

Blazing Clickers

Marian Kruse e Keller Breland estão entre os primeiros treinadores a usar clickers há cerca de setenta anos atrás. Desde então, este aparelho transformou-se num instrumento popular entre os treinadores de animais de todo o mundo. Usado correctamente, o clicker é um marcador preciso, de um evento que leva a uma comunicação clara, acerca da contingência entre o comportamento e a sua consequência, reforçando o comportamento que lhe segue. No entanto, alguns treinadores fazem um mau uso dos clickers, fazendo-os soar repetidamente sem uma entrega contígua de um primário ou de qualquer outro reforço para além do próprio som do clicker. Nós chamamos a esta prática *blazing clickers*. Sem um historial forte e obrigatório de click-comida e sem um plano sistemático, esta prática de apenas clicar sem a entrega de um reforço primário, leva à extinção clássica da força que o clicker detém como reforço, assim como a um progresso muito fraco na aprendizagem ou mesmo a agressão.

Uma comunicação clara e bilateral é a base de um treino animal bem sucedido. Através de uma comunicação clara, treinadores experientes, moldam com fluidez as respostas de um animal de uma aproximação para a outra, resultando num comportamento complexo e novo em minutos ao invés de semanas. Uma das ferramentas de comunicação mais importantes é o reforço condicionado, também conhecido por reforço secundário, marcador de eventos, marcador, estímulo que faz ponte e ponte. Reforços condicionados melhoram a comunicação bilateral porque podem ser entregues no instante em que o comportamento correcto ocorre. Esta associação temporal próxima entre o comportamento e o reforço é uma característica essencial para um reforço eficaz conhecida por contiguidade.

Marian Kruse e Keller Breland estiveram entre os primeiros treinadores de animais a usar clickers para melhorar o resultado dos treinos há mais de setenta anos. Junto com Bob Bailey e outros, eles exploraram uma grande variedade de outros reforços condicionados, tais como apitos, luzes, toque e palavras. Estes reforços condicionados são hoje em dia comuns nas nossas caixas de ferramentas de treino.

O que é interessante, é que a função precisa dos reforços condicionados ainda está a ser investigada (para uma boa discussão dos diferentes pontos de vista ver Pierce e Sheney, 2008).

No entanto, a um nível prático, nós já temos a informação de que precisamos acerca dos reforços condicionados:

- 1.** Como fazê-los - associar um estímulo neutro repetidamente com um reforço bem estabelecido, isto é condicionamento clássico ou respondente.
- 2.** Como desfazê-los - parar de associar um reforço condicionado com outro reforço, isto é extinção respondente.

À medida que trabalhamos em vários jardins zoológicos pelo mundo fora, observamos treinadores que inadvertidamente desfazem ou diminuem a força dos seus reforços

condicionados ao não associá-los com outro reforço bem estabelecido. A esta forma de treinar chamamos *blazing clickers*. *Blazing clickers* é definido pelo click aleatório e rápido de cada resposta correcta dentro de uma série de respostas correctas, sem seguir cada click com o seu reforço antecipadamente bem estabelecido (isto é, clicar sem recompensar).

Acerca das muitas discussões que tivemos com treinadores que clicam sem recompensar, acreditamos que esta forma de trabalhar resulta de diferentes conceitos errados acerca de processos básicos de comportamento relacionados com os reforços condicionados. O propósito deste artigo é melhorar a eficácia do seu treino falando dos cinco conceitos errados mais comuns associados aos *blazing clickers*, e adicionar as nossas vozes às dos treinadores que recomendam associar todos, ou quase todos os cliques a uma recompensa (ver por exemplo, Fernandez 2001; Ramirez 1999; Bob Bailey, comunicação pessoal 17 de Abril 2011, Karen Prior comunicação pessoal em 16 de Abril de 2011).

Termos

Para uma maior facilidade de comunicação durante este artigo, usaremos os seguintes termos:

1. A palavra click refere-se a qualquer reforço condicionado usado no treino para reforçar um comportamento com super contiguidade. É sinónimo de reforço condicionado ou secundário, estímulo que faz ponte, ponte, marcador de eventos e marcador.
2. A palavra recompensa refere-se a qualquer reforço bem estabelecido, condicionado ou não, usado para condicionar e manter a força do clicker. Recompensa é sinónimo de reforço bem estabelecido que, no caso do treino de animais, a maioria das vezes é comida.
3. O termo *blazing clickers* refere-se à prática de repetidamente clicar sem sistematicamente entregar o reforço, também conhecido por cliques isolados ou clicar sem recompensar.

Conceitos Errado mais comuns

Conceito Errado #1 – *Blazing Clickers* é uma boa aproximação porque o clicker é um reforço (um reforço secundário), por isso o animal não precisa de outro (recompensa).

Alguns treinadores afirmam que não precisam de seguir o click por uma recompensa porque o clicker é não só um marcador ou uma ponte, mas um genuíno reforço secundário. Porquê entregar dois reforços, quando apenas um é suficiente? É verdade que um reforço secundário bem condicionado pode ser tão forte, ou até mais forte, do que um reforço primário desde que exista um historial de condicionamento longo e forte. No entanto, uma diferença crítica entre um reforço primário e um reforço secundário é que um reforço primário é automaticamente reforçante – isto é, pré-programado; os reforços secundários, por sua vez, dependem da experiência, especificamente da associação repetida e seguida com outros reforços previamente bem estabelecidos, para adquirirem e manterem a sua força como reforços. Na verdade, o procedimento de tornar um reforço secundário ao seu estado neutro é desfazer a associação, isto é, repetidamente entregar o reforço secundário sem o reforço que

o apoia, processo conhecido por extinção respondente (um estímulo condicionado, CS, é apresentado sem o subsequente estímulo não condicionado US).

Alguns entre a associação consistente e a não associação, existirá o progressivo enfraquecimento do reforço secundário. Enquanto que os reforços secundários têm um “tempo de vida” esse tempo não é conhecido e aqueles reforços secundários que têm um “tempo de vida” longo, são resultado de um historial constituído por dezenas, se não centenas de associações (Pierce & Cheney, 2008). Cada vez que um click ocorre sem um reforço de apoio, é literalmente o mesmo que uma exposição a uma extinção respondente, e reforços secundários podem perder a sua força de reforço muito rapidamente, um problema que já observámos muitas vezes. À medida que o click deixa de garantir antecipar uma recompensa, os animais começam a investigar o ambiente para tentar encontrar pistas que lhes indiquem que o reforço (comida) está a caminho, tal como o movimento subtil do braço do treinador em direcção à bolsa ou balde de comida. Na verdade, no caso dos treinadores que trabalham fisicamente perto dos animais, muitas vezes vemos animais responderem à linguagem corporal do treinador, segundos antes do mesmo soar o clicker. Pode ser que ao trabalhar em proximidade, muitos animais respondem com muito maior rapidez ao que veem do que àquilo que ouvem.

Conceito Errado #2 – *Blazing Clickers* tornam o treino mais interessante para o animal. Se você recompensar sempre que clicar, a sessão torna-se muito previsível e os animais ficam aborrecidos com o treino.

Alguns treinadores explicaram-nos que o uso do *Blazing Clickers* é uma boa forma de manter os animais interessados no treino, de evitar o aborrecimento produzido pela sequência click recompensa. É verdade que variedade apimenta a vida, mas nós acreditamos que a pimenta deve vir da variedade e quantidade de reforços que você providencia, dos comportamentos que você treina e do ritmo que você imprime ao treino, ao invés dos *blazing clickers*.

Imagine o que é encontrar o seu frigorífico trancado 3 a 4 vezes por semana só para que as coisas se tornem mais interessantes para si. Uma outra hipótese para o porquê de um animal se pode tornar menos atento durante uma sessão de treino, pode ser aquilo que chamamos de “blazing behaviours”. Isso é o sinalizar seguido de respostas mundanas, respostas essas que não levam ao desenvolvimento de nenhuma habilidade específica ou que não melhoram de forma nenhuma a qualidade de vida do animal. Por exemplo, rotinamente vemos sessões de treino que se limitam a pedir rapidamente ao animal múltiplos comportamentos de target com diferentes partes do corpo, cada uma durando apenas uma fracção de um segundo. É mais ou menos como isto:

“Gracie, braço-click,dedo-click,ombro-click,orelha-click,pé-click,joelho-click,costas-click,
booooooooooom, recompensa, recompensa, recompensa”.

Isto sim é aborrecido! Quando observamos este tipo de treino a ser efectuado, encontramos-nos a questionar qual o objectivo de ensinar o animal a fazer “target” com tantas partes do corpo numa rápida sucessão de menos de 20 segundos? “Targetting” é extremamente útil quando o toque tem alguma duração. Quanto mais o comportamento tiver duração, mais fácil é levar o target a uma base para um bom tratamento médico ou “grooming”.

Conceito Errado #3 – *Blazing Clickers* controem comportamentos mais fortes do que consistentemente associar o click com recompensa porque a associação inconsistente é um esquema de reforço variável como uma máquina de casino.

Alguns treinadores acreditam que *blazing clickers* é um esquema de reforço variável que levará a comportamentos mais fortes uma vez que a entrega do reforço que o apoia é retida. Um esquema variável é um de muitos esquemas de reforço onde o número de respostas (com intervalo de tempo, duração, etc.) requeridas para adquirir o reforço, varia de acordo com uma média pré-estabelecida. É realmente verdade, que esquemas de reforço variáveis criam comportamentos mais persistentes naqueles comportamentos já fluentes, isto é, o comportamento torna-se mais lento a ser extinto. Existem, no entanto, dois erros neste raciocínio dos *blazing clickers*. Primeiro, se o clicker é realmente um reforço condicionado eficaz, não oferecer a recompensa não vai mudar o facto que você continua a usar um esquema de reforço contínuo de clicks. Se o click não é um reforço condicionado eficaz, então estamos perante a possibilidade real que o click é, para o animal, apenas um som sem sentido que o animal tem que discernir para perceber a contingência entre comportamento-consequência.

Também vale a pena considerar que não existe nenhum valor inerente ou absoluto em constuir persistência num comportamento quando este não é necessário. Comportamentos contingentes de sinais são um desses casos: você tem que pedir o comportamento, então, porquê perder uma oportunidade em aumentar a quantidade diária de reforço ao usar um esquema variável? Casos em que a persistência seja necessária, a melhor forma de o conseguir, é primeiro ensinar o comportamento novo através de um esquema de reforço contínuo (click-recompensa) para uma comunicação clara entre a contingência comportamento-consequência. Depois, gradualmente tornar os reforços mais escassos (conhecido como esticar o ratio de reforço) até chegarmos ao esquema de reforço desejável, tornando dessa forma a quantidade oferecida de comportamento imprevisível, enquanto aumentando a quantidade de comportamento desejado em vista do reforço. Por exemplo, se um treinador quer que um leão faça várias visitas à area com vista para o público durante o dia, um esquema de reforço variável seria a ferramenta ideal (isto é, clicar e recompensar nunca clicar sem recompensar!). Começando com um esquema de reforço contínuo, pedindo gradualmente um aumento variável no número de vezes que o leão passa pelo local, para adquirir reforço. Implementar este tipo de estratégia de treino leva tempo e um plano cuidadoso para manter a quantidade de reforço suficientemente alta para que o leão continue a querer trabalhar. Um esquema de duração variável pode ser usado para aumentar o tempo que o leão fica em frente à janela.

Conceito Errado #4 – *Blazing clickers* diminuem a agressão por frustração porque o animal aprende a não esperar uma recompensa de cada vez que ouve o clicker. Tudo pode correr mal se você fica sem comida antes da sessão terminar.

Alguns treinadores expressaram a sua preocupação que os animais treinados com consistência no click e entrega da recompensa, tornarão-se-ão agressivos quando a recompensa não for apresentada. Uma forma de resolver este problema é assegurarmo-nos que a recompensa é sempre apresentada depois do click. Isto requer planear a quantidade certa de reforços primários e espaça-los cuidadosamente durante cada sessão de treino, ou terminar uma

sessão de treino mais cedo antes de acabar o reforço primário (algo que só deveria acontecer uma vez). Você também pode planejar uma sessão mais curta com maiores quantidades de reforço ou menos repetições de cada comportamento. Isto tudo pode melhorar a motivação do seu aluno e ajudá-lo a evitar a repetição mundana de comportamentos descrita no número 2 acima falado.

Então, é possível ensinar um animal a oferecer muitos comportamentos maioritariamente em troca de reforços secundários, veja por exemplo, Alferink, Crossman & Cheney, 1973, que descreve o processo através do qual um reforço condicionado, neste caso uma luz, conseguiu manter um pássaro a dar 300 bicadas na ausência total de comida. No entanto, condicionar um reforço secundário tão forte requer a implementação de um plano sistemático que inclui centenas de bicadas sempre seguidas por luz e recompensa, um reforço primário muito forte e eventualmente esquemas de reforço variáveis cuidadosamente implementados para que a diminuição do ritmo de entrega de reforço não seja muito abrupto ou escasso (conhecido como “ratio strain”). Uma aproximação destas tão estruturada é muito diferente da escolha casual de não seguir o clique por uma recompensa baseando-se num relacionamento hipotético de que clicar e recompensar sempre seguido pode causar agressão. Por outro lado, existe uma abundância de factos de que um esquema de extinção pode provocar agressão (chamado de extinção induzida ou agressão induzida por frustração). Isso é uma preocupação quando um click perde a força de reforço por causa do click ser soado e a recompensa ser retida.

Conceito Errado #5 – *Blazing clicker* é bom para dizer ao animal que o que ele fez está certo e que ele deve continuar. Eles são suficientemente espertos para aprender que o click significa coisas diferentes.

Claro que não existe problema nenhum em ensinar a um animal um sinal de continuação (Keep going signal – KGS). É certamente emocionante ver um leão-marinho responder ao KGS e nadar mais uma volta, uma ave a dar mais voltas ao local, ou ver um elefante a manter a posição da pata para se poder limar as unhas. O que é problemático, no entanto, é quando um sinal, no caso o click, é usado para comunicar duas coisas completamente distintas. Um sinal de trânsito vermelho comunica aos condutores para colocarem o pé no travão e também para acelerarem. Ainda bem que temos luzes vermelhas e verdes!

É uma comunicação muito fraca ter um mesmo click significar duas coisas completamente distintas e opostas tais como, vem aí comida e continua a fazer esse comportamento. Já observamos animais a abandonar as sessões de treino (e isso sim é comunicar claramente com o treinador) quando o mesmo som é usado para as duas coisas.

Um click bem condicionado pode ter mais do que uma função para um dado comportamento. Pode marcar o comportamento correcto para que o animal entenda o que deve repetir para obter reforço de novo, pode servir de ponte entre o comportamento e o reforço e pode ser um estímulo discriminativo para terminar o comportamento e preparar-se para a comida que aí vem. Um KGS bem condicionado, por definição, nunca interrompe o comportamento. Quando usamos um KGS um sinal diferente deve ser usado para significar “comida vem aí!”.

Conclusão

Clicker, apitos e outros reforços condicionados são ferramentas valiosas para ajudar treinadores a comunicar com os animais a resposta precisa que este necessitam efectuar para obter uma recompensa. Quando um reforço condicionado é consistentemente associado com outro reforço de apoio ou primário, então a comunicação torna-se clara, a motivação permanece alta e os comportamentos são rapidamente aprendidos. No entanto, quando o som do clicker não é consistentemente associado a um reforço primário, a comunicação torna-se confusa, como provado pela diminuição da motivação, aumento da agressão e enfraquecimento da performance.

Quando o click começa a perder o seu significado por causa do seu uso repetido sem associação a uma recompensa, os animais começam a procurar por outros sinais que prevêm a consequência das suas acções. Muitas vezes eles observam a linguagem corporal que lhes diga que uma recompensa está iminente e isto vem dar força à contingência comportamento-consequência e tornar o click apenas num som. Enquanto é verdade que um reforço secundário não perde a sua capacidade de reforçar comportamento logo da primeira vez que é usado sem reforço primário, o número de vezes que clicamos sem recompensa que levarão à extinção não podem ser previstos e isto pode acontecer muito rapidamente. Portanto enquanto um click sem recompensa esporádico pode passar incólume, *blazing clickers* não são uma boa prática de treino. Quando o click não transporta nenhuma informação da qual o animal possa depender o resultado é comportamento errático.

REFERÊNCIAS

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 of 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: Shed Aquarium: p14.