Neuroscience and Biobehavioral Reviews 65 (2016) 1–20

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Review article

Motivational control of sign-tracking behaviour: A theoretical framework

Patrick Anselme [a,b,*]

[a] Department of Psychology, University of Liege, Belgium
[b] Department of Biopsychology, University of Bochum, Germany

ARTICLE INFO

Article history:
Received 5 December 2015
Received in revised form 25 February 2016
Accepted 21 March 2016
Available online 22 March 2016

Keywords:
Incentive motivation
Sign-tracking
Learning
Reward uncertainty
Dopamine

ABSTRACT

Learning and motivation are two psychological processes allowing animals to form and express Pavlovian associations between a conditioned stimulus (CS) and an unconditioned stimulus (UCS). However, most models have attempted to capture the mechanisms of learning while neglecting the role that motivation (or incentive salience) may *actively* play in the expression of behaviour. There is now a body of neurobehavioural evidence showing that incentive salience represents a major determinant of Pavlovian performance. This article presents a motivational model of sign-tracking behaviour whose aim is to explain a wide range of behavioural effects, including those related to partial reinforcement, physiological changes, competition between CSs, and individual differences in responding to a CS. In this model, associative learning is assumed to determine the ability to produce a Pavlovian conditioned response rather than to control the strength and the quality of that response. The model is in keeping with the incentive salience hypothesis and will therefore be discussed in the context of dopamine's role in the brain.

Contents

1. Introduction 2
2. How associative learning relates to performance 2
3. Partial reinforcement and the so-called frustration effect 4
4. Incentive salience and the neurobiology of sign- and goal-trackers 5
5. Computer modelling of incentive salience 6
6. Incentive hope: concept and model 7
   6.1. First mechanism: counterconditioning 7
   6.2. Second mechanism: Incentive hope 7
   6.3. Incentive hope, frustration, and dopamine 9
   6.4. Development of sign-tracking performance 9
7. Predictive scope of the model 11
   7.1. The effects of reward uncertainty 11
   7.2. Instant shifts in sign-tracking performance 12
   7.3. Inter-individual differences 12
   7.4. Stimuli competition 13
   7.5. Some original predictions 16
   7.6. Prediction 1 16
   7.7. Prediction 2 16
   7.8. Prediction 3 16
   7.9. Prediction 4 17

* Correspondence to: Department of Biopsychology, University of Bochum, 150 Universitätsstraße, D-44801 Bochum, Germany.
E-mail address: Patrick.Anselme@rub.de

http://dx.doi.org/10.1016/j.neubiorev.2016.03.014
0149-7634/

8. Conclusion 17
Acknowledgments 17
References 17

1. Introduction

Pavlovian conditioning is the process that enables a neutral cue (such as a metal lever) to become a conditioned stimulus (CS) through repeated association with an unconditioned stimulus (UCS). The transformation of a neutral cue into a CS is traditionally believed to be the consequence of associative learning, which makes the CS-UCS association stronger and stronger on each trial (Mackintosh, 1975; Pearce and Hall, 1980; Rescorla and Wagner, 1972). Non-associationist theories, which focus on the role of temporal intervals between trials and time accumulation in conditioning, have also emerged (Gallistel and Gibbon, 2000; Gibbon, 1977; Gibbon and Balsam, 1981). In both cases, the CS can be said to acquire its significance from the information it carries about the UCS (Gallistel, 2003; Rescorla, 1988): the CS predicts the UCS more than any other cue present during training.

However, is this account the whole story? Despite many successes, associationist and non-associationist theories of Pavlovian *learning* fail to correctly predict significant facts when they are used to account for Pavlovian *performance* (Anselme, 2015; Berridge, 2012; Lesaint et al., 2014; Meyer et al., 2014; Rescorla, 1988; Zhang et al., 2009). Firstly, they are incompatible with the evidence that partial reinforcement (PRF) enhances conditioned responding relative to continuous reinforcement (CRF). Although conditioned responding tends to develop more slowly early in training under PRF than under CRF, performance often reaches a higher asymptotic level under PRF than under CRF (e.g., Anselme et al., 2013; Boakes, 1977; Collins and Pearce, 1985; Gibbon et al., 1980; Gottlieb, 2004; Robinson et al., 2014b). Secondly, they are incompatible with the evidence that altering an animal's physiological state causes instant shift in conditioned responding despite an absence of additional training. Relearning the task in the new motivational state is unnecessary to impact behaviour (Robinson and Berridge, 2013; Tindell et al., 2005). Thirdly, those theories are limited to reproducing the averaged behaviour of a population of individuals, irrespective of the great variability that may exist among individuals in the expression of that behaviour. Because inter-individual differences are not considered, traditional interpretations cannot explain why, in autoshaping, some individuals become 'sign-trackers' (preferring to approach and to interact with the CS), while other individuals become 'goal-trackers' (preferring to approach and to interact with the food dish during CS presentations) or ambivalent responders (Flagel et al., 2007; Robinson and Flagel, 2009). In other words, current theories of Pavlovian learning do not reliably reflect what animals actually do.

In the absence of learning, a CS remains a meaningless (neutral) stimulus because no information about the UCS is carried (Rescorla, 1968). But the motivational (non-associative) component – eluded by most traditional models – of Pavlovian conditioning might exert *more direct* control on behaviour than does the learned association (see Berridge, 2012). The incentive salience hypothesis posits that animals, including humans, come to act according to what they 'want' as a result of increased dopamine release in their nucleus accumbens (Berridge and Robinson, 1998). In this view, reward 'wanting' plays a determining role in controlling behaviour in autoshaping. Thus, proposing a model that shows how motivational 'wanting' can be gradually attributed to a CS in order to reflect real behaviour has become a major preoccupation in behavioural neuroscience (Anselme, 2015; Berridge, 2012; Dayan and Berridge 2014; Meyer et al., 2014; Zhang et al., 2009). Sign-tracking responses in autoshaping are a traditional way of indexing incentive motivation (Anselme et al., 2013; Doremus-Fitzwater and Spear, 2011; Flagel et al., 2007), so that the motivational control of Pavlovian performance will be essentially (but not exclusively) discussed in a sign-tracking context throughout this paper. Briefly, autoshaping is an experimental procedure in which a metal lever (with rats) or an illuminated key (with pigeons) acting as a CS is presented for a few seconds to animals and, immediately after its presentation, some food becomes available in a food dish. The animal is not required to perform any action to be rewarded, so that CS-directed behaviours are assumed to reflect the animal's motivation for the food pellets. It is widely accepted that dopamine in the nucleus accumbens is necessary for the expression of sign-tracking responses (Blaiss and Janak, 2009; Day et al., 2006; Flagel et al., 2007; Meyer et al., 2012; Saunders and Robinson, 2012).

The model presented here – then referred to as the extended incentive hope model – aims to provide a unified solution to important problems, including the effects of reward uncertainty, physiological changes, individual differences, and stimuli competition. We will see that the model is both integrative and predictive, as it enables the formulation of original predictions. Elsewhere, the incentive hope model was shown to account for the effects of PRF on behaviour, although it was technically applicable to a wider range of phenomena involving no uncertainty (Anselme, 2015). Those phenomena, among others, are considered in the present article. The model's mathematical formulation has been slightly modified (two parameters, α and γ, have been added), enabling it to explain and predict additional facts (stimuli competition and individual differences) without altering its initial predictive power. Briefly, the extended incentive hope model describes how an animal's motivation ('wanting') towards a UCS is progressively attributed to a CS without explicit reference to any learning algorithm. It is argued that associative learning determines the ability to express a conditioned response, but does not modulate the properties (strength, duration, quality) of that response. The model shows how the incentive salience hypothesis might be used to account for the effects of reward uncertainty, as well as many other Pavlovian effects (related to sign-tracking) that are traditionally believed to result from associative learning only.

2. How associative learning relates to performance

Early models of Pavlovian conditioning postulated that the acquisition of conditioned responding gradually emerges over trials by means of an error-correcting mechanism (Mackintosh, 1975; Pearce and Hall, 1980; Rescorla and Wagner, 1972). These models have been supported by a large number of data, especially in situations involving stimuli competition (e.g., Kamin, 1969; Pavlov, 1927; Rescorla, 1971). Criticizing the principle of absolute temporal contiguity, several authors broke with the associationist tradition and have provided a new approach to Pavlovian conditioning, where time accumulation is more important than the number of trials to account for conditioning (Gallistel and Gibbon, 2000; Gibbon, 1977; Gibbon and Balsam, 1981). Let's briefly show the predictions that the main associationist and non-associationist models of Pavlovian conditioning make with respect to reward uncertainty.

- *The Rescorla-Wagner model* (Rescorla and Wagner, 1972). On any given trial, the associative strength (V) between a CS and a UCS is compared to the maximum conditioning possible (λ) with that

UCS and the difference ($\lambda - V$) is treated like an error to be corrected. Learning is maximal when $V = \lambda$, which represents the asymptote. The physical salience of the CS (α) and of the UCS (β) determines the slope of the learning curve. The statement that the V value reflects the strength of a conditioned response is incompatible with the excitatory effects of PRF (Gottlieb, 2004). Indeed, the model predicts that the error in learning will be higher under PRF ($\lambda - V > 0$) than under CRF ($\lambda - V = 0$) at the end of training, and hence that PRF should generate a lower performance than CRF.

- *The Pearce-Hall model* (Pearce and Hall, 1980; see also Pearce et al., 1982). The amount of attention paid to a CS is inversely related to its predictive accuracy, so that uncertain rewards attract attention more than certain rewards. Interestingly here, the error correction mechanism ($\lambda - V$) denotes the changes in CS effectiveness, not the changes in UCS effectiveness. Hence, sign-tracking is not only controlled by V (lower under uncertainty), but also by an orienting attentional response, noted α (higher under uncertainty). A problem with this view is that it is not known how V and α interact to determine sign-tracking performance (Collins and Pearce, 1985). Although we could speculate about how they might combine, a more fundamental problem occurs here: there is nothing in the model explaining how (and why) attention itself should impact behaviour. Attention is a very general process that may be recruited both for appetitive and aversive stimuli, so it is not specific to approach behaviour. As a result, attentional focus is necessary, but perhaps not sufficient, to approach and interact with the CS.
- *The Gallistel-Gibbon model* (Gallistel and Gibbon, 2000). The rate estimation theory (RET) predicts that the speed of acquisition depends on the ratio of the length of the ITI (I) to the duration of the CS (T). A conditioned response emerges when animals acquire the certainty that the absence of the CS predicts nonreinforcement – i.e., when the ratio of the cumulated ITI time to the cumulated CS time (which is, on average, the I/T ratio) exceeds a decision threshold. The RET aims to predict whether or not an animal will respond to a CS; it makes no prediction about the properties of conditioned responding (Bouton and Sunsay, 2003; Gottlieb, 2004; Kakade and Dayan, 2002). Thus, PRF and CRF are assumed to have similar effects on performance when measured as a function of the number of rewards received because reward probability does not alter the I/T ratio. Although this prediction may be correct, no explanation is provided with respect to the excitatory property of PRF measured as a function of the number of trials/sessions.

It is also important to note that it is out of the scope of these models to predict anything about the effects of motivational changes and of individual differences because they do not include physiological factors and are insensitive to behavioural variability. Several decades ago, Rescorla (1988) already pointed out the limits of learning models to explain Pavlovian performance. According to him, it is mistaken to believe that Pavlovian conditioning consists of the shifting of a response from one stimulus (the UCS) to another (the CS), because the unconditioned and conditioned responses may be different. A shock UCS may increase activity, while the associated CS decreases it. Also, the conditioned response often depends on the properties of the CS. In pigeons, an illuminated key induces direct pecking, while a diffuse tone enhances general activity. Thus, Rescorla (1988) suggested that the CS does not acquire the UCS's properties but simply acquires the *ability* to evoke a response. This statement means that associative models of Pavlovian conditioning do not provide an adequate explanation of performance, as explicitly recognised by Rescorla in the same article. I think that the idea that associative learning determines the ability to evoke a conditioned response is not incompatible with the principle of error correction ($\lambda - V$). It just means that associative strength must overcome a fixed threshold (θ) to allow the expression of the response. Regarding the Rescorla-Wagner model, $V_n = \alpha\beta(\lambda - V_{n-1})$, we could write:

if $V_n \leq \theta$, no conditioned response is expressed;

if $V_n > \theta$, a conditioned response is expressed

In other words, the *gradual* acquisition of a CS-UCS association by means of an error-correction mechanism can become an *all-or-none* process relative to a decision threshold. Alternatively, the I/T ratio and its decision threshold (Gallistel and Gibbon, 2000) might also be a good estimate of learning as a determinant of the ability to respond. It is not here the place to discuss these two options or to select one of them. For the model presented further, the exact learning algorithm underlying Pavlovian performance is unimportant. The model simply presupposes that, if a conditioned response emerges, it is because the learned significance of a CS relative to an UCS overcame a fixed decision threshold (Fig. 1).

While agreeing with Rescorla's general view, I would like to suggest that the differences in responding to distinct CSs may simply result from the animal's opportunity to interact with those CSs. When some interaction with a CS is possible (as in the case of a

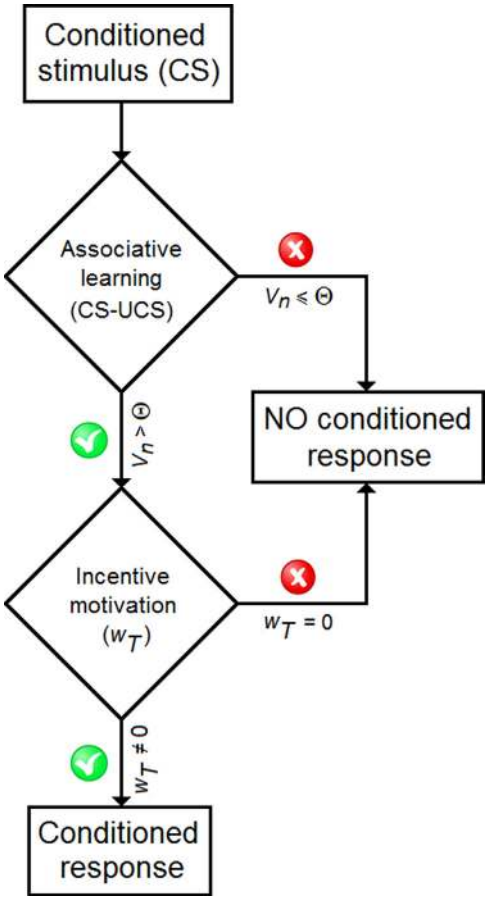

Fig 1. How associative learning and incentive motivation are assumed to control conditioned responding. If associative strength between a CS and a UCS is not sufficient (e.g., $V_n \leq \theta$), the learning process does not have the ability to produce a conditioned response. If associative strength is sufficient (e.g., $V_n > \theta$), the learning process has the ability to produce a conditioned response. But learning does not directly do it. For that, as seen further, the individual's motivation (w_T) for the UCS must differ from zero (approaching the CS involves $w_T > 0$, avoiding the CS involves $w_T < 0$). If $w_T = 0$, no conditioned response is elicited. In other words, incentive motivation rather than associative learning is responsible for the expression of a conditioned response, even though the ability to respond only depends on learning.

metal lever or of an illuminated key), the conditioned response often reflects the animal's motivation toward the UCS: rats trained with a food UCS tend to nibble the lever, while rats trained with a liquid UCS tend to lick the lever (Anselme et al., 2013; Davey et al., 1981). In contrast, when no interaction with a CS is possible (as in the case of a tone, a light, or a click), the animal is unable to adequately express its motivation toward the UCS. Thus, the conditioned response may result in general activity (locomotion), in a displacement activity (such as grooming) or in goal-tracking (food dish inspection). For example, in autoshaping, Meyer et al. (2014) found that a tone or a light leads rats to become goal-trackers – no sign-tracking activity develops with such CSs – even when they have a neurobiological profile of sign-trackers (for more details, see section 4). In other words, the CS reflects the UCS's motivational attractiveness, and the type of CS determines how that motivation is expressed in behaviour.

Associative models have had a renewal of interest in recent years, as many studies have suggested that the phasic activity of midbrain dopamine neurons consists of a prediction error signal similar to that presented in those models (Schultz, 1998). In other words, dopamine would be the neuromediator of reward-driven learning. In favour of this interpretation are notably the findings that, in monkeys, a strong phasic dopamine release is shown immediately after an unexpected reward (probability $p = 0.0$) and gradually decreases as the probability with which a CS predicts reward delivery increases. No release (beyond the background activity of dopamine neurons) is observed at $p = 1.0$. An incompletely learned CS-UCS association and an inaccurate predictability of reward delivery ($0 < p < 1$) would release phasic dopamine as a signal for the brain to learn more about the CS-UCS reward association, mimicking the expression $\lambda - V > 0$ highlighted by associative models (Mirenowicz and Schultz, 1994). Accordingly, dopamine continues to be produced following extensive training when the time interval between the CS and the UCS reward is randomly variable – i.e., impossible to learn (Schultz et al., 1993). We will see that this correlational interpretation of dopamine's role is challenged by the evidence that dopamine is a major cause of the attribution of a motivational salience to a CS (sections 4 and 5).

3. Partial reinforcement and the so-called frustration effect

Given the failure of associative learning models to account for the excitatory effects of PRF on sign-tracking performance, several psychological mechanisms have been proposed to add their effects to those of learning (Amsel, 1958; frustration; Capaldi, 1967; memory traces; Pearce and Hall, 1980; attention). Although such mechanisms can occasionally play a role in conditioning, secondary-response theories – which postulate that responding results from learning *plus* another mechanism – do not clearly show how the added processes can increase performance under PRF, whereas associative strength is weakened (Anselme, 2015). The added processes could be insufficient to enhance performance beyond the asymptotic level reached under CRF, as noted by Collins and Pearce (1985) with respect to their attentional model. Secondary-response theories are also hardly compatible with the evidence that, once established, PRF performance remains stable despite changes in the reinforcement schedule or in the initially trained CS (e.g., Crawford et al., 1985; Gottlieb, 2006; Robinson et al., 2014b). This conceptual problem disappears if we consider that associative learning causes the ability to evoke a CS-triggered response rather than the response itself. In this case, no secondary response is required to compensate for its detrimental effects under PRF.

Among the secondary-response theories, the most influential has been frustration theory (Amsel, 1958). According to Amsel, frustration basically results from the omission of an expected reward, but its development is a three-stage process (Amsel, 1958, 1992). First, unconditioned frustration directed to the food dish appears due to the absence of rewards on some trials. Second, as conditioning progresses, unconditioned frustration (also called frustration effect) becomes conditioned, directed toward the ambiguous CS. Conditioned frustration leads animals away from the frustrative CS. Third, conditioned frustration may be (or may not be) counterconditioned owing to the occasional associations of the CS with the delivery of rewards. The counterconditioning of frustration paradoxically leads animals to approach the frustrative CS and is assumed to explain the greater resistance of partially rewarded animals to extinction (PREE) compared to continuously rewarded animals. Hug and Amsel (1969) argued that counterconditioning is also responsible for the excitatory effects of PRF – also called partial reinforcement acquisition effect (PRAE). Those effects are assumed to be observed in two situations: (a) conditioned frustration was strong and converted into approach behaviour via counterconditioning and (b) conditioned frustration was initially weak (see also Amsel, 1992; p. 133).

The so-called frustration effect was demonstrated using two successive runways (Amsel and Roussel, 1952; Amsel and Ward, 1954; Amsel and Hancock, 1957). At the end of Runway 1 was a first goal box in which a reward was present or absent. At the end of Runway 2 was a second goal box in which a reward was always available. Amsel and colleagues showed that rats crossed Runway 2 at a higher speed when they experienced nonreward in the first goal box, suggesting that their increased performance resulted from frustration-based drive. The explanation provided by Hug and Amsel (1969) that reduced conditioned frustration can invigorate responding under PRF is hard to believe. If frustration was initially weak, why should this increase performance beyond the level of responding reached by animals trained under CRF? After all, frustration is a source of avoidance. Similarly, if frustration was high and became counterconditioned, no more than CRF performance would be obtained: at best, counterconditioning can cancel the detrimental behavioural effects of frustration, allowing PRF and CRF performances to be equivalent over training (see also section 6.1). The strong counterconditioned approach response is simply limited to compensating for the strong frustrative avoidance of the CS. In brief, I do not deny the temporary aversion caused by nonrewards (although aversion does not necessarily mean frustration), I am just suggesting that frustration and its counterconditioning are unlikely to be responsible for increased conditioned responding under PRF. I will not discuss the PREE in this article, but note that calling into question the existence of frustration under PRF should logically have implications for the interpretation of the PREE in terms of counterconditioned frustration.

A number of elements are hardly compatible with a possible development of frustration under PRF. First, the experience of frustration involves the violation of a 'positive' expectation such as an expected, desired food (Amsel, 1958; Papini, 2006). To develop a strong expectation of reward, an animal must often have received extensive training with that reward. However, McCain and McVean (1967) showed that enhanced running speed in Runway 2 could be obtained with PRF in the first goal box after only 16 trials. McCain (1968) even showed a difference in responding between rewarded and nonrewarded trials following a single reward. It is also noticeable that Amsel and Roussel (1952) found a frustration effect when animals were 100% rewarded in Runway 1, a situation that should not generate any frustration. Second, there is a potential conceptual problem with the suggestion that a frustration effect can develop in autoshaping conditions involving a 50% probability of reinforcement (Anselme, 2015). In this case, the probability (expectation) of reward on a given trial is equivalent to the probability (expectation) of nonreward. Thus, the resulting expectation is nil. If there is no

expectation to be violated, no frustration can emerge. Of course, it could be argued that the probability of reward is nevertheless equal to 50%. But if the animal cannot learn to expect reward delivery (like when probability is high) or nonreward (like when probability is low), past experience should return the message 'nothing to expect on the next trial' – the overall distribution is known but its details are not. So, provided that animals are trained under a 50% probability as soon as the first trials, the CS acquires an ambiguous meaning and there is no room for the occurrence of frustration in such a context. Take a concrete example: if an inheritance clause stipulates that you will receive $12,000, at a rate of $1000 every month during one year, you are likely to become frustrated if this sum is not received on your bank account at the expected time. In contrast, if the inheritance clause stipulates that the $12,000 will be received at a rate of $1000 on a random basis during the year, you have no reason to become frustrated following a prolonged period of monetary nonreward because you have nothing to expect relative to the time of its delivery. PRF is similar to the latter situation: the animal knows that a fixed number of pellets will be received within a training session, although it cannot predict their delivery on a given trial. Third (and accordingly), compared with CRF, an exposure to PRF conditions is known to reduce and even cancel the frustration induced by a successive negative contrast (SNC) procedure, where animals typically get access to a 4% sucrose solution after being trained with a 32% sucrose solution (Gonzalez and Bitterman, 1969; Mikulka et al., 1967; Pellegrini et al., 2004; Peters and McHose, 1974). This may indicate that animals accustomed to reward uncertainty develop less frustration once the reward amount is reduced. In a similar vein, Toates (2014) reports that, in humans, sexual frustration is more likely to develop in permissive settings (in which very liberal standards exist with respect to sexual rewards) than in non-permissive settings (in which those rewards can only be obtained following the compliance with strict conditions imposed by education-related inhibitions). Permissive settings may be compared to CRF conditions (allowing the individual to obtain rewards without great effort), while non-permissive settings look like PRF conditions (because obtaining rewards is harder). In non-permissive settings, the individual tends to consider that the setting is responsible for nonrewards and therefore accepts them more easily (Chappell et al., 1971).

In conclusion, frustration is likely to result from experimental procedures such as successive negative contrast and extinction because of the violation of expectations they generate, but its impact on performance under 50% reinforcement should not exist. If correct, this means that frustration cannot satisfactorily explain Pavlovian performance under PRF.

4. Incentive salience and the neurobiology of sign- and goal-trackers

Robinson and Berridge (2013) showed that salt non-deprived rats that received a salty solution (9%) directly into their mouth by means of oral cannulas acquired a strong aversion for the lever CS used to predict the delivery of the solution – the rats remained away from the lever. However, when those animals were retested two days later following the injection of two substances that mimic sodium deficiency/depletion brain signals (deoxycorticosterone and furosemide) and without additional training, they became avidly and immediately attracted by the previously aversive lever CS. Instant shift in behaviour was shown as soon as the first lever presentation during reexposure, irrespective of its fully learned aversive value. This result illustrates the fact that Pavlovian performance is directly altered following a change in the animal's motivational state, irrespective of the content of previous learning. The rats learned that the lever was a predictor of saltiness, but only the strength of the appetency (motivation) for salt determined the performance of those rats.

Instant shift in performance as a result of motivational change is impossible to understand by means of traditional learning positions, which presuppose that the CS-UCS association must be gradually relearned in the new motivational state to alter behaviour. A model like that of Hull (1943) has a motivational component (drive) that might potentially produce this effect. But this model (and drive-based models in general) is incompatible with some current data, showing that drives are unlikely to exist and that incentives depend on the individual's physiological states – i.e., they are not independent variables. Instant shift in performance is well captured by the incentive salience hypothesis (Berridge and Robinson, 1998). According to this hypothesis, what we 'want' (motivation) can be dissociated from what we 'like' (pleasure), although these two processes generally come together (we 'want' what we 'like' and 'like' what we 'want'). For example, drug addicts often report a strong desire for a drug they have limited pleasure to consume (Robinson and Berridge, 1993) and many people have sexual phantasms they do not attempt to satisfy (Toates, 2014). The incentive salience hypothesis addresses the 'wanting' (not the 'liking') process, which would mainly result from the release of dopamine in the nucleus accumbens from the ventral tegmental area. This process exerts a direct control on Pavlovian performance. Accordingly, the more a CS approached, the more dopamine (measured using fast-scan cyclic voltammerty) is released in the nucleus accumbens of rats during a 10-s CS presentation (Sunsay and Rebec, 2008, 2014). There is a body of evidence that the incentive salience of a CS is independent of the pleasure associated with UCS consumption and independent of the incentive salience of the UCS during acquisition (e.g., Cannon and Bseikri, 2004; Cagniard et al., 2006a, 2006b; Flagel et al., 2011a; Peciña et al., 2003; Robinson and Flagel, 2009; Robinson and Berridge, 2013; Smith et al., 2011; Tindell et al., 2005, 2009). In Robinson and Berridge's (2013) study about salt disgust/hunger reported above, there were dramatic increases in neuronal c-Fos expression (as a measure of neuronal activity) within several mesocorticolimbic regions when sodium-depleted rats were presented with the CS lever (e.g., +269% in the ventral tegmental area and +293% in the nucleus accumbens) despite aversive learning conditions.

Sign-trackers release more phasic dopamine in their nucleus accumbens than goal-trackers during training, and dopamine antagonism has a detrimental effect on the expression of sign-tracking responses in sign-trackers but does not alter goal-tracking responses in goal-trackers (Flagel et al., 2007, 2011a). Also, the phasic activity of dopamine neurons in response to repeated reward delivery decreases in sign-trackers but is unaltered in goal-trackers (Flagel et al., 2011a). According to the reward prediction error theory, such a difference indicates that sign-trackers are more effective in learning the predictive value of the CS than goal-trackers. However, this interpretation is incompatible with the evidence that the progression of response rates to the CS lever (for sign-trackers) and to the food dish (for goal-trackers) is comparable over training. In fact, sign- and goal-trackers can perfectly learn the predictive value of the CS, but the higher dopamine levels in the nucleus accumbens of sign-trackers make them to prefer the CS location to the food dish location (Flagel et al., 2007; Meyer et al., 2012; Robinson and Flagel, 2009). In contrast, goal-trackers seem to develop a more cognitive strategy, which consists of using the CS as an informational stimulus, unattractive in itself. Accordingly, the brain of goal-trackers shows thalamocortical and corticostriatal activation, while that of sign-trackers essentially displays subcortical activation (Flagel et al., 2011b). In addition, goal-trackers show stronger sustained attentional focus (stability, non-random performance, and self-control) than sign-trackers, caused by more effective top-down cholinergic control of attention (Paolone et al., 2013). Of course,

this is not to say that goal-trackers exhibit no 'wanting' with respect to the food pellets. The development of incentive salience for the food dish (goal-tracking) seems to depend on dopamine release as well (DiFeliceantonio and Berridge, 2012). But this behaviour might require lower dopamine levels than its sign-tracking counterpart. Only the animals that release more dopamine can 'break away' from the location of a primary reward and focus their motivation on more distal cues that predict the occurrence of that primary reward (Flagel et al., 2007, 2011a; Howe et al., 2013).

5. Computer modelling of incentive salience

The incentive salience hypothesis provides us a new picture of Pavlovian conditioning. In this section, I briefly describe four models of incentive salience and some of their limits. Subsequently, I will present a new model, the extended incentive hope model (see also Anselme, 2015), which circumvents some of those shortcomings while trying to go further.

A model developed by McClure et al. (2003) attempts to reconcile incentive salience with prediction error, calculated by means of the temporal difference (TD) algorithm (Sutton and Barto, 1981). Briefly, on a trial, prediction error is computed as a function of instantaneous reward value and of the estimated value of reward prediction gradually established through TD learning over past encounters with the CS-UCS pairings. Once prediction error becomes nil, reward prediction is maximal. In their model, McClure et al. (2003) equate the learned value of the CS acquired by the TD method with CS-triggered 'wanting'; the effectiveness of predictive learning would determine how much a CS is 'wanted'. Thus, any change in the incentive salience attribution to a CS must result from the relearning of the CS-UCS association in a new motivational state. Relearning is necessarily a gradual process hardly compatible with the instant shift in performance reported by Robinson and Berridge (2013) as well as other studies (e.g., Tindell et al., 2005, 2009). In fact, TD learning might be more relevant to instrumental conditioning, where a change in performance is known to depend on the opportunity to relearn an action-outcome association in a new motivational state (Dickinson and Dawson, 1987; Balleine, 1992). However, the development of instrumental performance is insensitive to dopaminergic agents such as pimozide and α-flupenthixol (Dickinson et al., 2000).

Interestingly, Dayan and Balleine (2002) proposed an integrative model that can account for the immediate sensitivity of Pavlovian conditioned responding to motivational shifts, in addition to instrumental conditioning and Pavlovian-instrumental transfer. Their model uses the TD learning algorithm for both forms of conditioning, but I will only consider the Pavlovian form here. It is suggested that appetitive UCSs activate an appetitive system via a motivational gate sensitive to the individual's physiological state, and that CSs can also do it more directly – shortcutting the motivational gate (see Dickinson and Balleine, 2002). The appetitive system's output is a dopaminergic prediction error signal, allowing Pavlovian control over actions and providing a teaching signal for values. Dopamine release modulates the energy level of Pavlovian consummatory habits, depending on hard-wired preparatory habits that control approach or withdrawal. This model shows that prediction error is perhaps not incompatible with instant shifts in motivational states. Nevertheless, although cognitive expectancies/predictions of future reward certainly play a role in Pavlovian conditioning, they might be poorly related to the computation of incentive salience (see section 4; for a review, see Berridge, 2012). The predictive efficiency of future reward provides the ability to approach or to avoid a CS, but this is unlikely to determine the strength of those behaviours.

A model proposed by Zhang et al. (2009) reflects the core principles behind the incentive salience hypothesis with more accuracy. This model also uses the TD algorithm. However, the motivational salience of a CS is only indirectly related to its learned predictive value. A κ factor is used to dynamically control incentive salience attribution to a CS depending on moment-to-moment changes in an individual's physiological state. Thus, cue-triggered 'wanting' is computed following physiological shift, independently of the opportunity for relearning the altered reward. If $\kappa > 1$, the motivational salience of a CS increases beyond its learned value and the reward-related CS becomes 'wanted'. If $\kappa < 1$, the motivational salience of the CS decreases below its learned value and the reward-related CS is avoided. Finally, if $\kappa = 1$, incentive salience mimics TD learning and the CS's motivational salience equals its learned value. The Zhang et al.'s model can effectively account for instant motivational shifts induced by natural physiological fluctuations (appetite versus satiety) and by pharmacological agents (agonists and antagonists) known to modulate dopamine levels. However, it remains of limited explanatory power beyond that particular fact. For example, strongly deprived animals typically show higher response rates than non-deprived animals (Hull, 1943; Reynolds and Pavlik, 1960). In the Zhang et al.'s model, the level of deprivation has no effect on incentive salience attribution if food deprived and food non-deprived animals are trained and tested under constant physiological state ($\kappa = 1$ in both cases because no change occurs between training and test).

Lesaint et al. (2014) developed a model of performance in autoshaping which reproduces a number of results, especially the proportion of sign- and goal-trackers found in the literature. They postulate the existence of two learning systems. A model-free (MF) system is required to evaluate features by trials and errors and favours approach to the lever CS as a predictor of food delivery. And a model-based (MB) system is used to learn the structure of the task and favours approach to the magazine until food is delivered. These systems make sign- and goal-tracking more or less likely depending on their respective contribution – controlled by a combination parameter, ω. When ω is high, the MF system is more strongly involved than the MB system, so that the emergence of sign-trackers is facilitated. In contrast, when ω is low, the MB system is recruited to a larger extent, generating more goal-trackers. This model can simulate the spectrum of responses found with real animals, from pure sign-tracking to pure goal-tracking – a dimension neglected by traditional models (Lesaint et al., 2014). However, the model does not take timing into consideration and is, therefore, unable to account for the gradual development of performance in autoshaping. Also, it is a learning rather than motivational model (see also Kaveri and Nakahara, 2014). For example, the model predicts that the detrimental effects of α-flupenthixol on Pavlovian performance can only occur after the relearning of the task, as in the case of McClure et al.'s (2003) interpretation of incentive salience.

Beyond the specific limits of each model, it is important to note that they all fail to account for the excitatory effects of PRF. In the McClure et al.'s (2003) and the Dayan and Balleine's (2002) models, CS-triggered 'wanting' is a function of UCS predictability. As a result, 'wanting' should be more intense when the probability of the UCS is 100% than when it is lower. In the Zhang et al.'s (2009) model, reward uncertainty can be interpreted in terms of a κ modulator, but there is no mechanism describing how uncertainty can alter the κ value (Anselme, 2015). We need a mechanism that explains how and why κ is increased rather than decreased under uncertainty. Finally, in Lesaint et al.'s (2014) model, the predictive power of the lever CS – resulting from the fact that the CS always precedes reward delivery and is never presented during the ITI – is assumed to play a major role in determining its attractiveness. Thus, the occasional absence of reward during training should logically reduce CS-triggered 'wanting'.

Overall, these predictions are not compatible with the hypothesis that the increased vigour of responding observed under PRF might reflect enhanced 'wanting'. Yet, in addition to the physiologically determined dopamine-dependent incentive properties of reward, dopamine release is boosted in the presence of reward uncertainty. For example, Hart et al. (2015) trained rats for six sessions of Pavlovian conditioned approach with an 8-s lever CS. They used fast-scan cyclic voltammetry to show the magnitude of the dopamine response to CS presentations across the full range of probabilities (0, 0.25, 0.5, 0.75, and 1). On session 2, it appeared that the magnitude of the dopamine response increased in proportion to probability, whether recorded during the early period (0.4–1.4 s), the peak period (1.5–2.5 s), or the late period (6.9–7.9 s) of CS presentations. However, on session 6, the dopamine response varied in a quadratic rather than linear manner during the peak and late periods: Dopamine release was higher at $p = 0.5$ (maximal uncertainty) and gradually decreased until reaching its minimal values at $p = 0$ and at $p = 1$ (no uncertainty). This result confirms previous findings by independent research teams (de Lafuente and Romo, 2011; Dreher et al., 2006; Fiorillo et al., 2003; Linnet et al., 2012; Preuschoff et al., 2006; Singer et al., 2012; Tan and Bullock, 2008; Zack et al., 2014). Conversely, the use of dopamine agonists, such as pramipexole, in patients suffering from Parkinson's disease is known to increase the risk of developing pathological gambling (Dodd et al., 2005; Voon et al., 2006).

6. Incentive hope: concept and model

Despite its apparent simplicity, Pavlovian conditioning has not yet been fully captured and a more integrative approach is to be proposed (Dayan and Berridge, 2014; Meyer et al., 2014). The model presented in this section is in keeping with the incentive salience hypothesis. It is an attempt to explain the development of sign-tracking performance under reward uncertainty, and can also be used in a number of situations involving no uncertainty.

The core concept behind this new framework is that of *incentive hope*, which can account for the excitatory effects of PRF. Incentive hope does not involve anything that could not be used to discuss animal behaviour – at least, with respect to mammals and birds – and can be defined as *the direct consequence of 'wanting' under uncertainty*. A mouse that spotted a food source can hope that it is still there the next day because the mouse 'wants' that food while its presence is not guaranteed (e.g., the food was possibly consumed by competitors meanwhile). Incentive hope is more likely to reflect the mouse's psychological state than expectation in this situation because there is no strong evidence allowing it to expect the presence or the absence of that food with a high probability. How can incentive hope enhance Pavlovian performance under PRF? Two distinct psychological mechanisms, described below, are assumed to operate in concert: counterconditioning and incentive hope.

6.1. First mechanism: counterconditioning

It is reasonable to think that reward uncertainty might lead to a noticeable attenuation of incentive salience attribution to a CS because of the presence of nonrewards – which are basically 'unwanted'. But counterconditioning can explain why this does not generally happen. Counterconditioning is the replacement of an undesired response to a stimulus by a desired response, resulting from its association with a reinforcer (positive or negative, depending on the valence of the desired response). A form of counterconditioning often used in psychotherapy is systematic desensitization, where an approach response is progressively induced towards an initially fearful/disliked stimulus.

In autoshaping, CSs associated with nonrewards may come to act as conditioned inhibitors (CS–), but there is evidence that such CSs often produces low conditioned inhibition – conditioned inhibition is found when responding to a CS+ (predicting reward) decreases after pairing it with a CS– (predicting nonreward). Counterconditioning can explain this fact: the occasional delivery of reward following a CS presentation leads animals to approach the CS, despite its association with occasional nonrewards. For example, Stagner and Zentall (2010) found that pigeons prefer a CS_A+/CS– compound to a CS_B+/CS_C+ compound, even when the CS– occurs four times as often as the CS_A+ (i.e., CS_A+/CS–: 20% chance of 10 pellets and 80% chance of zero pellets; CS_B+/CS_C+: 100% chance of 3 pellets). Using the same experimental design, Laude et al. (2014) showed that, early in training, pigeons tend to prefer the CS_B+/CS_C+ compound to the CS_A+/CS– compound. However, later in training, the preference completely reverses, suggesting a decrease in the inhibition induced by nonrewards as conditioning progresses. This process might explain why, in humans, problem gamblers show reduced sensitivity to conditioned inhibitors (Breen and Zuckerman, 1999; Brunborg et al., 2010).

The attenuation of conditioned inhibition for nonrewards over training might be a consequence of their counterconditioning owing to the occasional presence of rewards (e.g., Dickinson and Pearce, 1976). In Stagner and Zentall's (2010) paradigm, the attenuation of conditioned inhibition may have resulted in a change in preference from the optimal option (3 pellets per trial for sure) to the suboptimal option (2 pellets per trial, on average) because the suboptimal option potentially allowed the sudden delivery of a larger food amount. However, counterconditioning also acts in the absence of choice and of differential reward rates. In the case of serial autoshaping, counterconditioning can only allow performance under PRF to be similar to (i.e., not lower than) performance under CRF – by cancelling the detrimental effects of nonrewards (Fig. 2A). Counterconditioning cannot make performance higher under PRF than under CRF because this process has no excitatory properties. In psychotherapy, desensitizing an avoidance response to an initially fearful stimulus (e.g. the sight of spiders) is unlikely to produce subsequent stronger approach response than it would be if the stimulus was neutral. In pigeons autoshaping, counterconditioning should already have reached a strong (if not maximal) value at $p = 0.25$ since PRF responding is not lower than CRF responding (see Gottlieb, 2004, 2006).

At very low probability values (e.g., $p < 0.1$), the difficulty for an animal to learn a Pavlovian task results from the high proportion of nonrewarded trials, which may prevent the occurrence of counterconditioning – a difficulty that can be remedied by increasing the number of training sessions (e.g., Gallistel and Gibbon, 2000). At the probability values considered in this paper, counterconditioning is assumed to have a maximal impact, making performance under PRF and CRF identical. But an equation describing how counterconditioning depends on reward probability and past experience has to be developed.

6.2. Second mechanism: Incentive hope

While counterconditioning equates PRF and CRF responding, I argue that incentive hope explains why uncertainty can generate higher responding than certainty. Simply, receiving a reward when there was a possibility of missing it should raise the reward's incentive salience. If you were attached to an object you lost and then have found it after deploying some effort in seeking it, you may come to cherish that object more than if it had never been lost. The additional valuation attributed to a reward, given the eventuality of nonreward, is incentive hope (Fig. 2A). Like incentive motivation ('wanting'), incentive hope is not supposed to be a conscious process – even though the hopes we can report verbally inter-

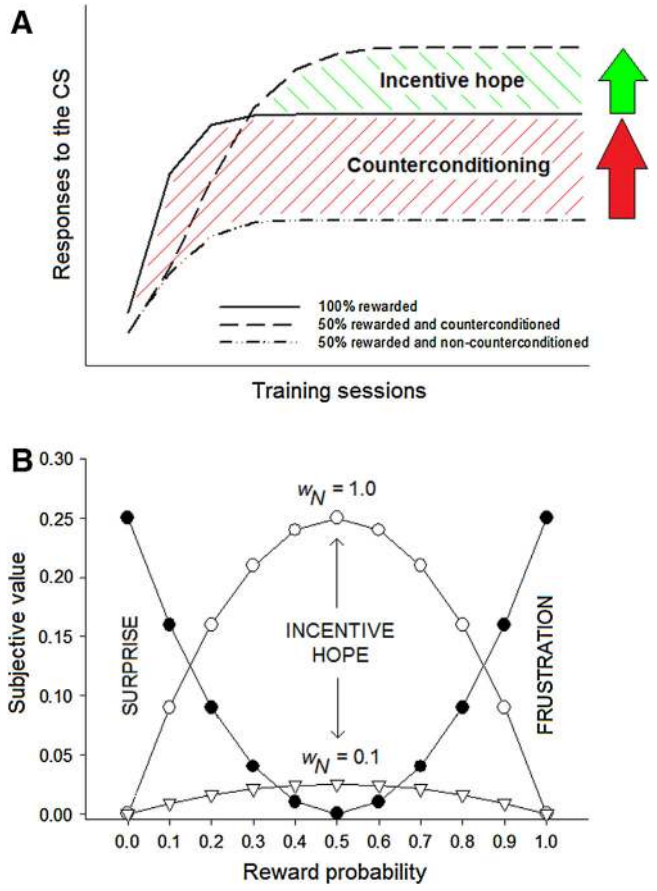

Fig. 2. Incentive hope. (A) Psycho-behavioural mechanism. The absence of counterconditioning under PRF might lead to a lower performance than that obtained under CRF. Counterconditioning gradually cancels the aversive influence of non-rewards, allowing the PRF and CRF performances to become equivalent. Nevertheless, receiving a reward when there is a possibility of nonreward (uncertainty) makes that reward more attractive than the same reward obtained with certainty. The additional valuation produced by uncertain rewards relative to certain rewards is incentive hope, which raises counterconditioned performance under PRF compared to performance under CRF. For clarity, the graph represents counterconditioning and incentive hope in successive order, but their action on behaviour is likely to be combined in reality – both processes come together. (B) Incentive hope and expectation as a function of reward probability. Incentive hope (white forms) is maximal when uncertainty is maximal ($p = 0.5$), while expectation (black circles) is maximal when uncertainty is nil ($p = 0$ or $p = 1$). Both surprise and frustration result from the violation of a strong expectation of nonreward and of reward, respectively. Reducing an animal's deprivation level (represented as decreased w_N values) attenuates the development of incentive hope (white circles: $w_N = 1.0$; white triangles: $w_N = 0.1$). Expectation is not assumed to vary according to deprivation level because anticipation is not a motivational parameter in itself.

act with our conscious thinking. It is not here the place to discuss whether conscious hopes may derive from incentive hope, but note that there is some indications that conscious desires would not exist without unconscious 'wanting' (for a recent discussion, see Anselme and Robinson, in press).

Given the low inhibitory power of nonrewards, it can be argued that incentive hope is a *sufficient condition* to cause a greater incentive salience of uncertain rewards compared with certain rewards. As said, incentive hope implies that a specific item (e.g., food) is both 'wanted' and not guaranteed, so that its basic mathematical formulation is:

$$\hbar = w_N \times p(1-p) \quad (1)$$

where $\hbar$ is incentive hope, w_N (need-induced 'wanting') is the incentive motivation resulting from the individual's physiological deficit, and p is the probability of reward delivery following a CS presentation. The w_N values $\in [-1,+1]$, ranging from total aversion to irresistible attraction. The expression $p(1-p)$ is the probabilistic representation of uncertainty, as it reaches its maximal value at $p = 0.5$ and becomes nil at $p = 0$ and at $p = 1$. Incentive hope can develop ($|\hbar| > 0$) provided that $w_N \neq 0$ and $0 < p < 1$, indicating that incentive hope is only possible under PRF. Under CRF, when reward is delivered with certainty ($p = 1$), the development of incentive hope would not make sense since the individual knows that a reward will be obtained after each CS presentation. In the same logic, a non-contingent CS-reward association ($p = 0$) also abolishes incentive hope because the individual knows that the CS is meaningless.

The concept of hope has already been used by psychologists in psychotherapy contexts. For example, Mowrer (1966) saw hope as a supplemental source of energy for the completion of an instrumental action. A more sophisticated interpretation of hope was suggested by Snyder (1994), for whom hopeful thinking depends on the individual's perceived capacity to find routes to a desired goal and to his or her motivation to use those routes. For hopeful thinking to develop, an individual must have a perceived 50% probability of successfully attaining goals, so that the goal-seeker is prepared for the possibility of loss (Snyder, 1996). In a sense, Snyder's concept of hope is akin to mine because it makes a reference to both motivation and uncertainty. However, the concept of incentive hope presented here refers to a motivational rather than cognitive process. No need to pursue any cognitive goal to experience incentive hope, the individual just has to 'want' a non-guaranteed stimulus. I am aware of using an anthropocentric terminology (we do not know about the subjective contents of animal minds), but we need concepts to understand what animals do. In this respect, the concept of hope is not more extravagant than many other concepts that have demonstrated their usefulness in the animal literature, such as curiosity, pleasure, and fear. Hope perfectly reflects the motivational state of an individual that 'wants' a reward whose delivery is unpredictable, and its use is necessary to account for the excitatory effect of reward uncertainty in motivational terms. To suggest that uncertainty causes a surge of incentive motivation is not enough because, as explained earlier, the modern theory of incentive motivation predicts that uncertainty should reduce the attractiveness of CSs. Incentive hope represents the psychological mechanism allowing reward uncertainty to enhance conditioned responding.

The concept of incentive hope is the exact opposite to that of expectation (Fig. 2B) – i.e., maximal in the absence of expectation on a given trial ($p = 0.5$) and nil when expectation is strong (close to $p = 1.0$ and $p = 0.0$). In this view, the violation of a strong expectation of reward ($p \cong 1$) should generate maximal frustration, while the violation of a strong expectation of nonreward ($p \cong 0$) should generate maximal surprise. This is not to say that incentive hope controls behaviour at intermediate probability values and that expectation does it at the extremes of the probability range. If this happened, the two aspects would be operationally undistinguishable and performance would be the same across the full range of probabilities. In fact, expectation is unlikely to be a motivation in itself because we can expect events for which we have no motivation at all (Berridge, 2012). For an expectation to be a motivation, the expected event must be 'wanted' – and also possibly hoped (Anselme and Robinson, in press). Also note that, in this context, frustration is likely to be rapidly counterconditioned and that it is unclear why surprise should cause a long-term increase in conditioned responding. This suggests that sign-tracking performance is a product of 'wanting' and incentive hope, not a product of expectation. Total 'wanting' (w_T) can be expressed as follows:

$$w_T = w_N + \hbar \quad (2)$$

6.3. Incentive hope, frustration, and dopamine

The incentive hope concept and frustration theory predict opposite results with respect to the involvement of dopamine in PRF. First of all, frustration attenuates dopamine release. Early studies showed that dopaminergic agents have no effects on SNC, whether injected during the first or the second postshift day (Flaherty, 1996). Since SNC is likely to generate unconditioned frustration (the first postshift day) and conditioned frustration (the next days), this suggested that frustration is not under dopaminergic control. But Phillips and colleagues used brain dialysis to monitor changes in dopamine efflux in the nucleus accumbens and showed that SNC causes attenuation in dopamine levels (Genn et al., 2004). Accordingly, they also demonstrated that successive positive contrast (SPC), where a 4% sucrose solution is suddenly switched to 32%, enhances dopamine levels (Phillips et al., 2008). The modulation of dopamine efflux is correlated with consummatory behaviour, whose vigour is decreased under SNC and increased under SPC. If PRF is a source of frustration, it should therefore have a similar detrimental effect on dopamine release. Second, the counterconditioning of frustration should also be dopamine-independent – at least in mammals (Papini, 2014). Amsel predicted that the PREE is due to the counterconditioning of frustration, and there is substantial evidence that the PREE is not under dopaminergic control. Since the excitatory effects of reward uncertainty are also assumed to result from the counterconditioning of frustration (Hug and Amsel, 1969; Amsel, 1992), those effects should logically be insensitive to dopamine manipulations as well. In the opposite, incentive hope is part of the 'wanting' process and its formulation (Eq. (2) predicts higher levels of dopamine in the nucleus accumbens under PRF. The incentive hope concept, but not frustration theory, predicts that reward uncertainty will enhance dopamine efflux (and performance) compared with reward certainty.

It is important to note that the studies that report the effects of brain lesions on performance under PRF fail to provide strong theoretical conclusions with respect to the processes involved. For example, lesions of the amygdala cancel the excitatory properties of PRF in a double runway (Henke, 1977). Given the role of the amygdala in aversive conditioning (Davis, 1992), it could be suggested that PRF-induced excitement is the result of an aversive experience. However, the amygdala is also recruited in incentive motivation, especially its basolateral part (Robbins and Everitt, 1996) and its central nucleus (DiFeliceantonio and Berridge, 2012; Mahler and Berridge, 2009; Robinson et al., 2014a). In Henke's (1977) study, the lesions were extensive, including the basolateral amygdala and also, perhaps in part, the central nucleus. Thus, it is here difficult to conclude anything about the psychological origin of reduced reactivity under PRF. Interestingly, Leszczuk and Flaherty (2000) showed that, following massive lesions of the nucleus accumbens, rats maintain their ability to discriminate the differences in reward magnitude (see also Bowman and Brown, 1998), suggesting that this brain region is not involved in the control of expectation. Perhaps more surprisingly, they also showed that rats with such lesions continue to exhibit a decrease in responding (SNC) when reward magnitude is suddenly reduced. One possible explanation might be that, because the SNC procedure is not a Pavlovian procedure in the strict sense (there is no CS), performance is controlled by expectation (as a cognitive process) rather than incentive motivation during training. Thus, when rats experience negative contrast, motivational processes are recruited and dopamine release is reduced in non-lesioned rats. In lesioned rats, dopamine release is also reduced but lesions have no effect on the expression of behaviour because the destruction of dopamine receptors mimics the reduction in dopamine release. Lesioned and non-lesioned rats therefore exhibit the same SNC pattern.

If dopamine can help distinguish incentive hope from frustration, it appears that incentive hope and the prediction error hypothesis make identical predictions with respect to dopamine release – which is, in both cases, assumed to be enhanced by reward uncertainty. However, the interpretations are different, and this has implications for the expression of behaviour. The prediction error hypothesis states that reward uncertainty enhances dopamine levels because of a learning failure of a CS-UCS association (see Schultz, 1998). This view cannot predict any increase in conditioned responding under reward uncertainty because a failure to learn is not a motivational factor in itself. In fact, animals exposed to such poor learning conditions might become demotivated by irreducible unpredictability rather than aroused in the task. Showing that dopamine is involved in uncertainty-induced conditioned responding would allow us to disentangle the hypothesis of incentive hope from that of prediction error.

Finally, note that the incentive hope hypothesis is not reducible to that of incentive salience. The incentive salience hypothesis suggests that rewards are 'wanted' and approached; it does not predict how uncertainty should impact motivation and behavior. Incentive hope is assumed to boost dopamine release in the nucleus accumbens – although other more distinctive (unknown) processes might also be involved. Is this different from 'wanting' manipulations by means of dopaminergic drugs? Yes, drugs directly alter 'wanting' at a neurobiological level, while incentive hope is the result of a psychological process. In other words, drugs can induce 'wanting' for an 'unwanted' stimulus (and vice-versa), while incentive hope has no effect if the individual does not 'want' a reward first. A demonstration that both processes are distinct is that enhanced drug-induced dopamine release in goal-trackers reinforces goal-tracking behavior (DiFeliceantonio and Berridge, 2012), while reward uncertainty transforms potential goal-trackers in sign-trackers (Robinson et al., 2015).

6.4. Development of sign-tracking performance

The extended incentive hope model consists of an equation that describes how the motivation (w_T) for a UCS is gradually transferred to a CS. In a first step, I show how the model captures the attribution of incentive salience from a single UCS to a single CS. The complete function that defines the UCS → CS motivational transfer is:

$$\Omega = [w_N + \hbar]/[\gamma + \alpha + e^{-\delta t+\theta}] \quad (3)$$

where γ is a parameter denoting the resistance to the UCS → CS motivational transfer, α is an attentional threshold, δ is a slope adjustment (learning-related) parameter, t is the session number (e.g., $t = 4$ means training session 4), and θ is a horizontal translation parameter. (The numerator's components were defined earlier.)

How does equation 3 work? The equation comprises two distinct parts. Firstly, its numerator is the method used to compute w_T, which determines the curve's asymptote – everything else being equal. Importantly, w_T denotes the motivation developed for the UCS (not that for its predictive CS), given the individual's physiological state and the probability of reward delivery. A noticeable property of the extended incentive hope model is thereby that asymptotic performance is a consequence of the individual's motivation for the UCS rather than a product of learning. For a deprived animal ($w_N > 0$), the asymptote should be higher under PRF at intermediate probability values ($\hbar > 0$) than under CRF ($\hbar = 0$). Secondly, the denominator of equation 3 describes how the UCS motivation is gradually attributed to the CS. This phenomenon is assumed to depend on three distinct processes: the memory consolidation of the CS-UCS association ($e^{-\delta t}$), the resistance to the transfer of motivation from the UCS to the CS (γ), and attentional focus (α). It must be noted that α plays a role in the control of performance only in a

context of stimulus competition (otherwise, $\alpha = 0$), a scenario that will be discussed further (section 7.4).

- Memory consolidation. Organisms can certainly learn when reality does not fit their expectations (error correction), but it is reasonable to suggest that organisms can also learn in the absence of any expectations (e.g., because they detect a correlation between two events). Memory consolidation is the mechanism allowing an unstable memory trace to become enhanced and stabilized over a period of minutes to years (Stickgold and Walker, 2005). It is now recognized that sleep is necessary for the consolidation processes (Stickgold and Walker, 2005), explaining why a session-per-session (rather than trial-per-trial) model is here used (the evidence that the number of trials also play a role in associative learning is recognized but not considered; but see Papini and Dudley, 1993; Gottlieb, 2008). The δ variable denotes the consolidation rate of the CS-UCS association. Its numerical value increases with the (mean) length of the ITI (k) and decreases with the proportion of nonrewarded trials within a session, computed according to the number of rewarded trials (n_R) relative to the total number of trials (n_T). We will see that δ may also increase or decrease as a result of the α value in a context of stimulus competition:

$$\delta = k - \alpha - (n_T - n_R)/n_T \quad (4)$$

Overall, the expression $e^{-\delta t}$ denotes the hypothetical change in memory consolidation when a CS-UCS association is learned from session to session. There is no a priori method allowing the number k to be set (this decision falls to the model's user), but the range of realistic k values remains limited (I suggest, $1 \leq k \leq 4$) and the selected value has no effect on the curve's asymptote. The longer the ITI, the higher is k and the steeper is the slope of the curve, as suggested by Gibbon et al.'s (1980) results. As predicted here, the ITI length has no effect on asymptotic performance (Gottlieb, 2006). The k value is compatible with the trial-spacing effect relative to a CS, but it is unrelated to any theoretical presupposition. The trial-spacing effect means that increasing the temporal interval that separates the CS-UCS pairings facilitates the development of conditioned responding (e.g., Barela, 1999). Nevertheless, it is reasonable to think that the computation of k is related somehow to Rescorla and Wagner's (1972) V value or to the Gallistel and Gibbon's (2000) I/T ratio, which might indicate when the learned ability to express a conditioned response can emerge (see section 2). Unlike ITI duration, nonrewarded trials ($n_T - n_R > 0$) impede the learning of a CS-UCS association. A high proportion of nonrewarded trials makes learning less reliable (Rescorla and Wagner, 1972). However, contrary to the Rescorla-Wagner model's prediction, nonrewarded trials here reduce the learning rate of the task (i.e., the curve's slope is gentler) without altering asymptotic performance. A high proportion of nonrewarded trials simply means that more trials are necessary to reach the asymptote (e.g., Gallistel and Gibbon, 2000). Learning 'inertia' induced by nonrewarded trials is obviously nil if all trials are rewarded, i.e. at $p = 1.0$ (because $n_T = n_R$), and increases as probability decreases (because $n_T > n_R$). The lower the probability of reward, the gentler is the slope of the curve.

- Resistance to motivational transfer. Associative learning is necessary for the ability to transfer motivational salience from the UCS to the CS, but it is not the cause of that transfer (Berridge, 2012; Meyer et al., 2012; Robinson and Flagel, 2009). Motivational transfer is assumed to depend on a resistance factor, γ, which may vary from individual to individual – and is likely to result from genetic predispositions that control cue-triggered dopamine release (Flagel et al., 2007). The γ factor consists of an improvement of the model that does not alter the predictions of its previous version (Anselme, 2015). Briefly, the higher γ ($\in]0, +\infty]$), the lower is the ability of the CS to become motivationally attractive. Pure goal-trackers have an elevated γ value ($\gamma \geq 20$), resulting in the absence of incentive salience attribution to the CS. In contrast, pure sign-trackers have a much smaller value ($\gamma \leq 2$), allowing a gradual attribution of incentive salience to the lever during training. Previously, a constant (equals to 1) without psychological meaning was used in place of γ (Anselme, 2015). This was not a problem because I mainly attempted to account for averaged sign-tracking responses in sign-trackers, and $\gamma = 1$ is an optimal value for the simulation of that behaviour. But the use of a mere constant may lead to misinterpret the origin of goal-tracking responses in goal-trackers (see section 7.3). Importantly, a mean γ value can be used to characterize a group of sign-trackers, but a γ value can also be assigned to a particular individual, and thus represent inter-individual differences across the entire spectrum of responses that may exist between pure sign-tracking and pure goal-tracking. The γ factor contributes to shape inter-individual differences in responding, although more research is needed to identify its exact brain mechanism.

Neurobiologically, a low γ value means an effective transfer of phasic dopamine from the UCS to the CS during training. The higher the γ value, the stronger is the resistance to that transfer. Although the transfer of phasic dopamine from the UCS to the CS is traditionally interpreted in terms of a prediction error signal (Schultz, 1998; Fiorillo et al., 2003; Dreher et al., 2006), I rather suggest that it consists of a shift in the incentive salience attribution (Anselme, 2013). Let's see how this may happen. Under CRF, the CS gains in predictive value as training progresses, and hence the UCS becomes more and more predictable. For this reason, the UCS loses its attractiveness – full predictability is dull, causing a suppression of phasic dopamine signals (Anselme, 2013). In contrast, because the CS is not preceded by any stimulus capable of predicting its occurrence, it comes to release phasic dopamine as its association with the UCS is learned. In this view, the transfer of phasic dopamine from the UCS to the CS is caused by the fact that the motivational, unpredictable stimulus has changed (moving from the UCS to the CS). At least, this is what might happen in sign-trackers. In goal-trackers, the reasons for the absence of a transfer of 'wanting' remain unclear. One possible explanation is that goal-trackers are less sensitive to uncertainty than sign-trackers, so that the uncertainty of the CS under CRF is not sufficient in itself to generate a motivational transfer of phasic dopamine signals. An empirical support to this view comes from autoshaping experiments under PRF (Fig. 3). We showed that reward uncertainty generates more sign-trackers and fewer goal-trackers than reward certainty (Robinson et al., 2015), a finding already suspected – but not demonstrated – in our previous works (Anselme et al., 2013; Robinson et al., 2014b). In other words, when the UCS is kept uncertain (in addition to the CS), individuals that should develop goal-tracking responses under reward certainty are able to attribute incentive salience to the CS. More uncertainty converts potential goal-trackers in sign-trackers. In a sense, uncertainty seems to have the reverse effect of sounds and lights CSs, which promote goal-tracking in individuals that have a pharmacological profile of sign-trackers (Meyer et al., 2014).

Note that equation 3 does not allow a step-by-step calculation of the change in CS attractiveness from one session, t–1, to the next, t. But a simple mathematical transformation can represent this gradual change (for details, see Anselme, 2015; Appendix):

$$\Omega(t) = \{[1 + e^{-\delta(t-1)+\theta}]/[1 + e^{-\delta t+\theta}]\}\Omega(t-1) \quad (5)$$

Eq. (5) can be used to compute an output at a given session (e.g. $t = 3$) on the basis of the result obtained at the previous session ($t = 2$). Finding out an algorithmic model of motivational dynamics

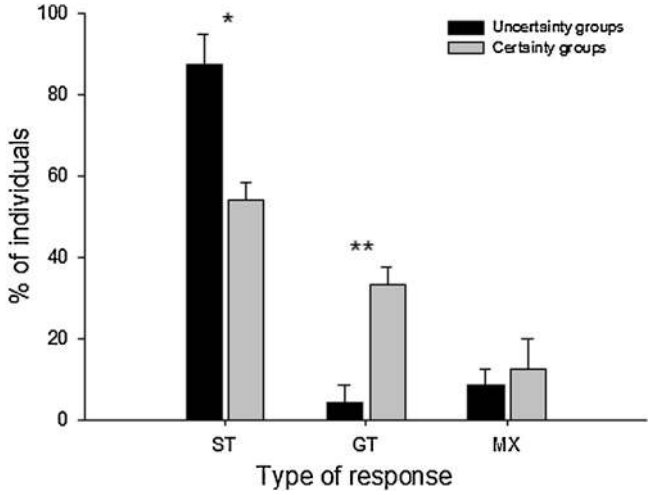

Fig. 3. Proportion of sign-trackers (ST), goal-trackers (GT), and mixed individuals (MX) trained under certainty and uncertainty conditions. Although more sign-trackers than goal-trackers are produced in both training conditions, uncertainty generates more sign-trackers and fewer goal-trackers than certainty. Modified from Robinson et al. (2015). Legend: * $p < 0.05$; ** $p < 0.01$.

is a topical question (Dayan and Berridge, 2014) and equation 5 is a first attempt to tackle this issue – at the session (rather than trial) level. In summary, the extended incentive hope model is an integrative approach to sign-tracking performance without any explicit reference to associative learning. Associative learning determines the ability to respond (e.g., $V_n > \theta$), memory consolidation controls the slope of the curve ($e^{-\delta t}$), and incentive motivation for the UCS (w_N and $\hbar$) and its transferability to the CS (γ) determine the asymptotic level of sign-tracking performance. This means that incentive motivation is not here assumed to compensate for any deficit in associative strength, a view that justified secondary-response theories.

7. Predictive scope of the model

7.1. The effects of reward uncertainty

This section will be brief because the capacity of the model to predict the effects of reward uncertainty on incentive salience attribution has already been discussed elsewhere (Anselme, 2015) – the model was initially designed in this aim. Nevertheless, it is here important to report the model's main predictions in this respect.

- Firstly, as depicted in Fig. 4A and 4B, the model correctly predicts a slower behavioural performance early in training under PRF than under CRF (Crawford et al., 1985; Gonzalez, 1973; Gottlieb, 2004; Papini and Overmier, 1984, 1985; Wasserman et al., 1975). The unreliability of the CS-UCS association caused by the presence of nonrewarded trials weakens the memory consolidation of that association, so that more trials are required for its acquisition. In the model, this phenomenon results from reduced δ value when $0 < p < 1$.
- Secondly, as also shown in Fig. 4, the model correctly predicts a higher asymptotic performance later in training under PRF than under CRF (Anselme et al., 2013; Boakes, 1977; Collins et al., 1983; Davey et al., 1982; Gibbon et al., 1980; Goodrich, 1959; Gottlieb, 2004, 2006; Robinson et al., 2014b, 2015). If there is a possibility of missing reward delivery (because $0 < p < 1$), incentive hope develops ($\hbar > 0$). Given that $\hbar$ is part of the total 'wanting' process (w_T) that controls the curve's asymptote, the w_T value (and hence the asymptote) reaches a higher level when reinforcement is partial rather than continuous.

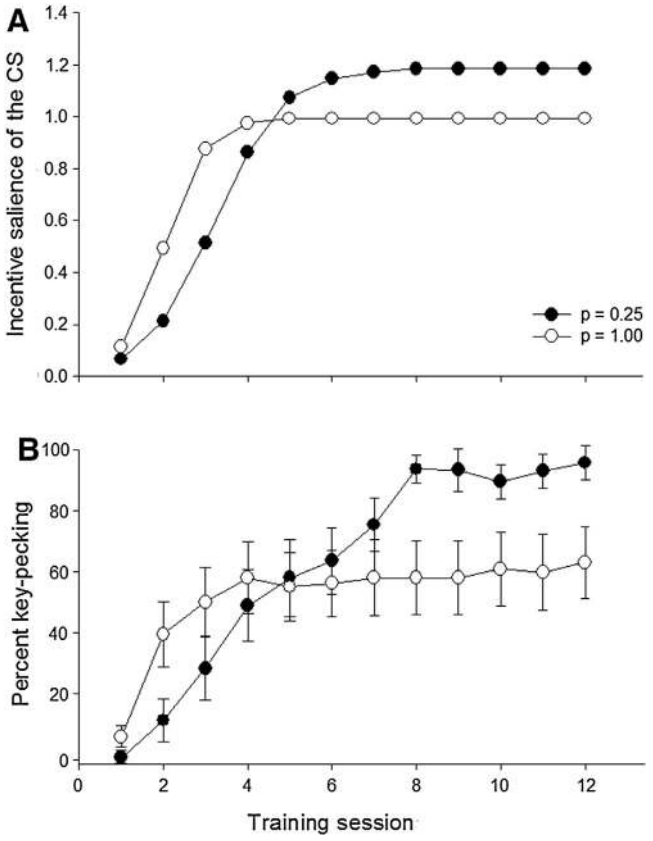

Fig. 4. Basic predictions of the extended incentive hope model. (A) Under PRF, the attribution of incentive salience to a CS is slower (because of reduced δ value) early in training but reaches a higher asymptotic value (because of enhanced w_T value) later in training compared with CRF. In this simulation, $w_N = 1$, $\gamma = 1$, $\alpha = 0$, $k = 2$, and $\theta = 3$. (B) Real percentages of key-pecking responses in pigeons trained at the same reward probabilities ($p = 0.25$ and $p = 1$). Modified from Gottlieb (2004).

- Thirdly, the model correctly predicts decreased approach behaviour during aversive conditioning (Skinner, 1938; Pelchat et al., 1983). Aversive conditioning means that the UCS is 'unwanted' ($w_N < 0$). As a result, there is a gradual loss in the CS attractiveness, which should be more marked under PRF than under CRF over the training sessions. Indeed, the hope that an aversive event (such as a shock) will not be delivered should exist when the event's probability is less than 100%, adding its aversive effect to that of the UCS itself, and should not exist when the event's probability is 100%. Nevertheless, if an action can prevent the aversive event, increased avoidance behaviour should occur (Foree and LoLordo, 1970) because relief is rewarding – in this case, w_N has a positive value and performance cannot be distinguished from that acquired following appetitive conditioning. Acquired impotence often occurs when punishment is unavoidable (Overmier and Seligman, 1967). This situation means that the 'unwanted' UCS ($w_N < 0$) and the 'wanted' relief ($w_N > 0$) may come to conflict, causing $w_T = 0$. Thus, the asymptote equals zero and no action is produced.
- Fourthly, in the case of physiological conditioning (heart rate, salivation, etc.) and of reflex conditioning (eye blink), the model correctly predicts weakened performance under PRF than under CRF (Fitzgerald, 1963; Powell et al., 2005; Reynolds, 1958; Ross, 1959; Sadler, 1968; Thomas and Wagner, 1964; Vardaris and Fitzgerald, 1969). This phenomenon results from the fact that physiological/reflex mechanisms are sensitive to the averaged *physical* salience (intensity, duration), not the incentive salience (motivation), of a UCS. If there is no 'wanting' involved in the control of performance (which only depends on a numerical value,

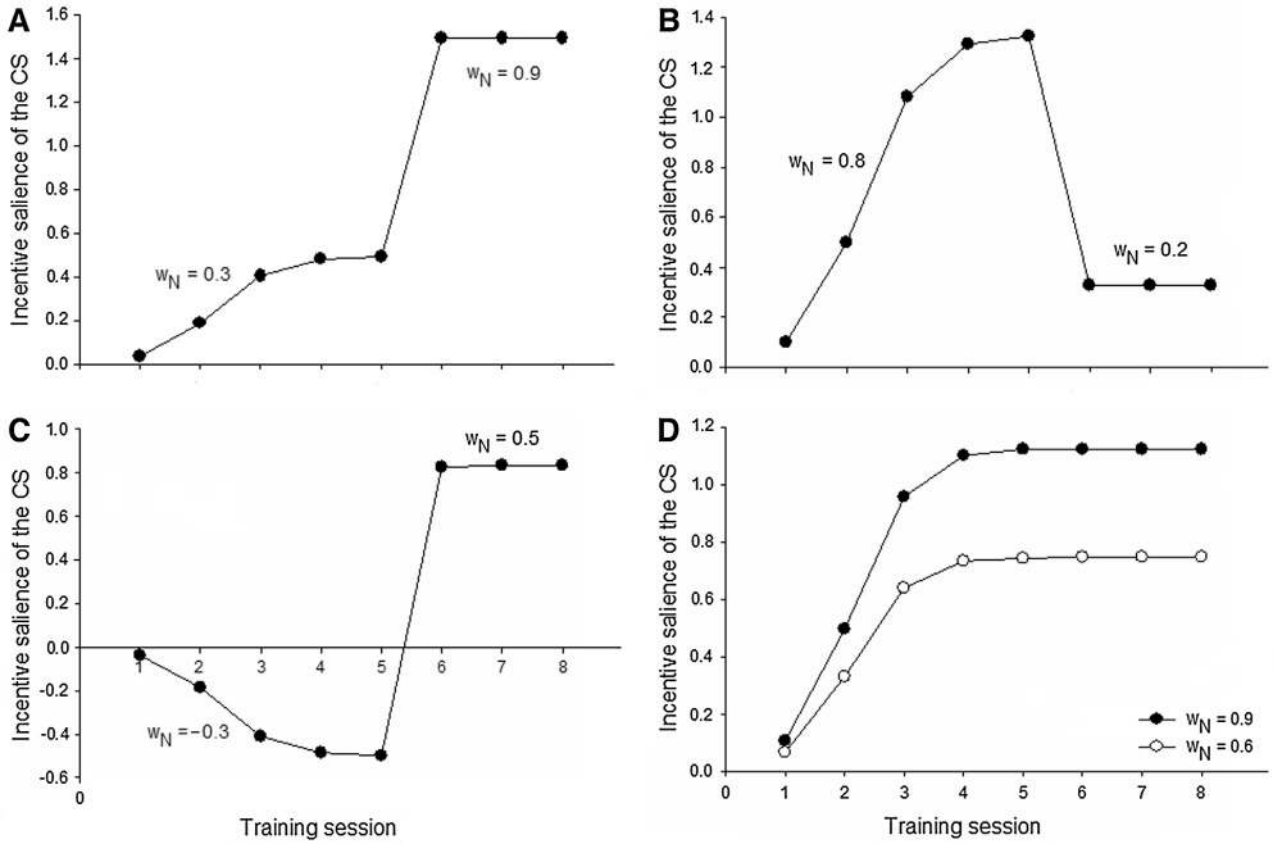

Fig. 5. Effects of deprivation level on performance. (A) In animals trained for five sessions while weakly deprived (w_N = 0.3), there is instant upshift in the incentive salience of the CS following an increase in their deprivation level for the last three sessions (w_N = 0.9). (B) Conversely, animal trained for five sessions while deprived (w_N = 0.8) show instant downshift in the incentive salience of the CS following a decrease in their deprivation level for the last three sessions (w_N = 0.2). (C) In animals trained for five sessions with an aversive CS (w_N = −0.3), there is instant upshift in the incentive salience of the CS following an increase in the CS appetency for the last three sessions (w_N = 0.5), such as reported in Robinson and Berridge (2013). All these results are akin to those obtained with the Zhang et al.'s (2009) model. (D) Animals continuously trained while strongly deprived (w_N = 0.9) show a higher asymptotic attribution of incentive salience to the CS than animals continuously trained while less deprived (w_N = 0.6). These effects should be marked using drugs that modulate dopamine levels (e.g. amphetamine) as well as hormones that represent food depletion/repletion signals (e.g., cholecystokinin). In these simulations, p = 1.0, γ = 0.6, k = 2, α = 0, and θ = 3.

S, describing the averaged physical salience of the UCS), asymptotic performance is necessarily higher in animals trained with CRF (e.g., S = 0.8 when p = 1.0) than in animals trained with PRF (here, S = 0.8/2 = 0.4 when p = 0.5). Note that the extended incentive hope model does not represent the physical salience of a stimulus, except in opposition to its incentive salience. The idea is that incentive salience is often more effective in motivating behaviour than physical salience (e.g., a foraging bird may suddenly fly away because of a slight noise in a bush nearby, while it will remain impassive during a clap of thunder).

7.2. Instant shifts in sign-tracking performance

From this section, I report predictions not discussed beforehand. Increasing or decreasing physiological deprivation without additional training has an immediate effect on sign-tracking performance (Robinson and Berridge, 2013; Tindell et al., 2005, 2009). As seen, this phenomenon is well captured by the Zhang et al.'s (2009) model, although its applicability to related phenomena remains limited.

The extended incentive hope model represents an animal's physiological state as w_N. Any discrepancy in physiological state that occurs between training and test (within subjects) or that characterizes two groups of individuals (between subjects) can be simulated by altering the w_N value. As w_N is involved in the calculation of the curve's asymptote, this has an immediate effect on performance. Fig. 5A and B depict the abrupt increase and decrease, respectively, in the incentive salience attribution to a CS that result from changes in physiological state between training and test. Fig. 5C represents the radical change from aversive to appetitive conditioning caused by the salt experience in non-deprived rats switched to deprivation (Robinson and Berridge, 2013). Fig. 5D shows that animals trained under distinct physiological states (deprived or non-deprived) will reach distinct asymptotic performances. Note that, in its current form, the extended incentive hope model cannot be fully compared with the Zhang et al.'s model, which is a trial-per-trial rather than session-per-session approach to motivational changes in Pavlovian conditioning. Nevertheless, it can reproduce the same effects at the session level.

7.3. Inter-individual differences

In addition to the distinction in the attribution of incentive salience between strong and weak sign-trackers, the γ factor also predicts the existence of variations between these two categories (Fig. 6A). For example, a sign-tracker that performs 120 lever presses during a session should have a higher resistance to motivational transfer than a sign-tracker that performs 300 lever presses during the same session. Ambivalent individuals (which both sign- and goal-track) are individuals with intermediate γ values. At an

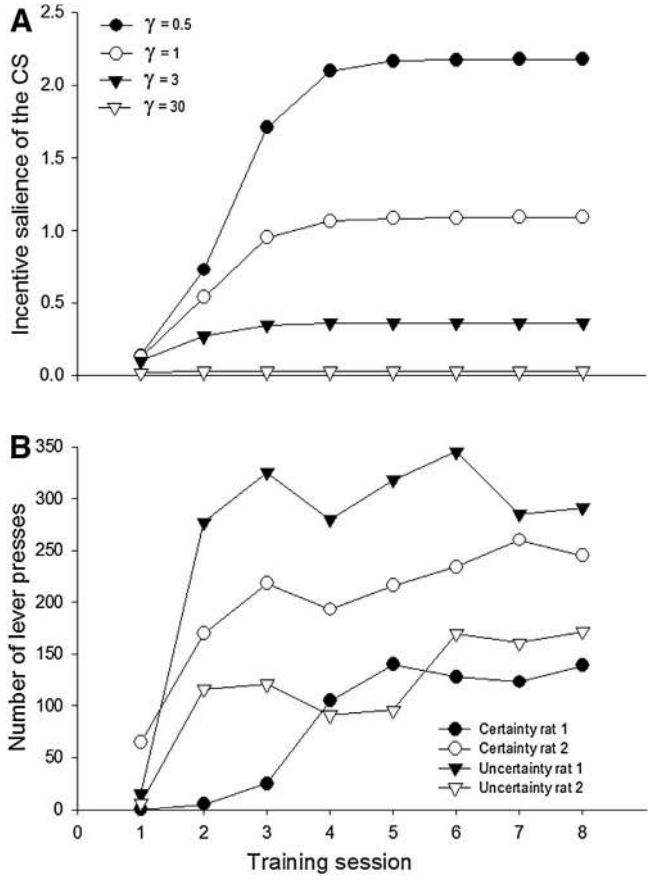

Fig. 6. Effects of γ on the attribution of incentive salience. (A) Changes in the γ value affects the ability to attribute incentive salience to a CS, ranging from pure sign-tracking (low γ) to pure goal-tracking (elevated γ). This parameter does not say anything about goal-tracking (the propensity to inspect the goal dish) in itself, but simply captures the fact that goal-trackers find the CS unattractive. (B) Variability in responding among sign-trackers exposed to certainty and uncertainty training conditions. It can be observed that certainty may lead some individuals to reach higher levels of responding than uncertainty. Part of data set from Anselme et al. (2013, Experiment 1). Note that part A is not a simulation of part B.

individual level, the influence of γ may trump that of $\hbar$ in determining the asymptote. For example, an individual with γ = 0.5 and $\hbar = 0$ (i.e., if $p = 1.0$) will reach a higher asymptotic performance than an individual with γ = 1 and maximal $\hbar = 0.25$ (i.e., if $p = 0.5$). This prediction is compatible with the empirical evidence that some rats trained under CRF may have higher sign-tracking performance than some other rats trained under PRF (Fig. 6B). Note that when two groups (rather than two individuals) are compared, γ is averaged and assumed to be equivalent for each individual, so that any asymptotic discrepancy between the two groups only results from a difference in $\hbar$.

The introduction of γ alters the explanation initially provided with respect to the lack of CS attractiveness in goal-trackers (Anselme, 2015). In accordance with the available data (e.g., Robinson and Flagel, 2009), I suggested that goal-trackers find the CS unattractive because they do not 'want' it. But the absence of 'wanting' was represented by a numerical value that denoted the physical salience of the CS, as in the case of physiological/reflex conditioning. This seemed to justify the apparent evidence that goal-tracking is less pronounced under PRF than under CRF (Gottlieb, 2005). However, can we really compare goal-tracking to a mere eye-blink response? It is likely that goal-trackers are able to 'want' the delivery of pellets, although to a lesser extent than sign-trackers due to reduced dopamine release (see section 4). But, as discussed below, the difference between goal- and sign-trackers may also rely on their ability to transfer that 'wanting' to the CS – as now captured by the γ factor. In the revised model, CS attractiveness does not develop in goal-trackers due to elevated γ value, despite the capacity of experiencing 'wanting' (w_T) relative to the UCS. As a result of an elevated γ value, the asymptotic performance of goal-trackers relative to the CS is (close to) zero. This modification does not alter the predictions drawn from the model's previous version but provides a more realistic description of the psychology of goal-trackers. It is important to note that the extended incentive hope model makes no prediction about goal-tracking responses properly speaking. For example, it does not distinguish between a pure goal-tracker (which never inspects the lever) and an animal totally uninterested in the experimental context (which also never inspects the lever). In short, the model should not be used to account for goal-tracking responses.

The γ factor is necessary to account for one particular fact: increasing dopamine release in goal-trackers does not transform them in sign-trackers but in stronger goal-trackers (DiFeliceantonio and Berridge, 2012). This means that a dopamine-dependent increase in w_T has the effect of increasing the goal-tracker's γ value, while the same increase in w_T should come along with a decrease in the sign-tracker's γ value. As a result, incentive salience attribution to a CS is reduced in the goal-tracker and enhanced in the sign-tracker, despite higher dopamine levels in both phenotypes. Also, very low values for γ (< 1) could be a way for the model to account for addiction, a complex process that makes the attractiveness of CSs so intense and so uncontrollable that behaviour turns pathological – alternatively, the model could also account for phobia and addiction, but not the dopamine-dependent strength of goal-tracking, by defining w_N's domain as $[-\infty, +\infty]$. For example, in crack cocaine addicts, irrational cue attraction sometimes leads people to inspect the floor for white pebbles because they look like crack cocaine. They may put those noncocaine pebbles in a pipe and try to light and smoke them, as if they were the drug reward itself (Rosse et al., 1993). Also note that exciting situations may act as conditioned stimuli that urge pathological gamblers to gamble (Sodano and Wulfert, 2010), even if this activity is likely to cause substantial monetary losses. Addicts in general might be individuals with a profile of sign-trackers (e.g., γ = 1) that have become very strong sign-trackers (γ < 1) due to the influence of stimuli favouring massive dopamine release. How γ could be controlled in the brain is an open question. We can speculate that its value for an individual depends on genetic predispositions that facilitate the transfer of reward properties to contextual cues.

7.4. Stimuli competition

Many models have succeeded in predicting a number of behavioural effects related to stimuli competition with two (or more) CSs, such as blocking and overshadowing. The α factor enables the extended incentive hope model to account for some of those effects as well. Before showing that, however, I would like to mention two traditional interpretations with respect to stimulus competition and to specify what is new with the present model. First, learning-focused models such as the Rescorla-Wagner model suggest that stimulus competition reflects a learning deficit. For example, in blocking (where responding to a CS2 is tested after the CS2 was paired with a previously trained CS1), the attenuated response to the CS2 is assumed to result from the fact that this stimulus brought no new information and was therefore not well learned as a predictor of the UCS. Such a view was challenged by a number of studies showing that conditioned responding to the blocked CS2 can emerge after extinguishing responding to the blocking CS1, indicating that the CS2-UCS association was correctly learned (Arcediano et al., 2001, 2004; Blaisdell et al., 1999). Similar observations were done in other stimulus competition con-

texts (see below). Second, in response to this theoretical issue, a performance-focused model – called comparator hypothesis – has been developed (Miller and Matzel, 1988; Denniston et al., 2001; Stout and Miller, 2007). According to the comparator hypothesis, the blocking stimulus (CS1) and the blocked stimulus (CS2) are both learned, but they compete for expression in behaviour on the basis of their associative strength relative to the UCS. As the strength of the CS1-UCS association is stronger than that of the CS2-UCS association, it impedes more the expression of CS2-directed behaviour than the CS2-UCS association impedes the expression of CS1-directed behaviour. Responding is therefore more pronounced for the CS1 than for the CS2. In line with the comparator hypothesis' prediction (but not the Rescorla-Wagner model's prediction), it was demonstrated that the blocked CS2 reduces the behavioural expression for the blocking CS1: when rats are exposed to the blocking CS1 at test (i.e. after experiencing the compound CS1-CS2), responding is attenuated compared to a control situation in which the CS1 was never presented in compound (Arcediano et al., 2004).

Like the comparator hypothesis, the extended incentive hope model sees stimulus competition as a performance rather than acquisition deficit. But it differs from the comparator hypothesis in that associative strength is here assumed to control the ability to respond to a CS (depending on a fixed decision threshold), not the strength of that response; winning competition is only indirectly related to the associative strength of a CS. Specifically, once an animal has learned a CS-UCS association (i.e. that the decision threshold is overcome), it comes to respond to the CS, and the increase in responding over the training sessions mainly reflects the gradual update (through γ) of motivational processes that control responding in the situation. The motivational salience of the CS stimulus grabs attention, reducing the individual's attentional resources for other stimuli. This reduction of attention for any other stimulus is assumed to be at the origin of blocking – and of any form of stimuli competition. Simply, as the blocked CS2 receives less attention than the blocking CS1, the animal is less motivated to respond to that stimulus. The factor α, mentioned in equations 3 and 4, represents the attentional control of motivational interactions. (On next section, an experiment is proposed to disentangle the extended incentive hope model from the comparator hypothesis with respect to blocking, see Prediction 3.)

Let us briefly introduce the topic of divided attention in animals and humans before showing how α can account for stimulus competition. Significant stimuli (UCSs and CSs) can influence an animal's behaviour because they are the object of the animal's motivation and attention. However, attentional resources are limited; the more attentional resources allocated to a stimulus, the less they can be allocated to another stimulus within the same temporal interval (for reviews of animal and human literature, see Baddeley, 1990; Zentall, 2005). A typical procedure that illustrates the limitation of attentional resources in animals is to present a stimulus, followed by a choice between two comparison stimuli. The initial stimulus is either simple (composed of one feature) or complex (composed of two features) while the comparison stimuli are composed of only one feature. Only one of the two comparison stimuli has the same feature as that previously seen. It is observed that pigeons do fewer comparison errors to process one-featured than two-featured initial stimuli (Maki and Leith, 1973; Zentall et al., 1997). Comparison errors result from the difficulty for pigeons to share their limited attentional capacity between the two dimensions – rather than from memory-retrieval deficits caused by the presence of more than one feature (Zentall et al., 1997). In animals as well as in humans, an extensive training is able to improve performance, but interference can never be totally eliminated (Shallice et al., 1985; Zentall et al., 1997). Several authors have insisted on the role of attention in modelling Pavlovian conditioning. In particular, Mackintosh (1975) suggested that the large amount of attention

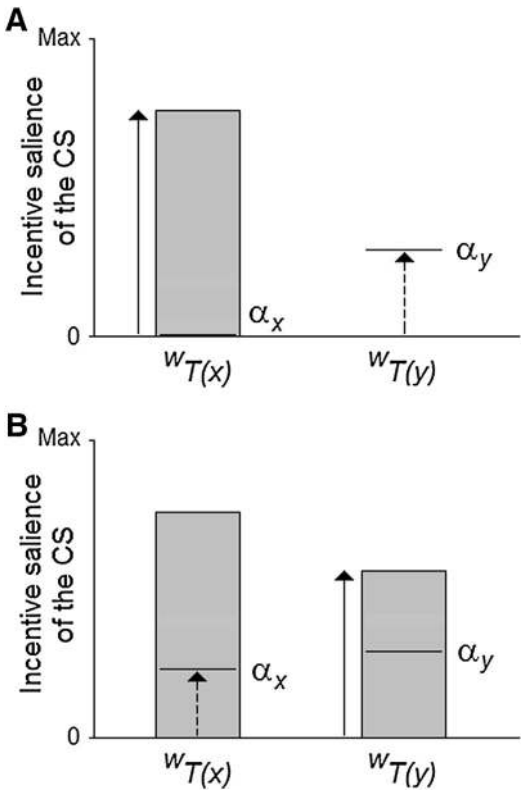

Fig. 7. Attentional control of motivational interactions (Anselme, 2007). (A) The increase in a motivation, $w_{T(x)}$, beyond its attentional threshold ($\alpha_x = 0$) raises the attentional threshold ($\alpha_y > 0$) of any potential concurrent motivation, $w_{T(y)}$. (B) If $w_{T(y)}$ is recruited and that $w_{T(y)} > \alpha_y$, this new motivation comes to conflict with $w_{T(x)}$. In turn, $w_{T(y)}$ tends to inhibit $w_{T(x)}$ by raising α_x.

devoted to a CS with a high predictive accuracy speeds up conditioning. In contrast, Pearce and Hall (1980) proposed that the amount of attention paid to a CS is inversely related to its predictive accuracy, so that an unreliable CS would attract more attention than a reliable CS.

The extended incentive hope model considers attention only when more than one CS is involved. The α ($\in [-\infty, k - (n_T - n_R)/n_T[$) denotes an attentional threshold in reference to a theory suggesting that motivational interactions depend on attentional control (Anselme, 2007). This theory has been shown to account for the abnormal activity patterns of displacement activities (Anselme, 2008) as well as some aspects of the effects of drugs on the processing of non-drug rewards (Anselme, 2009). Briefly, when a motivation, $w_{T(x)}$, is recruited for a stimulus in the absence of any interfering stimulus, the attentional threshold of that motivation is zero ($\alpha_x = 0$). But its recruitment raises the attentional threshold ($\alpha_y > 0$) of any concurrent motivation, $w_{T(y)}$, owing to limited attentional resources – even if no concurrent motivation is actually recruited (Fig. 7A). A motivation must overcome its own attentional threshold to influence behaviour ($w_{T(x)} > \alpha_x$). The enhanced α_y-threshold means that any (potential) concurrent motivation is (partly or totally) inhibited and that it will temporarily be more difficult for it to win the competition. Indeed, if an animal cannot pay much attention to a CS (because already attending to another CS), the incentive salience attribution to the new CS may be altered. However, if the concurrent motivation is sufficiently strong to distract the animal's attention ($w_{T(y)} > \alpha_y$), this will have the consequence of raising the attentional threshold (α_x) of the initial motivation, and hence to reduce its strength (Fig. 7B). For our purpose, the main idea behind these concepts is that, when two (or more) motivations are conflicting, their initial w_T value

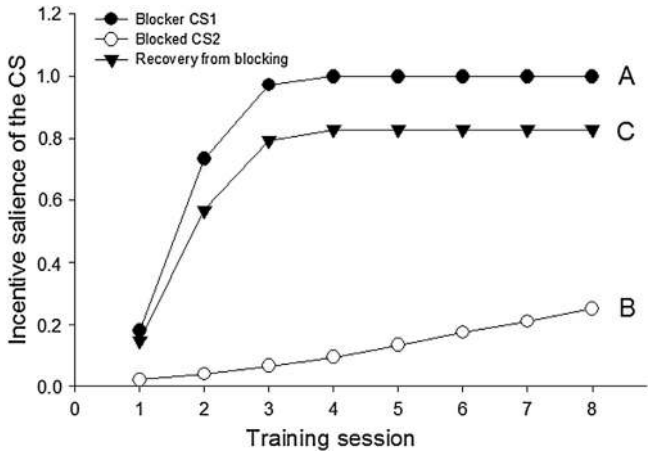

Fig. 8. Computer simulation of blocking. Curve A: Development of (normal) conditioned responding to the blocker CS1 (with $w_N = 1$, $p = 1$, $\gamma = 1$, $\alpha_1 = 0$, $k = 2.5$, and $\theta = 3$). Curve B: Development of conditioned responding to the blocked CS2 (with $\alpha_2 = 2$). Curve C: Recovery from blocking (with $\alpha_2 = 0.2$). The α_2 value (for CS2) is considerably reduced after extinction of the CS2-UCS association but is assumed to remain higher than zero owing to past experience with the blocker CS1. Unblocking is predicted to have similar effects to those of recovery from blocking. Although some of the aforementioned phenomena should logically follow others (e.g., recovery from blocking follows blocking), all the curves are represented on the same graph in order to permit comparisons between the effects they attempt to capture.

can be altered or camouflaged due to the dynamics of attentional thresholds. To understand stimuli competition, only the first part of the process (represented in Fig. 7A) is useful. On this basis, I now discuss four major phenomena related to stimuli competition and show how the α-threshold in the extended incentive hope model can solve them.

1. *Blocking* (Kamin, 1969). This effect occurs when a CS1 (e.g., light) repeatedly paired with a UCS is then part of a compound CS1-CS2 (e.g., light + tone) paired with the same UCS. A separate test for each CS indicates that CS1 produces a strong conditioned response, while CS2 generates no (or a smaller) conditioned response. The extended incentive hope model interprets the lack of responding to CS2 as a motivational phenomenon related to the dynamics of attentional thresholds (Fig. 8). The initial CS1-UCS association induced a motivation to approach or to avoid CS1, and this enhanced the α value of any potential concurrent motivation. Thus, the development (or the expression) of a motivation to approach or to avoid CS2 becomes more difficult. Blocking is assumed to result from a lack of motivation for the CS2 rather than from a learning failure. This view is compatible with the increased responding to the blocked CS2 observed following extinction of the blocker CS1–a phenomenon referred to as recovery from blocking. Several hundreds of extinction trials are often necessary to generate responding to the CS2, but the ability to recover from blocking is a clear indication that the CS2-UCS association was learned (Blaisdell et al., 1999). Recovery from blocking suggests that a CS2-UCS has been learned, and makes room for the possibility that it can be expressed in behaviour once the motivation for CS1 (at the origin of blocking) reduced after an extinction test – i.e., extinguishing the motivation for CS1 decreases the α-threshold value that camouflaged the motivation for CS2. The extended incentive hope model can also account for upward and downward unblocking (Holland, 1984; Khallad and Moore, 1996). Unblocking means that increased responding to the blocked CS2 is observed following sudden increase or decrease in magnitude of the UCS. This phenomenon results from a decrease in the α value, because the suddenness of the change in reward magnitude comes to refocus attention on CS2–as if the conditioning context was new. The extended incentive hope model predicts that upward unblocking and downward unblocking have similar effects on performance in autoshaping (because they reduce the α threshold in a similar way), a finding empirically demonstrated (Holland, 1984; Khallad and Moore, 1996).
2. *Overshadowing* (Pavlov, 1927). It is observed with a compound CS1-CS2 (e.g., light + tone) repeatedly paired with a UCS. When the elements of the compound are tested separately, one stimulus often produces stronger conditioned responding than the other, which is said to have been overshadowed during conditioning. A stimulus can overshadow another provided that it is physically more salient or that the animal is, for phylogenetic or ontogenetic reasons, more sensitive to that stimulus. As in the case of blocking, the overshadowed stimulus is traditionally assumed to be unlearned. The extended incentive hope model predicts that, if CS1 is motivationally more attractive than CS2, it will raise the attentional threshold of CS2. The development (or the expression) of a motivation to approach or avoid CS2 is then more difficult. Recovery from the overshadowed CS2 occurs following an extinction of responding to CS1, suggesting that a CS2-UCS was learned but could not express in behaviour (Matzel et al., 1985). As in the case of blocking, extinguishing the motivation for CS1 lowers the α-threshold value, allowing the motivation for CS2 to express in behaviour.
3. *Overexpectation* (Kremer, 1978). When two CSs are separately paired with a UCS, and then, become the elements of a compound CS1-CS2 paired with the same UCS, stronger responding occurs to the compound than to each individual CS. Then, when the CSs are tested separately, responding to each of them is lower than it was initially. Contrary to blocking and overshadowing, the animal is not initially trained with CS1 and CS2 occurring at the same time. But the association of the two stimuli (compound), as well as their subsequent testing following exposure to the compound, also generates some attentional conflicts. The incentive hope model predicts that a motivation should develop for each CS at training. When the two CSs are then combined to form a compound, there is a summation of their motivational salience ($w_{T(1)} + w_{T(2)}$), which increases asymptotic performance, and also an enhancement of the attentional threshold for each CS ($\alpha_1 > 0$ and $\alpha_2 > 0$). As a result, conditioned responding to the compound is higher than it was initially to each CS, but it should remain smaller than twice as much to the compound as to its elements (Fig. 9). At test, each CS generates a lower level of responding than initially because of its previous association with another significant CS. Concretely, this means that the CS's attentional threshold is reduced (because the other CS is no longer presented) but remains higher than zero (because the associative history of the presented CS distracts the animal).
4. *Superconditioning* (Rescorla, 1971). A CS is repeatedly paired with food (CS +), and then becomes part of a compound that predicts no food, so that the second CS acquires inhibitory properties (CS–). At test, the CS+ generates stronger conditioned responding when it was previously paired with the CS–than when it was presented alone. The extended incentive hope model suggests that the attribution of incentive salience to the CS–is nil because there is no motivational transfer to a CS without UCS (Eq. (3)'s numerator equals zero). However, the CS–is not just a neutral stimulus, in which case conditioned responding should be independent of the history of the CS+ (Williams and McDevitt, 2002). I postulate that the allocation of attention to conditioned inhibitors is redirected to the CS+. This surge of attentional focus is not incentive hope (because there is no uncertainty), but the behavioural effect may be similar (enhanced responding).The inhibitory properties of the CS–could be represented by a decrease of the attentional threshold of concurrent motivations below its neutrality level

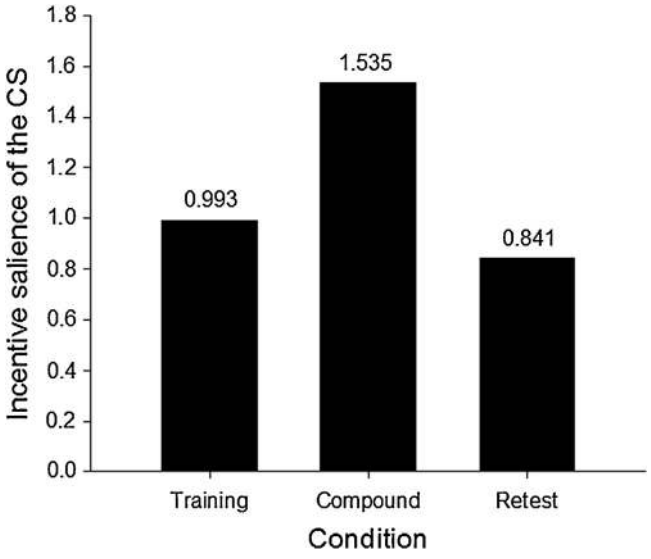

Fig. 9. Computer simulation of overexpectation. The histograms represent the asymptotic performance during the different experimental phases. Training denotes initial conditioned responding to each CS (responses are here supposed to be equivalent for the two CSs since they are associated with the same UCS). For this simulation, I used $w_N = 1$, $p = 1$, $\gamma = 1$, $\alpha = 0$, $k = 2$, and $\theta = 3$. Compound represents conditioned responding to CS1 and CS2 presented simultaneously ($w_{N(1)} + w_{N(2)} = 2$ and $\alpha = 0.3$). Retest is conditioned responding to each CS alone, after exposure to the compound ($w_N = 1$ and $\alpha = 0.18$).

(i.e., $\alpha < 0$ rather than $\alpha = 0$). This means that the motivation for the CS+ in the presence of the CS– overcomes its attentional threshold ($\alpha < 0$) to a larger extent than when the CS+ is not associated with the CS– ($\alpha = 0$). Superconditioning is a result of this change in the α value (Fig. 10).

The above-mentioned interpretations remain essentially theoretical and it would be useful to find out appropriate tests for the proposed mechanisms. It is also unclear whether the present model could account for more Pavlovian phenomena related to stimuli competition. But those analyses demonstrate that Pavlovian effects traditionally viewed as pure products of associative learning can be appropriately characterised using a motivational model.

7.5. Some original predictions

In addition to reproducing a number of well-studied phenomena, the extended incentive hope model can make original predictions. Four of them are presented here.

7.6. Prediction 1

The extended incentive hope model predicts that an alternation of cued rewards and nonrewards (e.g., 0-2-0-2-0-. . . pellets) should generate lower asymptotic performance than a random mixture of cued rewards and nonrewards (e.g., 2-0-0-0-2-0-2-2-. . . pellets). The reason is that alternated rewards are fully predictable ($p = 1.0$, $\hbar = 0$), while random rewards are not on a given trial ($p = 0.5$, $\hbar$ is maximal). This corroborates the evidence that when animals can anticipate the next trial, the anticipation of nonreward lowers performance compared to the anticipation of reward (Capaldi and Stanley, 1963; Couvillon et al., 1980; Gonzalez et al., 1966; Wall and Goodrich, 1964).

7.7. Prediction 2

In a blocking paradigm, it is predicted that strong sign-trackers will exhibit stronger conditioned suppression to a blocked CS2 than weaker sign-trackers, provided that none of these individuals produce goal-tracking responses. The reason is that strong sign-trackers should pay more attention to the blocking CS1 than weaker sign-trackers, raising the attentional threshold for CS2 at a higher level. As a result, strong sign-trackers should be less motivated to respond to CS2. In contrast, if the weaker sign-trackers are, in fact, ambivalent responders (both sign- and goal-tracking), the attention paid to CS1 *and* to the food dish should also raise the attentional threshold for CS2, causing a more pronounced conditioned suppression to the blocked CS2 in those animals – compared to weak sign-trackers that do not goal-track.

7.8. Prediction 3

If a blocking CS1 is an uncertain predictor of food, the extended incentive hope model predicts that it should block responding to CS2 more than if it is a certain predictor of food. Indeed, although an unreliable CS1-UCS association produces a weaker associative strength, the uncertain CS1 should give rise to a higher motivational

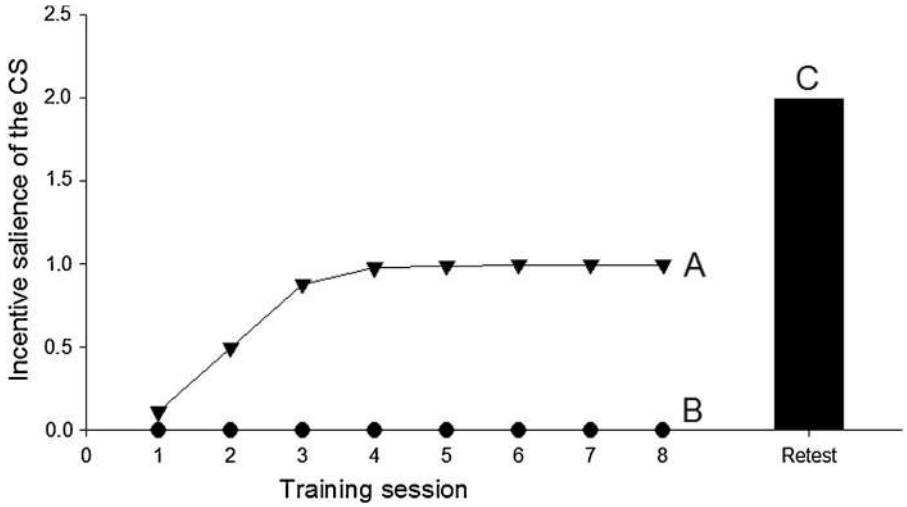

Fig. 10. Computer simulation of superconditioning. Curve A: Development of conditioned responding to an excitatory CS (CS+) presented alone. Curve B: Absence of conditioned responding to an inhibitory CS (CS–) presented in compound with the CS+. C (retest): Supernormal conditioned responding to the CS+ after training with the compound. For the superconditioned response: $w_N = 1$, $p = 1$, $\gamma = 1$, $\alpha = -0.5$, $k = 3$, and $\theta = 3$.

salience ($\hbar \neq 0$), and therefore raise the attentional threshold for CS2 more than a reliable CS1-UCS association. Less attention and less motivation are then allocated to CS2. This prediction contrasts with that of the comparator hypothesis, which would suggest that the blocking of responding to CS2 must be more pronounced when CS1 is a certain (reliable) predictor of food. Indeed, as the certainty of the CS1-UCS association results in a stronger associative strength, a greater impediment of the expression of a response to CS2 is expected.

7.9. Prediction 4

The extended incentive hope model predicts that, in a natural context, individuals exposed to a random distribution of food (because of winter conditions, because they are subordinates, or because they are poor foragers) should have greater fat reserves than animals exposed to a more predictable distribution of food. Indeed, if uncertainty motivates food-seeking behaviour, individuals experiencing uncertainty should consume more food items, and hence become fatter, than individuals experiencing food certainty. This prediction is confirmed by dozens of field studies, especially in small birds such as starlings, titmice, and blackbirds (e.g., Bauer et al., 2011; Cresswell, 2003; Cuthill et al., 1997; Gosler, 1996; Hake, 1996; MacLeod et al., 2007; Polo and Bautista, 2006; Pravosudov and Grubb, 1997; Ratikainen and Wright, 2013; Witter and Swaddle, 1995). Although some studies report a mere correlation between environmental unpredictability and fat reserves, others have explicitly shown that unpredictability promotes food seeking and consumption (e.g., Bauer et al., 2011; Haftorn, 1976; King and Farner, 1966; Pravosudov and Grubb, 1997). The interpretation suggested by the extended incentive hope model does not contradict the functional explanation proposed by behavioural ecologists that, when food is scarce, animals are less sure of meeting their daily energy budget, leading them to seek food more avidly. In contrast, the model provides a plausible psychological mechanism showing how this may happen.

8. Conclusion

The goal of the extended incentive hope model is twofold: explaining the effects of reward uncertainty on sign-tracking and capturing the well-documented evidence that sign-tracking performance mainly reflects CS-triggered motivation (rather than associative learning). This model is in keeping with the logic behind the Zhang et al.'s (2009) equation, although it offers a broader conceptual framework capable of tackling a wide range of Pavlovian phenomena. Nevertheless, the extended incentive hope model has limits that need to be overcome. First, the model makes no prediction with respect to the role of perception and memory in conditioning – and hence, does not account for a number of phenomena such as stimulus discrimination, stimulus generalisation, recovery from extinction, latent inhibition, etc. Second, the model does not explain the motivational dynamics that operates from trial to trial, so that the details of the interactions between several variables are not specified (and they are currently not known). A general model explaining and predicting motivational changes at the trial level is to be developed. Third, some of the model's variables are difficult to quantify a priori (w_N, α, γ, k). For some parameters (especially α and γ), one potential problem is that we do not know whether the exact same neurobiological determinants control them in distinct species. With respect to k, it might depend on traditional learning algorithms relative to a decision threshold, but it remains largely unspecified. Nevertheless, as shown throughout the paper, having exact values is unnecessary to enable the formulation of correct and original predictions.

Acknowledgments

This work was partly supported by a DFG (German Research Foundation) fellowship (AN-1067/1-1).

References

Amsel, A., Hancock, W., 1957. Motivational properties of frustration: III: Relation of frustration effect to antedating goal factors. J. Exp. Psychol. 53, 126–131.
Amsel, A., Roussel, J., 1952. Motivational properties of frustration: I: Effect on a running response of the addition of frustration to the motivational complex. J. Exp. Psychol. 43, 363–368.
Amsel, A., Ward, J.S., 1954. Motivational properties of frustration: II: Frustration drive stimulus and frustration reduction in selective learning. J. Exp. Psychol. 48, 37–47.
Amsel, A., 1958. The role of frustrative nonreward in noncontinuous reward situations. Psychol. Bull. 55, 102–119.
Amsel, A., 1992. Frustration Theory. Cambridge University Press, Cambridge.
Anselme, P., Robinson, M.J.F., 2016. 'Wanting', 'liking', and their relation to consciousness. J. Exp. Psychol. Anim. Learn. Cogn. (in press).
Anselme, P., Robinson, M.J.F., Berridge, K.C., 2013. Reward uncertainty enhances incentive salience attribution as sign-tracking. Behav. Brain Res 238, 53–61.
Anselme, P., 2007. Some conceptual problems with the classical theory of behaviour. Behav. Process. 75, 259–275.
Anselme, P., 2008. Abnormal patterns of displacement activities: a review and reinterpretation. Behav. Process 79, 48–58.
Anselme, P., 2009. The effect of drugs on the processing of natural rewards. Neurosci. Biobehav. Rev. 33, 314–335.
Anselme, P., 2013. Dopamine motivation, and the evolutionary significance of gambling-like behaviour. Behav. Brain Res. 256, 1–4.
Anselme, P., 2015. Incentive salience attribution under reward uncertainty: a Pavlovian model. Behav. Process. 111, 6–18.
Arcediano, F., Escobar, M., Matute, H., 2001. Reversal from blocking in humans as a result of posttraining extinction of the blocking stimulus. Anim. Learn. Behav. 29, 354–366.
Arcediano, F., Escobar, M., Miller, R.R., 2004. Is stimulus competition an acquisition deficit or a performance deficit? Psychon. Bull. Rev. 11, 1105–1110.
Baddeley, A., 1990. Human Memory: Theory and Practice. Lawrence Erlbaum Associates.
Balleine, B., 1992. Instrumental performance following a shift in primary motivation depends on incentive learning. J. Exp. Psychol. Anim. Behav. Process. 18, 236–250.
Barela, P.B., 1999. Theoretical mechanisms underlying the trial-spacing effect in Pavlovian fear conditioning. J. Exp. Psychol. Anim. Behav. Process. 25, 177–193.
Bauer, C.M., Glassman, L.W., Cyr, N.E., Romero, L.M., 2011. Effects of predictable and unpredictable food restriction on the stress response in molting and non-molting European starlings (*Sturnus vulgaris*). Comp. Biochem. Physiol. A 160, 390–399.
Berridge, K.C., Robinson, T.E., 1998. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Res. Rev. 28, 309–369.
Berridge, K.C., 2012. From prediction error to incentive salience: mesolimbic computation of reward motivation. Eur. J. Neurosci 35, 1124–1143.
Blaisdell, A.P., Gunther, U., Miller, R.R., 1999. Recovery from blocking achieved by extinguishing the blocking CS. Anim. Learn. Behav. 27, 63–76.
Blaiss, C.A., Janak, P.H., 2009. The nucleus accumbens core and shell are critical for the expression but not the consolidation, of Pavlovian conditioned approach. Behav. Brain Res. 200, 22–32.
Boakes, R.A., 1977. Performance on learning to associate a stimulus with positive reinforcement. In: Davis, H., Hurvitz, H.M.B. (Eds.), Operant Pavlovian Interactions. Erlbaum Associates, Hillsdale, NJ, pp. 67–97.
Bouton, M.E., Sunsay, C., 2003. Importance of trials versus accumulating time across trials in partially reinforced appetitive conditioning. J. Exp. Psychol. Anim. Behav. Process. 29, 62–77.
Bowman, E.M., Brown, V.J., 1998. Effects of excitotoxic lesions of the rat ventral striatum on the perception of reward cost. Exp. Brain Res. 123, 439–448.
Breen, R.B., Zuckerman, M., 1999. 'Chasing' in gambling behavior: personality and cognitive determinants. Person. Indiv. Diff. 27, 1097–1111.
Brunborg, G.S., Johnsen, B.J., Pallesen, S., Molde, H., Mentzoni, R.A., Myrseth, H., 2010. The relationship between aversive conditioning and risk-avoidance in gambling. J. Gambl. Studies 26, 545–559.
Cagniard, B., Balsam, P.D., Brunner, D., Zhuang, X., 2006a. Mice with chronically elevated dopamine exhibit enhanced motivation but not learning, for food reward. Neuropsychopharmacology 31, 1362–1370.
Cagniard, B., Beeler, J.A., Britt, J.P., McGehee, D.S., Marinelli, M., Zuhang, X., 2006b. Dopamine scales performance in the absence of new learning. Neuron 51, 541–547.
Cannon, C.M., Bseikri, M.R., 2004. Is dopamine required for natural reward? Physiol. Behav 81, 741–748.
Capaldi, E.J., Stanley, L.R., 1963. Temporal properties of reinforcement aftereffects. J. Exp. Psychol. 65, 169–175.
Capaldi, E.J., 1967. A sequential theory of instrumental training. In: Spence, K.W., Spence, J.T. (Eds.), The Psychology of Learning and Motivation. Academic Press, New York, pp. 67–156.

Chappell, D., Geis, G., Schafer, S., Siegel, L., 1971. Forcible rape: a comparative study of offenses known to the police in Boston and Los Angeles. In: Henslin, J.M. (Ed.), Studies in the Sociology of Sex. Appleton-Century-Crofts, New York, pp. 169–190.
Collins, L., Pearce, J.M., 1985. Predictive accuracy and the effects of partial reinforcement on serial autoshaping. J. Exp. Psychol. Anim. Behav. Process. 11, 548–564.
Collins, L., Young, D.B., Davies, K., Pearce, J.M., 1983. The influence of partial reinforcement on serial autoshaping with pigeons. Q. J. Exp. Psychol. 35B, 275–290.
Couvillon, P.A., Brandon, S.E., Woodard, W.T., Bitterman, M.E., 1980. Performance of pigeons in patterned sequences of rewarded and nonrewarded trials. J. Exp. Psychol. Anim. Behav. Process. 6, 137–154.
Crawford, L.L., Steirn, J.N., Pavlik, W.B., 1985. Within- and between-subjects partial reinforcement effects with an autoshaped response using Japanese quail (*Coturnix coturnix japonica*). Anim. Learn. Behav 13, 85–92.
Cresswell, W., 2003. Testing the mass-dependent predation hypothesis: in European blackbirds poor foragers have higher overwinter body reserves. Anim. Behav. 65, 1035–1044.
Cuthill, I.C., Hunt, S., Cleary, C., Clark, C., 1997. Colour bands dominance, and body mass regulation in male zebra finches (*Taeniopygia guttata*). Proc. R. Soc. Lond. 264, 1093–1099.
Davey, G.C., Phillips, S., Cleland, G.G., 1981. The topography of signal-centered behaviour in the rats: the effects of liquid and solid food reinforcers. Behav. Anal. Letters 1, 331–337.
Davey, G.C.L., Cleland, G.G., Oakley, D.A., 1982. Applying Konorski's model of classical conditioning to signal-centered behavior in the rat: some functional similarities between hunger CRs and sign-tracking. Anim. Learn. Behav 10, 257–262.
Davis, M., 1992. The role of the amygdala in fear and anxiety. Annu. Rev. Neurosci. 15, 353–375.
Day, J.J., Wheeler, R.A., Roitman, M.F., Carelli, R.M., 2006. Nucleus accumbens neurons encode Pavlovian approach behaviors: evidence from an autoshaping paradigm. Eur. J. Neurosci. 23, 1341–1351.
Dayan, P., Balleine, B.W., 2002. Reward motivation, and reinforcement learning. Neuron 36, 285–298.
Dayan, P., Berridge, K.C., 2014. Model-based and model-free Pavlovian reward learning: revaluation, revision, and revelation. Cogn. Affect. Behav. Neurosci. 14, 473–492.
de Lafuente, V., Romo, R., 2011. Dopamine neurons code subjective sensory experience and uncertainty of perceptual decisions. Proc. Natl. Acad. Sci. USA 108, 19767–19771.
Denniston, J.C., Savastano, H.I., Miller, R.R., 2001. The extended comparator hypothesis: learning by contiguity, responding by relative strength. In: Mowrer, R.R., Klein, S.B. (Eds.), Handbook of Contemporary Learning. Erlbaum, Mahwah, NJ, pp. 65–117.
DiFeliceantonio, A.G., Berridge, K.C., 2012. Which cue to 'want'? Opioid stimulation of central amygdala makes goal-trackers show stronger goal-tracking, just as sign-trackers show stronger sign-tracking. Behav. Brain Res. 230, 399–408.
Dickinson, A., Balleine, B.W., 2002. The role of learning in the operation of motivational systems. In: Pashler, H., Gallistel, C.R. (Eds.), Learning, Motivation and Emotion, Vol. 3 of Steven's Handbook of Experimental Psychology. John Wiley & Sons, New York, pp. 497–533.
Dickinson, A., Dawson, G.R., 1987. The role of the instrumental contingency in the motivational control of performance. Q. J. Exp. Psychol., 77–93.
Dickinson, A., Pearce, J.M., 1976. Preference and response suppression under different correlations between shock and a positive reinforcer in rats. Learn. Motiv. 7, 66–85.
Dickinson, A., Smith, J., Mirenowicz, J., 2000. Dissociation of Pavlovian and instrumental incentive learning under dopamine antagonists. Behav. Neurosci. 114, 468–483.
Dodd, M.L., Klos, K.J., Bower, J.H., Geda, Y.E., Josephs, K.A., Ahlskog, J.E., 2005. Pathological gambling caused by drugs used to treat Parkinson disease. Arch. Neurol. 62, 1377–1381.
Doremus-Fitzwater, T.L., Spear, L.P., 2011. Amphetamine-induced incentive sensitization of sign-tracking behavior in adolescent and adult female rats. Behav. Neurosci 125, 661–667.
Dreher, J.-C., Kohn, P., Berman, K.F., 2006. Neural coding of distinct statistical properties of reward information in humans. Cereb. Cortex 16, 561–573.
Fiorillo, C.D., Tobler, P.N., Schultz, W., 2003. Discrete coding of reward probability and uncertainty by dopamine neurons. Science 299, 1898–1902.
Fitzgerald, R.D., 1963. Effects of partial reinforcement with acid on the classically conditioned salivary response in dogs. J. Comp. Physiol. Psychol. 56, 1056–1060.
Flagel, S.B., Watson, S.J., Robinson, T.E., Akil, H., 2007. Individual differences in the propensity to approach signals vs goals promote different adaptations in the dopamine system of rats. Psychopharmac 191, 599–607.
Flagel, S.B., Clark, J.J., Robinson, T.E., Mayo, L., Czuj, A., Willuhn, I., Akers, C.A., Clinton, S.M., Phillips, P.E.M., Akil, H., 2011a. A selective role for dopamine in stimulus-reward learning. Nature 469, 53–57.
Flagel, S.B., Cameron, C.M., Pickup, K.N., Watson, S.J., Akil, H., Robinson, T.E., 2011b. A food predictive cue must be attributed with incentive salience for it to induce c-Fos mRNA expression in cortico-striatal-thalamic brain regions. Neuroscience 196, 80–96.
Flaherty, C.F., 1996. Incentive Relativity. Cambridge University Press, Cambridge.
Foree, D.D., LoLordo, V.M., 1970. Signalled and unsignalled free-operant avoidance in the pigeon. J. Exp. Anal. Behav. 13, 283–290.
Gallistel, C.R., Gibbon, J., 2000. Time, rate, and conditioning. Psychol. Rev. 107, 289–344.
Gallistel, C.R., 2003. Conditioning from an information processing perspective. Behav. Process. 62, 89–101.
Genn, R.F., Ahn, S., Phillips, A.G., 2004. Attenuated dopamine efflux in the rat nucleus accumbens during successive negative contrast. Behav. Neurosci. 118, 869–873.
Gibbon, J., Balsam, P., 1981. Spreading associations in time. In: Locurto, C.M., Terrace, H.S., Gibbon, J. (Eds.), Autoshaping and Conditioning Theory. Academic Press, New York, pp. 219–253.
Gibbon, J., Farrell, L., Locurto, C.M., Duncan, H.J., Terrace, H.S., 1980. Partial reinforcement in autoshaping with pigeons. Anim. Learn. Behav. 8, 45–59.
Gibbon, J., 1977. Scalar expectancy theory and Weber's law in animal timing. Psychol. Rev 84, 279–325.
Gonzalez, R.C., Bitterman, M.E., 1969. Spaced-trials partial reinforcement effect as a function of contrast. J. Comp. Physiol. Psychol. 67, 94–103.
Gonzalez, R.C., Bainbridge, P., Bitterman, M.E., 1966. Discrete trials lever pressing in the rat as a function of pattern reinforcement, effortfulness of response, and amount of reward. J. Comp. Physiol. Psychol. 61, 110–122.
Gonzalez, F.A., 1973. Effects of partial reinforcement (25%) in an autoshaping procedure. Bull. Psychon. Soc. 2, 299–301.
Goodrich, K.P., 1959. Performance in different segments of an instrumental response chain as a function of reinforcement schedule. J. Exp. Psychol. 57, 57–63.
Gosler, A.G., 1996. Environmental and social determinants of winter fat storage in the great tit *Parus major*. J. Anim. Ecol. 65, 1–17.
Gottlieb, D.A., 2004. Acquisition with partial and continuous reinforcement in pigeon autoshaping. Learn. Behav. 32, 321–334.
Gottlieb, D.A., 2005. Acquisition with partial and continuous reinforcement in rat magazine approach. J. Exp. Psychol. Anim. Behav. Process. 31, 319–333.
Gottlieb, D.A., 2006. Effects of reinforcement and time between reinforced trials on terminal response rate in pigeon autoshaping. Behav. Process. 72, 6–13.
Gottlieb, D.A., 2008. Is the number of trials a primary determinant of conditioned responding? J. Exp. Psychol. Anim. Behav. Process. 34, 185–201.
Haftorn, S., 1976. Variation in body weight: wing length and tail length in the great tit *Parus major*. Nor. J. Zool. 4, 241–271.
Hake, M., 1996. Fattening strategies in dominance-structured greenfinch (*Carduelis chloris*) flocks in winter. Behav. Ecol. Sociobiol 39, 71–76.
Hart, A.S., Clark, J.J., Phillips, P.E.M., 2015. Dynamic shaping of dopamine signals during probabilistic Pavlovian conditioning. Neurobiol. Learn. Mem. 117, 84–92.
Henke, P.G., 1977. Dissociation of the frustration effect and the partial reinforcement extinction effect after limbic lesions in rats. J. Comp. Physiol. Psychol. 91, 1032–1038.
Holland, P.C., 1984. Unblocking in Pavlovian appetitive conditioning. J. Exp. Psychol. Anim. Behav. Process. 10, 476–497.
Howe, M.W., Tierney, P.L., Sandberg, S.G., Phillips, P.E.M., Graybiel, A.M., 2013. Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. Nature 500, 575–579.
Hug, J.J., Amsel, A., 1969. Frustration theory and partial reinforcement effects: the acquisition-extinction paradox. Psychol. Rev. 76, 419–421.
Hull, C.L., 1943. Principles of Behavior. Appleton-Century, New York.
Kakade, S., Dayan, P., 2002. Acquisition and extinction in autoshaping. Psychol. Rev. 109, 533–544.
Kamin, L.J., 1969. Predictability, surprise, attention, and conditioning. In: Campbell, B.A., Church, R.M. (Eds.), Punishment and Aversive Behavior. Appleton-Century-Crofts, New York.
Kaveri, S., Nakahara, H., 2014. Dual reward prediction components yield Pavlovian sign- and goal-tracking. PLoS One 9, e108142.
Khallad, Y., Moore, J., 1996. Blocking, unblocking, and overexpectation in autoshaping with pigeons. J. Exp. Anal. Behav. 65, 575–591.
King, J.R., Farner, D.S., 1966. The adaptive role of winter fattening in the white crowned sparrow with comments on its regulation. Am. Nat. 100, 403–418.
Kremer, E.F., 1978. The Rescorla-Wagner model: losses in associative strength in compound conditioned stimuli. J. Exp. Psychol. Anim. Behav. Process. 4, 22–36.
Laude, J.R., Stagner, J.P., Zentall, T.R., 2014. Suboptimal choice by pigeons may result from the diminishing effect of nonreinforcement. J. Exp. Psychol. Anim. Learn. Cogn. 40, 12–21.
Lesaint, F., Sigaud, O., Flagel, S.B., Robinson, T.E., Khamassi, M., 2014. Modelling individual differences in the form of Pavlovian conditioned approach responses: a dual learning systems approach with factored representations. PLoS Comput. Biol. 10, e1003466.
Leszczuk, M.H., Flaherty, C.F., 2000. Lesions of the nucleus accumbens reduce instrumental but not consummatory negative contrast in rats. Behav. Brain Res. 116, 61–79.
Linnet, J., Mouridsen, K., Peterson, E., Møller, A., Doudet, D.J., Gjedde, A., 2012. Striatal dopamine release codes uncertainty in pathological gambling. Psychiatry Res. 204, 55–60.
MacLeod, R., Lind, J., Clark, J., Cresswell, W., 2007. Mass regulation in response to predation risk indicate population declines. Ecol. Lett. 10, 945–955.
Mackintosh, N.J., 1975. A theory of attention: variations in the associability of stimuli with reinforcement. Psychol. Rev 82, 276–298.

Mahler, S.V., Berridge, K.C., 2009. Which cue to 'want'? Central amygdala opioid activation enhances and focuses incentive salience on a prepotent reward cue. J. Neurosci. 29, 6500–6513.
Maki, W.S., Leith, C.R., 1973. Shared attention in pigeons. J. Exp. Anal. Behav. 19, 345–349.
Matzel, L.D., Schachtman, T.R., Miller, R.R., 1985. Recovery of an overshadowed association achieved by extinction of the overshadowing stimulus. Learn. Motiv. 16, 398–412.
McCain, G., McVean, G., 1967. Effects of prior reinforcement or nonreinforcement on later performance in a double alley. J. Exp. Psychol. 73, 620–627.
McCain, G., 1968. The partial reinforcement effect after minimal acquisition: single pellet reward. Psychon. Sci 13, 151–152.
McClure, S.M., Daw, N.D., Montague, P.R., 2003. A computational substrate for incentive salience. Trends in Neurosci. 26, 423–428.
Meyer, P.J., Lovic, V., Saunders, B.T., Yager, L.M., Flagel, S.B., Morrow, J.D., Robinson, T.E., 2012. Quantifying individual variation in the propensity to attribute incentive salience to reward cues. PLoS One 7, e38987.
Meyer, P.J., Cogan, E.S., Robinson, T.E., 2014. The form of a conditioned stimulus can influence the degree to which it acquires incentive motivational properties. PLoS One 9, e98163.
Mikulka, P.J., Lehr, R., Pavlik, W.B., 1967. Effect of reinforcement schedules on reward shifts. J. Exp. Psychol. 74, 57–61.
Miller, R.R., Matzel, L.D., 1988. The comparator hypothesis: a response rule for the expression of associations. In: Bower, G.H. (Ed.), The Psychology of Learning and Motivation, 22. Academic Press, San Diego, pp. 51–92.
Mirenowicz, J., Schultz, W., 1994. Importance of unpredictability for reward responses in primate dopamine neurons. J. Neurophysiol. 72, 1024–1027.
Mowrer, O.H., 1966. Integrity therapy A self-help approach. Psychother: Theory, Res. Pract. 3, 114–119.
Overmier, J.B., Seligman, M.E.P., 1967. Effects of inescapable shock upon subsequent escape and avoidance responding. J. Comp. Physiol. Psychol. 63, 28–33.
Paolone, G., Angelakos, C.C., Meyer, P.J., Robinson, T.E., Sarter, M., 2013. Cholinergic control over attention in rats prone to attribute incentive salience to reward cues. J. Neurosci. 33, 8321–8335.
Papini, M.R., Dudley, R.T., 1993. Effects of the number of trials per session on autoshaping in rats. Learn. Motiv. 24, 175–193.
Papini, M.R., Overmier, J.B., 1984. Autoshaping in pigeons: effects of partial reinforcement on acquisition and extinction. Rev. Interam. Psicol 18, 75–86.
Papini, M.R., Overmier, J.B., 1985. Partial reinforcement and autoshaping of the pigeon's key-peck behavior. Learn. Motiv 16, 109–123.
Papini, M.R., 2006. Role of surprising nonreward in associative learning. Jpn. J. Anim. Psychol. 56, 35–54.
Papini, M.R., 2014. Diversity of adjustments to reward downshifts in vertebrates. Int. J. Comp. Psychol. 27, 420–445.
Pavlov, I.P., 1927. Conditioned Reflexes. University Press, Oxford: Oxford.
Pearce, J.M., Hall, G., 1980. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychol. Rev 87, 532–552.
Pearce, J.M., Kaye, H., Hall, G., 1982. Predictive accuracy and stimulus associability: development of a model for Pavlovian conditioning. In: Commons, M.L., Herrnstein, R.J., Wagner, A.R. (Eds.), Quantitative Analyses of Behavior, III. Ballinger, Cambridge, MA, pp. 241–255.
Peciña, S., Cagniard, B., Berridge, K.C., Aldridge, J.W., Zhuang, X., 2003. Hyperdopaminergic mutant mice have higher 'wanting' but not 'liking' for sweet rewards. Behav. Neurosci 23, 9395–9402.
Pelchat, M.L., Grill, H.J., Rozin, P., Jacobs, J., 1983. Quality of acquired responses to tastes by *Rattus norvegicus* depends on type of associated discomfort. J. Comp. Psychol. 97, 140–153.
Pellegrini, S., Muzio, R.N., Mustaca, A.E., Papini, M.R., 2004. Successive negative contrast after partial reinforcement in the consummatory behavior of rats. Learn. Motiv. 35, 303–321.
Peters, D.P., McHose, J.H., 1974. Effects of varied preshift reward magnitude on successive negative contrast effects in rats. J. Comp. Physiol. Psychol. 86, 85–95.
Phillips, A.G., Vacca, G., Ahn, S., 2008. A top-down perspective on dopamine motivation, and memory. Pharmacol. Biochem. Behav. 90, 236–249.
Polo, V., Bautista, M.L., 2006. Daily routines of body mass gain in birds: 2: An experiment with reduced food availability. Anim. Behav 72, 517–522.
Powell, D.A., Churchwell, J., Burriss, L., 2005. Medial prefrontal lesions and Pavlovian eyeblink and heart rate conditioning: effects of partial reinforcement on delay and trace conditioning in rabbits (*Oryctolagus cuniculus*). Behav. Neurosci 119, 180–189.
Pravosudov, V.V., Grubb, T.C., 1997. Management of fat reserves and food caches in tufted titmice (*Parus bicolor*) in relation to unpredictable food supply. Behav. Ecol. 8, 332–339.
Preuschoff, K., Bossaerts, P., Quartz, S.R., 2006. Neural differentiation of expected reward and risk in human subcortical structures. Neuron 51, 381–390.
Ratikainen, I.I., Wright, J., 2013. Adaptive management of body mass by Siberian jays. Anim. Behav. 85, 427–434.
Rescorla, R.A., Wagner, A.R., 1972. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreniforcement. In: Black, A.H., Prokasy, W.F. (Eds.), Classical Conditioning II: Current Theory and Research. Appleton-Century-Crofts, New-York, pp. 64–99.
Rescorla, R.A., 1968. Probability of shock in the presence and absence of CS in fear conditioning. J. Comp. Physiol. Psychol. 66, 1–5.
Rescorla, R.A., 1971. Variations in the effectiveness of reinforcement and nonreinforcement following prior inhibitory conditioning. Learn. Motiv. 2, 113–123.
Rescorla, R.A., 1988. Pavlovian conditioning: it's not what you think it is. Am. Psychol 43, 151–160.
Reynolds, W.F., Pavlik, W.B., 1960. Running speed as a function of deprivation period and reward magnitude. J. Comp. Physiol. Psychol. 53, 615–618.
Reynolds, W.F., 1958. Acquisition and extinction of the conditioned eyelid response following partial and continuous reinforcement. J. Exp. Psychol. 55, 335–341.
Robbins, T.W., Everitt, B.J., 1996. Neurobehavioural mechanisms of reward and motivation. Curr. Opin. Neurobiol. 6, 228–236.
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.
Robinson, M.J.F., Berridge, K.C., 2013. Instant transformation of learned repulsion into motivational 'wanting'. Curr. Biol 23, 282–289.
Robinson, T.E., Flagel, S.B., 2009. Dissociating the predictive and incentive motivational properties of reward-related cues through the study of individual differences. Biol. Psychol. 65, 869–873.
Robinson, M.J.F., Warlow, S.M., Berridge, K.C., 2014a. Optogenetic excitation of central amygdala amplifies and narrows incentive motivation to pursue one reward above another. J. Neurosci. 34, 16567–16580.
Robinson, M.J.F., Anselme, P., Fischer, A.M., Berridge, K.C., 2014b. Initial uncertainty in Pavlovian reward prediction persistently elevates incentive salience and extends sign-tracking to normally unattractive cues. Behav. Brain Res 266, 119–130.
Robinson, M.J.F., Anselme, P., Suchomel, K., Berridge, K.C., 2015. Amphetamine-induced sensitization and reward uncertainty similarly enhance the incentive salience of conditioned cues. Behav. Neurosci. 129, 502–511.
Ross, L.E., 1959. The decremental effect of partial reinforcement during acquisition of the conditioned eyelid response. J. Exp. Psychol. 57, 74–82.
Rosse, R.B., Fay-McCarthy, M., Collins Jr., J.P., Risher-Flowers, D., Alim, T.N., Deutsch, S.I., 1993. Transient compulsive foraging behavior associated with crack cocaine use. Am. J. Psychiatr. 150, 155–156.
Sadler, E.W., 1968. A within- and between-subjects comparison of partial reinforcement in classical salivary conditioning. J. Comp. Physiol. Psychol. 66, 695–698.
Saunders, B.T., Robinson, T.E., 2012. The role of dopamine in the accumbens core in the expression of Pavlovian-conditioned responses. Eur. J. Neurosci 36, 2521–2532.
Schultz, W., Apicella, P., Ljungberg, T., 1993. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. J. Neurosci. 13, 900–913.
Schultz, W., 1998. Predictive reward signal of dopamine neurons. J. Neurophysiol. 80, 1–27.
Shallice, T., McLeod, P., Lewis, K., 1985. Isolating cognitive modules with the dual-task paradigm: are speech perception and production separate processes. Q. J. Exp. Psychol. 37A, 507–532.
Singer, B.F., Scott-Railton, J., Vezina, P., 2012. Unpredictable saccharin reinforcement enhances locomotor responding to amphetamine. Behav. Brain Res. 226, 340–344.
Skinner, B.F., 1938. The Behavior of Organisms. Appleton, New York.
Smith, K.S., Berridge, K.C., Aldridge, J.W., 2011. Disentangling pleasure from incentive salience and learning signals in brain reward circuitry. Proc. Natl. Acad. Sci. 108, E255–E264.
Snyder, C.R., 1994. The Psychology of Hope: You can Get There from Here. Free Press, New York.
Snyder, C.R., 1996. To hope, to lose, and to hope again. J.Pers. Interpers. Loss 1, 1–16.
Sodano, R., Wulfert, E., 2010. Cue reactivity in active pathological, abstinent pathological, and regular gamblers. J. Gambl. Stud. 26, 53–65.
Stagner, J.P., Zentall, T.R., 2010. Suboptimal choice behavior by pigeons. Psychon. Bull. Rev. 17, 412–416.
Stickgold, R., Walker, M.P., 2005. Memory consolidation and reconsolidation: what is the role of sleep? Trends in Neurosci. 28, 408–415.
Stout, S.C., Miller, R.R., 2007. Sometimes-competing retrieval (SOCR): A formalization of the comparator hypothesis. Psychol. Rev. 114, 759–783.
Sunsay, C., Rebec, G.V., 2008. Real-time dopamine efflux in the nucleus accumbens core during Pavlovian conditioning. Behav. Neurosci 122, 358–367.
Sunsay, C., Rebec, G.V., 2014. Extinction and reinstatement of phasic dopamine signals in the nucleus accumbens core during Pavlovian conditioning. Behav. Neuroci. 128, 579–587.
Sutton, R.S., Barto, A.G., 1981. Toward a modern theory of adaptive networks: expectation and prediction. Psychol. Rev. 88, 135–170.
Tan, C.O., Bullock, D., 2008. A local circuit model of learned striatal and dopamine cell responses under probabilistic schedules of reward. J. Neurosci. 28, 10062–10074.
Thomas, E., Wagner, A.R., 1964. Partial reinforcement of the classically conditioned eyelid response in the rabbit. J. Comp. Physiol. Psychol. 58, 157–158.
Tindell, A.J., Berridge, K.C., Zhang, J., Peciña, S., Aldridge, J.W., 2005. Ventral pallidal neurons code incentive motivation: amplification by mesolimbic sensitization and amphetamine. Eur. J. Neurosci 22, 2617–2634.
Tindell, A.J., Smith, K.S., Berridge, K.C., Aldridge, J.W., 2009. Dynamic computation of incentive salience: 'Wanting' what was never 'liked'. J. Neurosci. 29, 12220–12228.
Toates, F., 2014. How sexual desire works. In: The Enigmatic Urge. University Press, Cambridge: Cambridge.

Vardaris, R.M., Fitzgerald, R.D., 1969. Effects of partial reinforcement on a classically conditioned eyeblink response in dogs. J. Comp. Physiol. Psychol. 67, 531–534.
Voon, V., Hassan, K., Zurowski, M., de Souza, M., Thomsen, T., Fox, S., Lang, A.E., Miyasaki, J., 2006. Prevalence of repetitive and reward-seeking behaviors in Parkinson disease. Neurology 67, 1254–1257.
Wall, A.M., Goodrich, K.P., 1964. Differential responding on reinforcement and nonreinforcement trials occurring in fixed repeated patterns. Psychon. Sci. 1, 193–194.
Wasserman, E.A., Hunter, N.B., Gutowski, K.A., Bader, S.A., 1975. Autoshaping chicks with heat reinforcement: the role of stimulus–reinforcer and response–reinforcer relations. J. Exp. Psychol. Anim. Behav. Process. 104, 158–169.
Williams, B.A., McDevitt, M.A., 2002. Inhibition and superconditioning. Psychol. Sci. 13, 454–459.
Witter, M.S., Swaddle, J.P., 1995. Dominance, competition, and energetic reserves in the European starling. Sturnus vulgaris. Behav. Ecol. 6, 343–348.
Zack, M., Featherstone, R.E., Mathewson, S., Fletcher, P.J., 2014. Chronic exposure to a gambling-like schedule of reward predictive stimuli can promote sensitization to amphetamine in rats. Front. Behav. Neurosci. 8, 36.
Zentall, T.R., Sherburne, L.M., Zhang, Z., 1997. Shared attention in pigeons: retrieval failure does not account for the element superiority effect. Learn. Motiv. 28, 280–308.
Zentall, T.R., 2005. Selective and divided attention in animals. Behav. Process. 69, 1–15.
Zhang, J., Berridge, K.C., Tindell, A.J., Smith, K.S., Aldridge, J.W., 2009. A neural computational model of incentive salience. PLoS Comput. Biol. 5, e1000437.