How clicker training works: Comparing Reinforcing, Marking, and Bridging Hypotheses

Lynna C. Feng*, Tiffani J. Howell, Pauleen C. Bennett

School of Psychology and Public Health, La Trobe University, P.O. Box 199, Bendigo, Victoria 3552, Australia

ARTICLE INFO

ABSTRACT

Clicker training is a popular technique used in companion animal training. It employs a handheld signalling device called a clicker, which emits an audible "click" noise when pressed. Trainers press the clicker when an animal performs a desired behaviour, usually following the click with presentation of a food reward. The clicker is purported to facilitate learning, but scientific evidence to support this claim is limited. Of five studies comparing a clicker-type signal + food group with a food-only control group, only one found that animals in the signal + food group learned faster. Further investigation is therefore required to better understand the circumstances under which clickers might help or hinder learning. To inform future studies, it is important to consider mechanisms by which the clicker may function. In this paper three proposed mechanisms are presented, which we term the Reinforcing Hypothesis, Marking Hypothesis, and Bridging Hypothesis. To begin understanding which (if any) of these three mechanisms is the means by which clickers may operate, we evaluate relevant laboratory animal studies. Based on available behavioural and neuropsychological evidence, it is concluded that clickers and other clicker-like stimuli likely function as conditioned reinforcers, but may also have marking and bridging properties. Ways to investigate how this translates to clicker use in applied settings are identified.

© 2016 Elsevier B.V. All rights reserved.

Contents

1. Where clicker training comes from

B.F. Skinner's Theory of Operant Conditioning proposes that animals learn to "operate" their world based on the consequences of their behaviours. According to this theory, behaviours followed immediately by a desirable consequence (reinforcement) become more likely to occur again, while behaviours followed immediately by an undesirable consequence (punishment) become less likely to occur (Skinner, 1938). Even very brief delays in the delivery of both desirable and undesirable consequences have been found to impair the rate at which animals learn to perform novel behaviours (as reviewed in Lattal, 2010). However, in laboratory studies, presenting a signal which predicts the consequences during the time delay can reduce this learning impairment. For example, Grice (1948) found that rats provided with immediate reinforcement were able to learn a black-white visual discrimination task in a median of

* Corresponding author.
E-mail address: l.feng@latrobe.edu.au (L.C. Feng).

20 trials, while those with a 5-s delayed reinforcement required a median of 580 trials. Consistent with the argument that a reward-predicting signal could reduce this learning impairment, rats with a 5-s delayed reinforcement that was signalled by an immediate reward-predictor signal required significantly fewer trials (median of 155 trials) to learn the discrimination task than the 5-s delayed reinforcement group without the reward-predictor signal (Grice, 1948).

In applied animal training outside of the laboratory, delay of consequences is often inevitable due to uncontrollable factors, such as the handler's position relative to the animal performing the behaviour. An immediate, reward-predicting, signal, like the one used by Grice (1948), seems likely to be beneficial in such situations, where immediate reinforcement is not feasible. Accordingly, many animal trainers have adopted such a reward-predicting signal. This technique was popularized as "clicker training" by dolphin trainer turned dog trainer, Karen Pryor, in her book called *Don't Shoot the Dog!* (Pryor, 1999; historical account by Gillaspy et al., 2014). Clicker training employs a clicker: a hand-held device that makes a clicking sound when pressed. The clicker is pressed when a desired behaviour occurs, and is typically followed by presentation of a food reward as soon as possible (Bailey and Bailey, 1996b; Ramirez, 1999; Pryor, 2005).

Among clicker training professionals, the clicker is often referred to interchangeably as a secondary reinforcer, marker, or bridging stimulus (Pryor, 2009). These terms are defined later in this paper, but it is important to acknowledge here that they are not synonymous and may be misleading, since each is suggestive of a different underlying mechanism. For the purposes of this review, the overarching term "SIGNAL" is used to refer to clickers and other stimuli used in the same context. A SIGNAL is any auditory, visual, or tactile stimulus which is not a primary reinforcer (holding intrinsic value), and which is administered immediately after a target behaviour. A PREDICTOR SIGNAL, meanwhile, is any SIGNAL which is, or has previously been, intentionally paired or conditioned to predict the presentation of a primary reinforcer.

2. Does clicker training facilitate learning?

Animal trainers suggest that use of a clicker should result in animals learning new tasks more quickly (Pryor, 1999). As evidence of this, many proof-of-concept studies have been conducted, demonstrating that training protocols using a previously conditioned PREDICTOR SIGNAL can be used to train a myriad of behaviours, such as teaching pigeons to play ping pong, dolphins to detect mines, and dogs to identify samples from cancer sufferers (Bailey and Bailey, 1996a; Willis et al., 2004; McCulloch et al., 2006). However, only a limited number of studies have empirically compared PREDICTOR SIGNAL use to a control group trained without an intermediary PREDICTOR SIGNAL, in order to assess whether or not such signals are, in fact, facilitating learning. The results of these studies are surprisingly inconclusive.

To our knowledge, there have been only four peer-reviewed articles and one academic thesis reporting studies which empirically evaluated the efficacy of a clicker-type SIGNAL as compared to a primary reinforcer-only control group in companion animal species. In all of these studies the audible SIGNAL was first paired with a primary reinforcer (either food or water), ostensibly meaning that the comparisons were between a PREDICTOR SIGNAL group and a control group. Of these studies, two were performed with horses (McCall and Burgin, 2002; Williams et al., 2004), one with dwarf goats (Langbein et al., 2007), and the remaining two with companion dogs (Smith and Davis, 2008; unpublished honour's thesis Blandina, n.d.). Only one of these five studies (Langbein et al., 2007) found that using a PREDICTOR SIGNAL resulted in higher rates of task acquisition when compared to using the primary reinforcer alone. Langbein et al. (2007) found that, for a shape discrimination task, the goats in a PREDICTOR SIGNAL group required significantly fewer trials to reach criterion compared to those in a control group. Conversely, Blandina (n.d.) reported that dogs in a food-only condition reached a higher shaping stage in a down-stay shaping protocol than those in a PREDICTOR SIGNAL group. Unfortunately, Blandina did not report inferential statistics to indicate whether or not these differences were significant. The three other studies comparing a PREDICTOR SIGNAL to a primary reinforcer-only control all assessed a nose-target task. None found a significant difference between a PREDICTOR SIGNAL group and a food-only control group in rates of task acquisition.

The above evidence suggests that there are contexts in which a PREDICTOR SIGNAL facilitates learning, as expected, but potentially other contexts where PREDICTOR SIGNALS have no impact or even inhibit learning. It is important to consider the possibility, however, that the PREDICTOR SIGNALS used in these applied studies were not the same as the PREDICTOR SIGNALS used in laboratory animal research. Skinner (1938) reported that a hungry rat learned that a SIGNAL predicted food in a single presentation. Pairing the SIGNAL with a less desirable primary reinforcer, such as flavoured water, however, meant that it took twenty presentations to elicit a response. Further, Skinner stated that he chose to present between 50 and 200 SIGNAL-primary reinforcer pairings to ensure sufficient conditioning (Skinner, 1938). Although all of the above companion animal studies specifically exposed subjects to the SIGNAL-primary reinforcer pairing before testing, no food deprivation protocols were implemented and, where reported, animals were exposed to just twenty SIGNAL-primary reinforcer pairings. None of the studies systematically evaluated whether this number of SIGNAL-primary reinforcer pairings effectively produced a PREDICTOR SIGNAL. In addition, these studies were conducted assessing different tasks with different species and methodologies. As such, there are innumerable potential reasons for the discrepant results. Further investigation is therefore required to better understand the circumstances under which clicker-type SIGNALS might aid or obstruct learning. To inform such investigations, a sound understanding of the mechanisms proposed to underlie SIGNAL function and, more specifically, PREDICTOR SIGNAL function, is imperative.

3. Proposed mechanisms of the clicker

Many mechanisms could potentially explain how SIGNALS facilitate learning. In this review, we discuss three which are reflective of the terminology most commonly used by applied animal trainers to describe clicker-type SIGNALS: secondary reinforcer, event marker, and bridging stimulus (Ramirez, 1999; Pryor, 2009). We refer to the corresponding hypotheses for these mechanisms as the Reinforcing Hypothesis, Marking Hypothesis, and Bridging Hypothesis, respectively.

Traditionally, PREDICTOR SIGNALS have been called secondary reinforcers (as suggested by Skinner, 1938). According to Skinner's Reinforcing Hypothesis (1938), a clicker (or other conditioned stimulus) becomes a secondary reinforcer by taking on the reinforcing quality of the primary reinforcer (usually a food reward) with which it is paired. Alternatively, Lieberman et al.'s Marking Hypothesis (1979) suggests that salient cues or SIGNALS, presented in close temporal and contextual proximity to a goal behaviour, emphasize that instant in time, which then facilitates learning. Finally, Kaplan and Hearst (1982) proposed the Bridging Hypothesis, which suggests that a stimulus or SIGNAL presented between the animal performing a target behaviour and the subsequent reward acts to fill or "bridge" the gap between the behaviour and the consequence.

These three hypotheses are not mutually exclusive, and it is likely that more than one of these mechanisms contribute to the way in which SIGNALS facilitate learning, at least when reinforcement is delayed. However, each of the hypotheses leads to different predictions, meaning that they can be tested and compared. Based on the Reinforcing Hypothesis, a SIGNAL, which has been repeatedly paired with a primary reinforcer, becomes a PREDICTOR SIGNAL which takes on the reinforcing properties of the primary reinforcer with which it was paired. It follows that the PREDICTOR SIGNAL would then be expected to reinforce behaviours even when it is not subsequently followed by the primary reinforcer. This is not consistent with the Marking Hypothesis, however, since a marking SIGNAL is not expected, much less required, to have any reinforcing qualities to be effective. In the absence of this requirement, the SIGNAL would be expected to facilitate learning even without prior pairing with a primary reinforcer. According to the Bridging Hypothesis, meanwhile, SIGNALS which fill a large portion of any delay between behaviour and primary reinforcer might be expected to be more effective than those that fill a shorter portion of the delay. What follows is a discussion of the merits of these three hypothesised mechanisms, based on evidence from studies conducted to assess their relevance to explaining how SIGNALS function.

3.1. The Reinforcing Hypothesis

Secondary reinforcement refers to the use of a "secondary reinforcer", as a consequence, in order to increase the likelihood of a behaviour re-occurring. Unlike a primary reinforcer (e.g. food) which has intrinsic value, a secondary reinforcer is a previously neutral stimulus (such as the "click" of a clicker), which, upon repeated pairings with a primary reinforcer (e.g. 135 pairings in Egger and Miller, 1962), is thought to take on the reinforcing value of the primary reinforcer (Skinner, 1938). That is, it becomes a sort of "reward" in itself. Theoretically, any perceivable, neutral stimulus could be conditioned to become a secondary reinforcer. Thereafter, it would be able to provide immediate reinforcement to the animal regardless of its location or position. Spence (1947) proposed that all learning in delayed reinforcement paradigms is a result of conditioned secondary reinforcers providing immediate reinforcement. In this context, a PREDICTOR SIGNAL with reinforcing properties provides an alternative source of immediate reinforcement, which mitigates the effects of primary reinforcement delay.

The standard methods by which to assess whether or not a SIGNAL functions as a secondary reinforcer are to measure frequency of task performance, resistance of the task to extinction, and ability of the SIGNAL to condition a new behaviour (as reviewed by Williams, 1994). Skinner (1938) suggested that a PREDICTOR SIGNAL should increase response frequency compared to using a primary reinforcer alone, and that, even when presented alone, the PREDICTOR SIGNAL should reinforce novel behaviours with no history of being directly reinforced by a primary reinforcer. In addition, Bugelski (1938) proposed that a PREDICTOR SIGNAL's reinforcing properties should result in behaviours' increased resistance to extinction. Researchers have specifically investigated the ability of a PREDICTOR SIGNAL to function in each of these three capacities (see Table 1).

Table 1 summarizes the findings of a number of studies which assessed one or more of the Reinforcing Hypothesis predictions. These studies overwhelmingly confirm that use of a PREDICTOR SIGNAL increases rate of responses, increases resistance to extinction, and successfully reinforces novel task performance. For example, Egger and Miller (1962) tested how a PREDICTOR SIGNAL affected the rate of rats' lever pressing behaviour. Rats were presented with 135 pairings of a SIGNAL followed by food. This conditioned PREDICTOR SIGNAL was then used in a reward-sequence (PREDICTOR SIGNAL + food) and compared to a neutral tone (no prior history of predicting rewards) + food control. The rats were tested on a known lever-pressing behaviour and demonstrated significantly higher rates of pressing when the consequence following lever-pressing was the PREDICTOR SIGNAL reward-sequence rather than the control. While the neutral tone may have gradually come to predict the reward throughout the course of the testing phase, this effect was apparently not strong enough to eclipse the stronger reinforcing effect that the previously conditioned PREDICTOR SIGNAL provided.

PREDICTOR SIGNALS have also been shown to increase a behaviour's resistance to extinction. In an extinction protocol, animals are given the opportunity to perform a previously reinforced behaviour, but the "ultimate" reinforcement is withheld (Skinner, 1938). The outcome is typically that the rate of behaviour performance rises for a short amount of time (called an extinction burst) then drops off to near zero levels (Skinner, 1933). Bugelski (1938) demonstrated that rats' bar pressing behaviour took longer to extinguish (i.e. their responses persisted for a longer period of time) in animals that heard a PREDICTOR SIGNAL after bar pressing than in those who were given no auditory feedback after the same behaviour. This evidence is congruent with the Reinforcing Hypothesis, as this resistance to extinction is attributed to the PREDICTOR SIGNAL providing reinforcement not provided to the control group. However, the behaviour eventually does extinguish, suggesting that the PREDICTOR SIGNAL reverts to being a neutral SIGNAL, unable to permanently maintain the behaviour in the face of withheld primary reinforcement (e.g. food reward).

Finally, as predicted by the Reinforcing Hypothesis, a conditioned PREDICTOR SIGNAL can be used on its own (without being immediately followed by a food reward) to train a novel task. Skinner (1938) demonstrated this in four rats who were presented with an auditory SIGNAL paired with a food reward. After 60 combined presentations of the sound and food, the rats were presented with a novel task: a lever was inserted in the testing cage. Any lever-presses were followed by the PREDICTOR SIGNAL, but without the subsequent food reward. Even though these rats had never been given primary reinforcement for pressing the lever, the resulting response curves were visually similar to rats undergoing an extinction protocol after being reinforced for the lever pressing behaviour: there was initially a high rate of responses, but this response rate dropped to nearly zero within 30–50 trials (Skinner, 1933).

Skinner argued that these results indicated that the PREDICTOR SIGNAL was functioning as a secondary reinforcer, which itself became extinguished (thus reverting to being merely a SIGNAL) when used without being followed by a food reward for a prolonged period (Skinner, 1938). These results suggest that, while the PREDICTOR SIGNAL can be used alone to reinforce a novel behaviour, practical applications would be limited by the fact that the PREDICTOR SIGNAL, and as a result the behaviour, become rapidly extinguished.

Research conducted by McCall and Burgin (2002) confirmed that a conditioned PREDICTOR SIGNAL cannot be used on its own for an extended period of time. Horses 'reinforced' with a PREDICTOR SIGNAL previously paired with food had a significantly higher rate of performance of a flap-pushing behaviour than control horses, who received a neutral SIGNAL, in the first 30 min of a testing phase, but performance fell to the same rate as the control group in a second 30-min session. This suggests that, while PREDICTOR SIGNALS likely acquire some reinforcing properties, applied use of PREDICTOR SIGNALS for an extended period without subsequent reward may not be effective. Overall, then, it seems likely that a SIGNAL, when conditioned to become a PREDICTOR SIGNAL, does function in a reinforcing capacity, until repeated presentations in the absence

Table 1

Findings from studies on PREDICTOR SIGNAL effectiveness as a secondary reinforcer in laboratory animals. All studies were summarized based on the following three questions: 1) ↑ rate of operant performance: does using a PREDICTOR SIGNAL in addition to the primary reinforcer result in more repetitions of the target behaviour compared to just using the primary reinforcer alone? 2) ↑ resistance to extinction: once primary reinforcers are no longer delivered for target behaviour performance, does continued use of a PREDICTOR SIGNAL sustain the target behaviour (i.e., decrease rate of "extinction" of the behaviour)? 3) Novel task performance (PS alone): can the PREDICTOR SIGNAL act alone to reinforce a new behaviour?

Species	Reference	↑ rate of operant performance	↑ resistance to extinction	Novel task performance (PS alone)
Cynomolgus monkey (*Macaca fascicularis*)	Gaffan and Harrison (1987)	–	–	√
Pigeon (*Columba livia*)	Zimmerman and Hanford (1966)	–	√	–
	Zimmerman et al. (1967)	–	√	–
Rat (*Rattus norvegicus*)	Bugelski (1938)	–	√	–
	Skinner (1938)	–	–	√
	Saltzman (1949)	–	–	√
	Zimmerman (1957)	–	√	–
	Zimmerman (1959)	–	√	√
	Egger and Miller (1962)	√	–	–
	Halford and Halford (1969)	√	–	–
	Davis and Smith (1976)	√	√	–
	Numan et al. (1976)	–	√	–
	Tombaugh et al. (1982)	–	√	–
	Slawecki et al. (1999)	√	√	–

(√ = supporting evidence; – = not assessed; no studies reported null results).

of the paired primary reinforcer diminish the predictive ability of the SIGNAL. This provides support for the Reinforcing Hypothesis, but leaves open the possibility that alternative mechanisms may also partially explain SIGNAL effectiveness.

3.2. The Marking Hypothesis

Animal trainers often refer to clickers as "event-markers" that "take a snapshot" of an animal's behaviour in the instant the click is heard (Pryor, 1999; Martin and Friedman, 2011). Lieberman et al. (1979) originally proposed what they termed the Marking Hypothesis, which suggests that salient cues (SIGNALS) presented in close temporal and contextual proximity to a goal behaviour act to emphasize this moment, which then facilitates learning. The Marking Hypothesis proposes that an arousing SIGNAL initiates an active memory search for preceding events that might causally relate to this cue (Lieberman et al., 1979). Lieberman et al. suggested that individuals (both human and animal) would be more likely to focus attention on distinctive events, and this extra attention may result in a stronger memory trace, leading to easier recall of the event. This marking function is distinguished from the reinforcing effect described above, because this mechanism does not require prior conditioning of the marker or SIGNAL with a primary reinforcer.

If the Marking Hypothesis is valid, SIGNALS should facilitate learning during delayed reinforcement even when they are neutral, having not yet been conditioned to predict the desired resource. However, assessment of naïve animals proves difficult because, as soon as the animal performs the target behaviour and receives the SIGNAL followed by the desired resource for the first time, the conditioning process has begun. To assess event-markers while controlling for the possibility that the marking stimulus could be reinforcing a particular choice, researchers have performed experiments in which they marked both correct and incorrect choices (Lieberman et al., 1979). For example, Thomas et al. (1983) implemented a T-maze training apparatus where animals had to learn to preferentially run down one arm of the T over the other to obtain a subsequent reward. In this study, the authors included a neutral but distinctive cue (e.g. a sound, handling, or a light) presented as a rat performed the choice-making behaviour (i.e., turning to run down one arm of the T-maze), whether the rats were correct or incorrect in their choice. They found that animals receiving the SIGNAL learned the task in fewer trials than those without the SIGNAL. These results suggest that a neutral stimulus presented at the time

of learning is able to facilitate task acquisition; however, in clicker training, the clicker SIGNAL is generally intentionally paired with a primary reinforcer, thus becoming a PREDICTOR SIGNAL (Pryor, 2005). One might assume that a PREDICTOR SIGNAL would mark a behaviour as effectively as a SIGNAL, but Thomas et al.'s paradigm did not actually assess whether or not PREDICTOR SIGNALS do function in a marking capacity. It is also possible that, because the neutral stimulus was paired with the primary reinforcer on the trials when the rat made a correct choice, it may have acquired some reinforcing properties. As such, the demonstrated marking effect could be irrelevant to clickers in their applied context.

To investigate this further, Williams (1991) trained rats on a two-choice paradigm and compared rates of learning in three groups of rats: 1) those who received a SIGNAL after both correct and incorrect responses, with correct responses only being subsequently reinforced with a food reward (in which case the SIGNAL, as described above, would not be entirely predictive of a food reward), 2) those who received a SIGNAL only after correct responses, which was then followed by a food reward (which would begin pairing the SIGNAL to the subsequent primary reinforcer to develop a PREDICTOR SIGNAL), and 3) those in a control group, who received no SIGNAL and only a food reward following correct responses. Williams found that rats who received the SIGNAL only after correct responses had significantly faster rates of task acquisition than rats in the other two conditions. This evidence suggests that, contrary to the Marking Hypothesis, a SIGNAL that reliably predicts a food reward is more effective than a non-predictive SIGNAL. In fact, the rats who received the non-predictive SIGNAL did not perform better than the no-SIGNAL control group.

From this, Williams concluded that the SIGNAL facilitated learning via its reinforcing properties, and that the Marking Hypothesis is not a sufficient alternative to the Reinforcing Hypothesis. However, this conclusion should be considered with caution. Williams' failure to demonstrate marking effects contradict the results reported by Lieberman et al. (1979) and Thomas et al. (1983). This could, indeed, be because the Marking Hypothesis does not sufficiently explain the mechanism behind SIGNAL function. However, it could equally be due to methodological differences such as salience of the SIGNAL that meant Williams was assessing a different phenomenon than reported by Lieberman et al. and Thomas et al. Furthermore, the SIGNAL used in Williams' study had not been previously paired with the primary reinforcer, suggesting that, at least at first, the SIGNAL itself did not have reinforcing capabilities and could not have been facilitating learning via secondary reinforcement. Additional

research should explore potential reasons for Williams' results and further assess the potential role of Marking in SIGNAL, and more specifically, PREDICTOR SIGNAL function.

3.3. The Bridging Hypothesis

The Bridging Hypothesis suggests that a stimulus presented between an animal performing a target behaviour and a subsequent reward could act to fill or bridge the gap between the behaviour and the consequence (Kaplan and Hearst, 1982). According to this hypothesis, as opposed to providing immediate reinforcement or marking a moment in time, the SIGNAL simply fills a temporal gap between the target behaviour and the delayed reinforcement. To study the putative bridging effect of a SIGNAL, researchers have used an autoshaping protocol (for example Rescorla, 1982). Autoshaping refers to the spontaneous acquisition of a food-seeking reflexive behaviour (like pecking behaviour in pigeons) at a specific target, when the target is repeatedly presented in close temporal association with a food reward (Brown and Jenkins, 1968). As with other forms of learning discussed previously, when autoshaping, the longer the temporal gap between the target presentation and the food reward, the slower the rate of task acquisition (Kaplan and Hearst, 1982). In the context of SIGNAL effects, Kaplan and Hearst (1982) found that pigeons acquired the spontaneous pecking behaviour more quickly when an auditory SIGNAL was presented during the gap between the target and food reward than when the gap was unfilled.

Based on these results, Kaplan and Hearst proposed the Bridging Hypothesis, where the SIGNAL (auditory stimulus) bridged the gap between the target and subsequent food reward. However, these results could also have been attributed to the Marking Hypothesis, as the gap-filling SIGNAL could have been marking the behaviour without bridging the gap. Kaplan and Hearst did not assess SIGNALS that bridged differing amounts of the gap between the behaviour and primary reinforcer. In addition, the auditory stimulus assessed in Kaplan and Hearst had no prior conditioning to be paired with the subsequent food reward. As such, the auditory stimulus was arguably not a PREDICTOR SIGNAL.

To parse out the bridging versus marking mechanisms of a previously conditioned PREDICTOR SIGNAL, Rescorla (1982) designed a special autoshaping protocol. In this case the pigeons were first conditioned to two different PREDICTOR SIGNALS (short and long auditory stimulus) predicting food. Both PREDICTOR SIGNALS were paired with the subsequent food reward an equal number of times, under the assumption that this would mean that the auditory stimuli were equally conditioned to predict the food. The longer-lasting auditory stimulus between the light and the food reward subsequently resulted in stronger light pecking behaviour in pigeons. This suggested that the auditory stimulus was likely to be bridging the gap between the light and food presentation, as opposed to only functioning as a reinforcer.

Rescorla argued that these results provided evidence that a conditioned PREDICTOR SIGNAL has some bridging effect, as the PREDICTOR SIGNAL that filled more of the temporal gap between the cue light and subsequent food reward was more effective in reinforcing the pigeons' light pecking behaviour. However, it is possible that, as per the Reinforcing Hypothesis, if the SIGNALS took on reinforcing qualities, a longer version of the SIGNAL may have had greater reinforcing power than a shorter one. This has not been specifically addressed thus far, as the superior effect of the longer tone has been primarily attributed to the bridging effect (Rescorla, 1982).

To the best of our knowledge, the Bridging Hypothesis of PREDICTOR SIGNAL function has only been assessed in the context of the autoshaping protocol presented above. Since autoshaping protocols rely solely on the spontaneous generation of appetitive behaviours, the extent to which the Bridging Hypothesis applies to PREDICTOR SIGNALS used in more common training paradigms needs to be rigorously studied.

3.4. Summary

Based on the above evidence, it appears that SIGNALS like clickers most likely function in a reinforcing capacity, provided they are first paired with a primary reinforcer to the extent required to imbue them with reinforcing capabilities, and provided this 'charge' is maintained so that the reinforcing properties are not extinguished. Less certain is whether they also have bridging and marking capabilities. However, assessing these hypotheses using behavioural evidence requires specially designed and tightly controlled paradigms which differ markedly from the intended applied context of delayed reinforcement learning. As such, the extent to which the conclusions from laboratory based studies still hold when the PREDICTOR SIGNAL is used as a means of facilitating delayed reinforcement learning in applied settings remains an open question. Instead, we can begin assessing PREDICTOR SIGNALS in their intended contexts by not only examining behavioural evidence but neurological activity as well. Technological advances in studying brain activity related to rewards and learning could provide an alternative source of insight into how PREDICTOR SIGNALS function.

4. Neurochemical approach

Historically, the only feasible way to study mechanisms of learning was to look at behavioural outcomes. Continued technological advancement, however, has provided another window into developing our understanding of learning and behaviour. The neurotransmitter dopamine is thought to play a role in the experience of rewards and motivation (Domjan, 2003). As such, dopamine has been a focus of researchers studying such phenomena for many years (e.g. Berridge, 2012). Historically, there were two main competing hypotheses with regards to dopamine's role in rewards and motivation: the Incentive Salience Hypothesis, (proposed by Robinson and Berridge, 1993) and the Reward Prediction Error Hypothesis (RPEH) (proposed by Schultz, 1998). More recent evidence suggests that, while dopamine is involved in determining the salience of incentives, it does not play a role in the hedonic experience of pleasure, which is the basis for the Incentive Salience Hypothesis (Berridge and Kringelbach, 2013). As such, for the purposes of this review, we will focus on Schultz's RPEH. This is presently the predominant hypothesis regarding mechanisms underlying dopamine's function in reward prediction (as reviewed by Colombo, 2014).

Based on the RPEH, changes in levels of dopamine occur when there is a mismatch between expected and received rewards. Hollerman and Schultz (1998) found that, when a previously neutral signal is paired with a desirable resource, such that the signal functions to predict presentation of the desirable resource (i.e., becomes a PREDICTOR SIGNAL), dopamine release occurs at the time of PREDICTOR SIGNAL presentation. Furthermore, no additional rise in levels of dopamine occurs when the animal receives the subsequent desirable resource, most likely since it matched the animal's expectations. This evidence supports the idea that the PREDICTOR SIGNAL acts as a secondary reinforcer, since the PREDICTOR SIGNAL takes on the role of eliciting the pleasurable effects, at the neurological level, of the original desirable resource (primary reinforcer). However, when rats are presented with a conditioned PREDICTOR SIGNAL but not given the expected desirable resource, thus producing a reward prediction error, the rats experience a rise in dopamine when the PREDICTOR SIGNAL is presented but, at the

time the expected primary reinforcer is withheld, dopaminergic activity is actually suppressed to below baseline levels (Schultz, 2016). This expectancy violation may precipitate the rapid decline of performance observed when the primary reinforcer is no longer provided after the SIGNAL.

As previously discussed, one argument based on the Reinforcing Hypothesis is that the PREDICTOR SIGNAL should take on the reinforcing quality of the primary reinforcer with which it was paired. Here, when the primary reinforcer (desirable resource) is predicted but withheld, a drop in dopamine is experienced. This contradicts the suggestion that the PREDICTOR SIGNAL provides the same reinforcing quality as the primary reinforcer, as the dopamine drop following the lack of primary reinforcer is not present following presentation of a primary reinforcer itself. The RPEH also suggests that PREDICTOR SIGNALS and reward predictions are continuously being updated with every experience the animal has with the signal. Thus, repeated presentations of the PREDICTOR SIGNAL without being followed by a primary reinforcer would result in an updated prediction and reduced activation of dopaminergic neurons. This is consistent with behavioural evidence that conditioned PREDICTOR SIGNALS are only able to reinforce novel behaviours for a short amount of time without being followed by the desirable resource before it no longer facilitates learning (as found by McCall and Burgin, 2002).

No studies have specifically discussed dopaminergic response patterns in the context of the Marking or Bridging Hypotheses. As previously discussed, the marking mechanism of a neutral SIGNAL is suggested to be caused by an instant of arousal following the SIGNAL. In fact, research suggests that the dopamine response has two components (as reieved by Stauffer et al., 2015). In a dopaminergic response to a stimulus, there is initially a non-selective increase in dopamine, followed by either a continued increase (if the stimulus predicts a reward) or a drop in dopaminergic activity (if the stimulus does not predict a reward). The non-selective initial response is thought to prime the brain's sensitivity to subsequent processing of potential rewards. Perhaps this dopaminergic activity contributes to the "marking" mechanism of the SIGNAL.

In terms of the Bridging Hypothesis, differences in dopaminergic responses to PREDICTOR SIGNALS that either fully or partially fill the temporal gap between behaviour and subsequent reward have not yet been investigated. However, based on the hypothesis, if a longer-filling PREDICTOR SIGNAL produces a greater dopaminergic response than a shorter-filling one, there would be neurochemical evidence for PREDICTOR SIGNAL function under the Bridging Hypothesis. Perhaps the greater salience, or some other factor of the longer (versus shorter) PREDICTOR SIGNAL produces a greater dopamine response, thus providing a greater reward. This would be worth further investigation in order to better understand factors involved in choosing an optimal PREDICTOR SIGNAL.

Overall, further research into the role of dopamine and other neurological activity related to rewards and learning would help to better interpret the results from behavioural tests. They could also improve understanding of how PREDICTOR SIGNALS attenuate learning deficits brought on by delayed reinforcement.

5. What next?

Based on a critical review of the current literature, it appears that the function of SIGNALs such as clickers, in laboratory animals at least, fits well with the Reinforcement Hypothesis, but bridging and marking mechanisms could potentially work in conjunction with a reinforcing mechanism to facilitate learning. While bridging or marking explanations are likely not competitive *alternatives* to the secondary reinforcement explanation for the effects of a clicker, they are potentially *complementary* mechanisms in explaining how animals learn. This is important because, if clicker-type SIGNALS operate through multiple mechanisms under laboratory conditions, not all of these may apply under the less-controlled conditions that operate in most companion animal training contexts. Inconsistent results in the limited number of studies in companion animals suggest that further investigation is required to better understand the contexts in which SIGNALS could facilitate learning.

In particular, the evidence suggests that, for a SIGNAL to function under the Reinforcing Hypothesis, it must be paired with a primary reinforcer until it becomes a PREDICTOR SIGNAL. The number of trials this takes and the conditions affecting this pairing could potentially be assessed by recording dopamine response in addition to the behaviour following successful SIGNAL-primary reinforcer conditioning. The literature also suggests that while the PREDICTOR SIGNAL can be used alone to reinforce novel behaviours and maintain behaviours in the face of withheld primary reinforcement, the SIGNAL can only do this while it retains its predictive, and subsequently, reinforcing, value. Further investigation into ways of maintaining SIGNALS' reinforcing abilities, looking at both behavioural and neurochemical evidence, could better provide animal trainers with methods of applying clicker-type SIGNALS without continuous primary reinforcement. Additional research is required to investigate the implications of various factors, such as the number of SIGNAL-primary reinforcer pairings required prior to using the SIGNAL in training novel tasks, variable ratios of SIGNAL-primary reinforcer pairing, and quality and intensity of the SIGNAL. For example, perhaps the short, sharp sound emitted by a clicker would be the most effective at marking, but a tone or whistle that can be temporally extended could have a greater bridging effect. A better understanding of if and how any of these or other factors mediate the ability of SIGNALS to facilitate learning in the face of delayed or even withheld primary reinforcement could have implications in applied settings.

Finally, it would be inappropriate to discard other potential mechanisms underlying clicker-type SIGNAL function. In the 1960 book *Plans and the Structure of Behavior*, Miller et al. (1960) proposed the Test-Operate-Test-Exit (TOTE) cognitive feedback mechanism to explain how humans plan and make decisions. In the TOTE unit, an individual begins with a particular goal in mind, and tests whether or not this goal has been met. If it has, the individual exits the feedback loop. If not, the individual operates on, or manipulates the environment in an effort to meet the goal and the goal is re-tested. The sequence of Test-Operate-Test repeats until the goal condition has been met, at which point the individual exits the TOTE unit (Miller et al., 1960). Such a recursive mental model requires higher order cognitive processes (Kopp, 2012), but it is certainly possible that at least some animals have this cognitive capacity (Smith et al., 2012).

In the context of SIGNAL processing, perhaps the SIGNAL acts as a cue to exit the TOTE unit. In other words, the animal understands that the SIGNAL means the animal has completed the goal. If so, the SIGNAL might be a precise method of communication from the trainer to the animal the instant the correct behaviour was performed. In fact, when teaching companion animal owners about clicker training, animal trainers often explain that the clicker communicates to the animal, "that behaviour was what I wanted" (e.g. Pryor, 2006). Whether or not this is in fact how the animal perceives the SIGNAL has not been assessed. These and other proposed cognitive mechanisms underlying clicker-type SIGNAL function are potentially insightful avenues of further investigation. Perhaps future applications of fMRI technology would allow a clearer window into the cognitive processes involved in perception of clicker-type SIGNALS.

Conflict of interest

The authors declare that there are no conflict of interest.

Acknowledgments

References

Bailey, R., Bailey, M., 1996a. Patient like the Chipmunks. Eclectic Science Productions, Hot Springs, AR.

Bailey, R.E., Bailey, M.B., 1996b. Patient like the Chipmunks. Eclectic Science Productions, Hot Springs, AR.

Berridge, K.C., Kringelbach, M.L., 2013. Neuroscience of affect: brain mechanisms of pleasure and displeasure. Curr. Opin. Neurobiol. 23, 294–303, http://dx.doi.org/10.1016/j.conb.2013.01.017.

Berridge, K.C., 2012. From prediction error to incentive salience: mesolimbic computation of reward motivation. Eur. J. Neurosci. 35, 1124–1143, http://dx.doi.org/10.1111/j.1460-9568.2012.07990.x.

Blandina, A.G., n.d. To click or not to click: Positive reinforcement methods on the acquisition of behavior, Unpublished honours thesis, University of Florida.

Brown, P.L., Jenkins, H.M., 1968. Auto-shaping of the pigeon's key-peck. J. Exp. Anal. Behav. 11, 1–8, http://dx.doi.org/10.1901/jeab.1968.11-1.

Bugelski, R., 1938. Extinction with and without sub-goal reinforcement. J. Comp. Psychol. 26, 121–134, http://dx.doi.org/10.1037/h0057091.

Colombo, M., 2014. Deep and beautiful. The reward prediction error hypothesis of dopamine. Stud. Hist. Philos. Biol. Biomed. Sci. 45, 57–67, http://dx.doi.org/10.1016/j.shpsc.2013.10.006.

Davis, W.M., Smith, S.G., 1976. Role of conditioned reinforcers in the initiation, maintenance and extinction of drug-seeking behavior. Pavlovian J. Biol. Sci. 11, 222, http://dx.doi.org/10.1007/bf03000316.

Domjan, M., 2003. The Principles of Learning and Behavior, 5th ed. Thomson Brooks/Cole Publishing Co, Belmont, CA, US.

Egger, M.D., Miller, N.E., 1962. Secondary reinforcement in rats as a function of information value and reliability of the stimulus. J. Exp. Psychol. 64, 97–104, http://dx.doi.org/10.1037/h0040364.

Gaffan, D., Harrison, S., 1987. Amygdalectomy and disconnection in visual learning for auditory secondary reinforcement by monkeys. J. Neurosci. 7, 2285–2292.

Gillaspy, J.A., Brinegar, J.L., Bailey, R.E., 2014. Operant psychology makes a splash—in marine mammal training (1955–1965). J. Hist. Behav. Sci. 50, 231–248, http://dx.doi.org/10.1002/jhbs.21664.

Grice, G.R., 1948. The relation of secondary reinforcement to delayed reward in visual discrimination learning. J. Exp. Psychol. 38, 1–16, http://dx.doi.org/10.1037/h0061016.

Halford, G.S., Halford, J.M., 1969. Secondary reinforcement: signal or substitute reward? A preliminary investigation. Aust. J. Psychol. 21, 145, http://dx.doi.org/10.1080/00049536908257777.

Hollerman, J.R., Schultz, W., 1998. Dopamine neurons report an error in the temporal prediction of reward during learning. Nat. Neurosci. 1, 304–309, http://dx.doi.org/10.1038/1124.

Kaplan, P.S., Hearst, E., 1982. Bridging temporal gaps between CS and US in autoshaping: insertion of other stimuli before, during, and after CS. J. Exp. Psychol. Anim. B 8, 187–203, http://dx.doi.org/10.1037/0097-7403.8.2.187.

Kopp, B., 2012. A simple hypothesis of executive function. Front. Hum. Neurosci. 6, http://dx.doi.org/10.3389/fnhum.2012.00159.

Langbein, J., Siebert, K., Nuernberg, G., Manteuffel, G., 2007. The impact of acoustical secondary reinforcement during shape discrimination learning of dwarf goats (Capra hircus). Appl. Anim. Behav. Sci. 103, 35–44, http://dx.doi.org/10.1016/j.applanim.2006.04.019.

Lattal, K.A., 2010. Delayed reinforcement of operant behavior. J. Exp. Anal. Behav. 93, 129–139, http://dx.doi.org/10.1901/jeab.2010.93-129.

Lieberman, D.A., McIntosh, D.C., Thomas, G.V., 1979. Learning when reward is delayed: a marking hypothesis. J. Exp. Psychol. Anim. B 5, 224–242, http://dx.doi.org/10.1037/0097-7403.5.3.224.

Martin, S., Friedman, S.G., 2011. Blazing clickers. Conference Paper at Animal Behavior Management Alliance.

McCall, C.A., Burgin, S.E., 2002. Equine utilization of secondary reinforcement during response extinction and acquisition. Appl. Anim. Behav. Sci. 78, 253–262, http://dx.doi.org/10.1016/S0168-1591(02)00109-0.

McCulloch, M., Jezierski, T., Broffman, M., Hubbard, A., Turner, K., Janecki, T., 2006. Diagnostic accuracy of canine scent detection in early- and late-stage lung and breast cancers. Integr. Cancer Ther. 5, 30–39, http://dx.doi.org/10.1177/1534735405285096.

Miller, G.A., Galanter, E., Pribram, K.H., 1960. Plans and the Structure of Behavior. Holt, Rinchart and Winston, Inc.

Numan, R., Banerjee, U., Smith, N., Lal, H., 1976. Secondary reinforcement property of a stimulus paired with morphine administration in the rat. Pharmacol. Biochem. Behav. 5, 395–399, http://dx.doi.org/10.1016/0091-3057(76)90102-7.

Pryor, K., 1999. Don't Shoot the Dog! The New Art of Teaching and Training, 2nd ed. Bantam Books, New York, New York.

Pryor, K., 2005. Getting Started: Clicker Training for Dogs. Sunshine Books.

Pryor, K., 2006. What Is Clicker Training?, Karen Pryor Clicker Training, (Web page) Accessed August 20, 2015, Retrieved from http://www.clickertraining.com/what_is_clicker_training.

Pryor, K., 2009. Reaching the Animal Mind: Clicker Training and what it Teaches us About all Animals. Scribner, New York, New York.

Ramirez, K., 1999. Animal Training: Successful Animal Management Through Positive Reinforcement. Shedd Aquarium Society, Chicago, Illinois.

Rescorla, R.A., 1982. Effect of a stimulus intervening between CS and US in autoshaping. J. Exp. Psychol. Anim. B 8, 131–141, http://dx.doi.org/10.1037/0097-7403.8.2.131.

Robinson, T.E., Berridge, K.C., 1993. The neural basis of drug craving: an incentive-sensitization theory of addiction. Brain Res. Rev. 18, 247–291, http://dx.doi.org/10.1016/0165-0173(93)90013-P.

Saltzman, I.J., 1949. Maze learning in the absence of primary reinforcement: a study of secondary reinforcement. J. Comp. Physiol. Psychol. 42, 161–173, http://dx.doi.org/10.1037/h0056337.

Schultz, W., 1998. Predictive reward signal of dopamine neurons. J. Neurophysiol. 80, 1–27.

Schultz, W., 2016. Dopamine reward prediction-error signalling: a two-component response. Nat. Rev. Neurosci., http://dx.doi.org/10.1038/nrn.2015.26.

Skinner, B., 1933. The abolishment of a discrimination. Proc. Natl. Acad. Sci. U. S. A. 19, 825–828, http://dx.doi.org/10.1073/pnas.19.9.825.

Skinner, B.F., 1938. The Behavior of Organisms: An Experimental Analysis. Appleton-Century, Oxford, England.

Slawecki, C.J., Samson, H.H., Chappell, A., 1999. Presentation of an ethanol-paired stimulus complex alters response patterns during extinction. Pharmacol. Biochem. Behav. 62 (1), 127–135, http://dx.doi.org/10.1016/S0091-3057(98)00134-8.

Smith, S.M., Davis, E.S., 2008. Clicker increases resistance to extinction but does not decrease training time of a simple operant task in domestic dogs (Canis familiaris). Appl. Anim. Behav. Sci. 110, 318–329, http://dx.doi.org/10.1016/j.applanim.2007.04.012.

Smith, J.D., Couchman, J.J., Beran, M.J., 2012. The highs and lows of theoretical interpretation in animal-metacognition research. Philos. Trans. R. Soc. Lond. B Biol. Sci. 367, 1297–1309, http://dx.doi.org/10.1098/rstb.2011.0366.

Spence, K.W., 1947. The role of secondary reinforcement in delayed reward learning. Psychol. Rev. 54, 1–8, http://dx.doi.org/10.1037/h0056533.

Stauffer, W.R., Lak, A., Kobayashi, S., Schultz, W., 2015. Components and characteristics of the dopamine reward utility signal. J. Comp. Neurol. 524, 1699–1711, http://dx.doi.org/10.1002/cne.23880.

Thomas, G.V., Lieberman, D.A., McIntosh, D.C., Ronaldson, P., 1983. The role of marking when reward is delayed. J. Exp. Psychol. Anim. B 9, 401–411, http://dx.doi.org/10.1037/0097-7403.9.4.401.

Tombaugh, T.N., Grandmaison, L.J., Zito, K.A., 1982. Establishment of secondary reinforcement in sign tracking and place preference tests following pimozide treatment. Pharmacol. Biochem. Behav. 17, 665–670, http://dx.doi.org/10.1016/0091-3057(82)90342-2.

Williams, J.L., Friend, T.H., Nevill, C.H., Archer, G., 2004. The efficacy of a secondary reinforcer (clicker) during acquisition and extinction of an operant task in horses. Appl. Anim. Behav. Sci. 88, 331–341, http://dx.doi.org/10.1016/j.applanim.2004.03.008.

Williams, B.A., 1991. Marking and bridging versus conditioned reinforcement. Anim. Learn. Behav. 19, 264–269, http://dx.doi.org/10.3758/bf03197885.

Williams, B.A., 1994. Conditioned reinforcement: experimental and theoretical issues. Behav. Analyst 17, 261–285.

Willis, C., Church, S., Guest, C., 2004. Olfactory detection of human bladder cancer by dogs: proof of principle study. BMJ 329, 712–714, http://dx.doi.org/10.1136/bmj.329.7468.712.

Zimmerman, J., Hanford, P.V., 1966. Sustaining behavior with conditioned reinforcement as the only response-produced consequence. Psychol. Rep. 19, 391–401, http://dx.doi.org/10.2466/pr0.1966.19.2.391.

Zimmerman, D.W., 1957. Durable secondary reinforcement: method and theory. Psychol. Rev. 64, 373–383, http://dx.doi.org/10.1037/h0041885.

Zimmerman, D.W., 1959. Sustained performance in rats based on secondary reinforcement. J. Comp. Physiol. Psychol. 52, 353–358, http://dx.doi.org/10.1037/h0045807.

Zimmerman, J., Hanford, P.V., Brown, W., 1967. Effects of conditioned reinforcement frequency in an intermittent free-feeding situation. J. Exp. Anal. Behav. 10, 331–340, http://dx.doi.org/10.1901/jeab.1967.10-331.