyo, también referido a un solo click.

MALENTENDIDOS COMUNES

Malentendido #1 – Los clickers ardientes es un buen enfoque porque el clicker es un refuerzo (un refuerzo secundario), así el animal no necesita otro (el premio).

Algunos entrenadores dicen que no necesitan seguir el click con un premio porque el clicker no es solo un marcador o un puente, sino un refuerzo secundario de buena fe también. ¿Porque dar dos refuerzos cuando se puede dar uno? Es cierto que un refuerzo secundario bien-condicionado puede ser tan, o incluso más fuerte, que un refuerzo primario, dando un largo y fuerte historial condicional. Sin embargo, una diferencia fundamental entre los refuerzos primarios y secundarios es que los refuerzos primarios se refuerzan de forma automática – pre-instalado por así decirlo; los reforzadores secundarios dependen de la experiencia, específicamente asociaciones cercanas repetidas con otros reforzadores bien establecidos para adquirir y mantener su fuerza de refuerzo. De hecho, el procedimiento para retornar un reforzador secundario a su estado neutral es des-asociando, es decir, entregando el refuerzo secundario repetidamente sin un refuerzo de apoyo, conocido como extinción de la respuesta (un estímulo condicionado, EC, es presentado sin el subsecuente estímulo no condicionado, EN).

En algún lugar entre la asociación constante y la no asociación está el debilitamiento progresivo del refuerzo secundario. Mientras los refuerzos secundarios tienen una “vida útil”, el período de esa vida útil no es conocida y los refuerzos secundarios que tienen una larga vida útil son el resultado de una historia consistente de click-premio formado por decenas, si no es que cientos, de asociaciones (Pierce y Cheney, 2008). Cada vez que ocurre un click sin un refuerzo de apoyo, es casi literalmente un ensayo de la extinción de respuesta y los refuerzos secundarios pueden perder su fuerza para reforzar muy rápidamente, un problema que hemos observado muchas veces. Como el click no logra prever de manera fiable un premio, los animales analizan otras pistas del medio ambiente de que el refuerzo de apoyo (comida) está en camino, como el sutil movimiento de la mano del entrenador hacia la bolsa o cubeta de premios. De hecho, en el caso donde los entrenadores trabajan muy de cerca con sus animales, a menudo vemos a los animales responder al lenguaje corporal de sus entrenadores antes o después de que el click suene (independiente del click). Puede ser que dada la estrecha proximidad, muchos animales responden más rápido a lo que ven que lo que escuchan.

Malentendido #2 – Los Clickers ardientes hacen el entrenamiento más interesante para el animal. Si premias cada vez que haces click, la sesión es demasiado predecible y los animales se aburren con el entrenamiento.

Algunos entrenadores nos han explicado que usar esos clickers ardientes es una buena manera de mantener a los animales interesados en el entrenamiento, para alejar el aburrimiento producido por la repetición monótona del constante click-premio. Es cierto que la variedad es la sazón de la vida pero pensamos que la sazón debe venir de la variedad y cantidad de refuerzos que se ofrecen, los comportamientos que entrenas, y el ritmo con el que se los entrena, más que de clickers ardientes.

Imagina encontrar tu refrigerador bloqueado 3 ó 4 veces a la semana solo para mantener las cosas interesantes para ti. Vale la pena considerar una hipótesis diferente del porque un animal puede volverse desatento en una sesión de entrenamiento, la cual llamaremos comportamientos ardientes. Ese es el disparo rápido de señalamientos de respuestas rutinarias, respuestas que conducen a ninguna habilidad importante ni mejora la calidad de vida del animal. Por ejemplo, de manera rutinaria vemos sesiones de entrenamiento

compuestas de múltiples señales entregados rápidamente para tocar el target con diferentes partes del cuerpo, cada toque dura solo una fracción de segundo. Esto suena algo así:

“Gracie, brazo-click, dedo-click, hombro-click, oreja-click, pie-click, rodilla-click, espalda-click, bieeeeeeeneen, premio-premio-premio.”

¡Hablando de monotonía! Cuando se observa este tipo de sesión de entrenamiento, nos preguntamos exactamente cuál es el propósito de enseñar a un animal a tocar tantas partes del cuerpo en una sucesión rápida en menos de 20 segundos. El targeting es más útil cuando el toque se mantiene durante un período de tiempo. Como un comportamiento de mayor duración, el targeting puede ser fácilmente aprovechado para importantes comportamientos médicos y de manejo.

Malentendido #3 – Los Clickers ardientes construyen comportamientos más fuertes que la constante asociación click-premio porque la asociación inconstante es un programa variable de refuerzo como una máquina tragamonedas.

Algunos entrenadores piensan que los clicker ardientes es un programa variable de refuerzo que debe llevar a un comportamiento más fuerte desde que el refuerzo de apoyo es retenido intermitentemente. Un programa variable es uno de varios programas intermitentes de refuerzo donde el número de respuestas (o intervalo de tiempo, duración, etc.) requieren de cambios de refuerzo en torno a un promedio establecido. Es cierto que los programas intermitentes construyen la persistencia del comportamiento fluido, es decir, el comportamiento es más lento de extinguirse. Sin embargo, hay dos malentendidos enrollados dentro de esta lógica para los clicker ardientes. Primero, si el clicker es realmente un refuerzo condicionado efectivo, reteniendo el premio no cambia el hecho de que tú todavía estás utilizando un programa de refuerzo continuo de clicks. Si el click no es un refuerzo condicionado efectivo, nos enfrentamos a la posibilidad muy real de que el click es un ruido sin sentido que el animal tiene que clasificar para encontrar la contingencia conducta-consecuencia.

También vale la pena considerando que no hay ningún valor absoluto o inherente en la construcción de la persistencia del comportamiento cuando no se necesita. Los comportamientos señalados son un ejemplo de ello: Tienes que presentarles una señal, así que ¿por qué perder la oportunidad de incrementar la cantidad diaria de reforzamiento para un animal utilizando un programa variable? Cuando se requiere persistencia, el mejor enfoque es enseñar primero el nuevo comportamiento con un refuerzo continuo (click-premio) para una comunicación más clara de la contingencia comportamiento-consecuencia. A continuación, gradualmente se afinan los refuerzos a lo largo del tiempo (conocido como estiramiento de la proporción de refuerzo) del programa variable deseado cambiando la cantidad del comportamiento impredeciblemente, mientras que se incrementa la cantidad del comportamiento requerido por un refuerzo global. Por ejemplo, si un entrenador quiere que un león haga varios viajes hacia la ventana de vista del público cada día, un programa de proporción de refuerzo variable (es decir, el click-premio juntos, ¡no solo hacer click!) sería la herramienta correcta. Comenzando con un programa continuo para conseguir un alto rango de paseos por la ventana primero, el entrenador puede entonces gradualmente afinar el refuerzo de click-premio al requerir un incremento, aunque variable, número de pasadas para obtener el refuerzo. La implementación de esta estrategia de entrenamiento toma tiempo y una planificación cuidadosa para mantener la tasa de refuerzo lo suficientemente alta para que el león siga participando en el entrenamiento. Un programa de duración variable puede ser utilizado para incrementar la longitud de tiempo que el león se quede en la ventana.

Malentendido #4 – Los Clickers ardientes reducen la agresión por frustración porque el animal aprende a no esperar un premio cada vez. Al contrario, todo el infierno se desatará si te quedas sin comida antes de que la sesión haya terminado.

Algunos entrenadores han expresado la preocupación de que los animales entrenados con un historial de constante click-premio se vuelven agresivos cuando no se les presenta un premio. Una manera de resolver este problema es asegurándose que el premio se presente siempre después de un click. Para esto es necesario planificar la cantidad correcta de los refuerzos de apoyo y analizar para sacarlos cuidadosamente durante cada sesión de entrenamiento, o terminar antes de tiempo una sesión de entrenamiento porque te quedaste sin comida (algo que debería suceder solo una vez). También puedes planificar para una sesión de entrenamiento

más corta con grandes cantidades de refuerzos y un menor número de comportamientos o repeticiones. De esta manera mejoras la motivación de tu aprendiz y te ayuda a evitar las repeticiones rutinarias de comportamiento ardientes descritos arriba, en el #2.

Es posible enseñar a un animal a que ofrezca un montón de comportamientos principalmente por los refuerzos secundarios, véase por ejemplo, Alferink, Crossman, & Cheney, 1973, el cual describe el proceso por el cual un refuerzo condicionado, una sola luz de embudo, llegó a mantener a las palomas picoteando el disco 300 veces en presencia de comida gratis. Sin embargo, tanto el refuerzo condicionado como el secundario requieren la implementación sistemática de un plan que incluya cientos de asociaciones click-premio, un fuerte refuerzo de apoyo, y eventualmente programas variables cuidadosamente entregados para evitar el estiramiento de la proporción de refuerzo muy bruscamente o demasiado lejos (conocido como proporción de tensión). Dicho enfoque estructurado es muy diferente de elegir al azar para no seguir cada clic con un refuerzo de apoyo valioso basado en la relación hipotética entre el entrenamiento constante de click-premio y la agresión animal. Por otro lado, hay abundante información de que un programa de extinción puede provocar agresión (llamado extinción inducida o agresión inducida por frustración). Esto es una preocupación cuando un click ha perdido su fuerza como refuerzo de solo hacer click y el premio está siendo retenido de manera intermitente.

Malentendido #5 – Los Clickers ardientes son buenos para decirle a un animal que lo que acaba de hacer es correcto y el animal deberá seguirlo haciendo. Ellos son en definitiva lo suficientemente inteligentes para aprender que un click significa diferentes cosas.

Por supuesto no hay problema para enseñar a los animales una señal de seguir adelante (SSA). Sin duda es emocionante ver a un león marino responder a la SSA para nadar otra vuelta-rápida, o a una guacamaya volando unas pocas vueltas más amplias alrededor del teatro, o a un elefante manteniendo su pierna en posición para un recorte de la uña del pie. Sin embargo es un problema, cuando una señal, el click, es utilizado para significar dos comunicaciones completamente diferentes. Una luz roja del semáforo no puede señalar eficazmente a los conductores para que frenen y también para acelerar. ¡Gracias a Dios tenemos luces rojas y verdes!

Es una comunicación muy poco clara tener que el mismo click signifique dos cosas completamente diferentes, tales como que la comida viene en camino y seguir haciendo este comportamiento. Hemos visto a animales alejarse de la estación de entrenamiento (eso sí que es una comunicación algo clara), cuando el mismo click se utilizó de ambas maneras.

Un click bien condicionado puede muy bien servir más que una función para un comportamiento determinado. Puede marcar el comportamiento correcto así el animal aprende qué repetir para obtener de nuevo un refuerzo comestible, puede marcar el puente en el comportamiento para el refuerzo comestible, y puede ser un estímulo discriminativo para terminar el comportamiento y prepararse para la comida. Un SSA bien-condicionado no interrumpe la fluidez del comportamiento, por definición. Cuando se utiliza un SSA, una señal diferente es necesaria para comunicar, “¡La comida viene ahora!”

CONCLUSIÓN

Clickers, silbatos y otros refuerzos condicionados son herramientas valiosas que ayudan a los entrenadores a comunicar a los animales la respuesta precisa que necesitan repetir para conseguir un premio. Cuando un refuerzo condicionado es asociado sin problema con un bien-establecido refuerzo de apoyo entonces la comunicación es clara, la motivación se mantiene alta y los comportamientos son aprendidos rápidamente. Sin embargo, cuando un click no es asociado sistemáticamente con un refuerzo de apoyo, la comunicación se vuelve poco clara, como se evidencia por la disminución de la motivación, aumento de la agresividad y bajo rendimiento

Cuando el click comienza a perder significado a causa de un uso repetido y sin un premio, los animales comienzan a buscar otros estímulos para predecir sus resultados. A menudo observan pistas del lenguaje corporal que predicen el premio es inminente así se fortalece aún más la contingencia comportamiento- consecuencia y el clic es sólo ruido. Aunque es verdad un reforzador secundario no pierde su capacidad para fortalecer el comportamiento la primera vez que se usa sin un refuerzo de apoyo, el número de solo hacer click para la extinción no se puede predecir, y puede suceder muy rápidamente. Así que, mientras sea posible hay que alejarnos del ocasional click solo, los clickers ardientes no es la mejor práctica de entrenamiento. Cuando el click no lleva información de la que el animal puede depender, el resultado es un comportamiento poco confiable.

1 Martin, S. & Friedman, S.G. (2011, Noviembre). *Los clickers ardientes*. Trabajo presentado en el Congreso de la Alianza por el Manejo de la Conducta Animal, Denver. Co.

REFERENCIAS

Alferink, L.A., Crossman, E. K., & Cheney, C.D. (1973). Control of responding by a conditioned reinforcer in the presence of free food. *Animal Learning and Behavior*, 1, 38-40.

Fernandez, E.J., (2001). Click or Treat: A Trick or Two in the Zoo. *American Animal Trainer Magazine*, 2, 41-44. Shedd Aquarium

Pierce, W. D., and Cheney, C.D. (2008). *Behavior analysis and learning*. (4th ed.). New York, . NY: Psychology Press, 221-240.

Ramirez, K. (1999). *Animal training: Successful animal management through positive reinforcement*. Chicago, IL: Shedd Aquarium: p.14.