Behavioural Processes 210 (2023) 104893

Contents lists available at ScienceDirect

Behavioural Processes

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

Models of conditioned reinforcement and abnormal behaviour in captive animals

Vera Vinken [a,d,*], Lena Lidfors [b], Jenny Loberg [b,c], Anna Lundberg [b], Johan Lind [d], Markus Jonsson [d], Stefano Ghirlanda [d,e,f], Magnus Enquist [d,g]

[a] Biosciences Institute, Newcastle University, Newcastle upon Tyne, UK
[b] Department of Animal Environment and Health, Swedish University of Agricultural. Sciences, P.O. Box 234, SE-53223 Skara, Sweden
[c] Foundation Nordens Ark, Åby Säteri SE-45693, Sweden
[d] Centre for the Study of Cultural Evolution, Stockholm University, 10691 Stockholm, Sweden
[e] Department of Psychology, Brooklyn College, 2900 Bedford Ave, Brooklyn, NY 11210, USA
[f] Departments of Psychology and Biology, Graduate Center of CUNY, 365 5th Avenue, 10016 NY, USA
[g] Department of Zoology, Stockholm University, 11418 Stockholm, Sweden

ARTICLE INFO

Keywords:
Abnormal behaviour
Associative learning
Stereotypic behaviour
Mathematical model
Conditioned reinforcement
Animal welfare

ABSTRACT

Abnormal behaviours are common in captive animals, and despite a lot of research, the development, maintenance and alleviation of these behaviours are not fully understood. Here, we suggest that conditioned reinforcement can induce sequential dependencies in behaviour that are difficult to infer from direct observation. We develop this hypothesis using recent models of associative learning that include conditioned reinforcement and inborn facets of behaviour, such as predisposed responses and motivational systems. We explore three scenarios in which abnormal behaviour emerges from a combination of associative learning and a mismatch between the captive environment and inborn predispositions. The first model considers how abnormal behaviours, such as locomotor stereotypies, may arise from certain spatial locations acquiring conditioned reinforcement value. The second model shows that conditioned reinforcement can give rise to abnormal behaviour in response to stimuli that regularly precede food or other reinforcers. The third model shows that abnormal behaviour can result from motivational systems being adapted to natural environments that have different temporal structures than the captive environment. We conclude that models including conditioned reinforcement offer an important theoretical insight regarding the complex relationships between captive environments, inborn predispositions, and learning. In the future, this general framework could allow us to further understand and possibly alleviate abnormal behaviours.

1. Introduction

Animals in captivity often develop stereotypic behaviour or other types of abnormal behaviours. Stereotypic behaviours have historically been described as 'repetitive, unvarying sequences of movements without any obvious goal or function' (Odberg, 1987; Mason, 1991a) and can reflect issues with the physical and psychological health of the animal (Mason, 1991b; Rushen and Mason, 2006). As discussed below, we are interested in modelling the role of conditioned reinforcement in the development and/or persistence of (elements of) abnormal behaviours. We use the broader term 'abnormal behaviour' in this paper, as our models are not specific to any examples and could be relevant to behaviours not categorised as 'stereotypic behaviour'. With this, we refer to behaviours that deviate from behaviours seen in wild-living animals (Wiepkema, 1985; Broom and Fraser, 2015). These behaviours are observed in many taxonomic groups and have many different expressions such as stereotypic pacing in carnivores (Clubb and Mason, 2003), licking/biting of non-food objects in ungulates (Bashaw et al., 2001) and in horses (Sarrafchi and Blokhuis, 2013), self- and other-directed aggression in birds (Mellor et al., 2018) and circular swimming in marine mammals (Gygax, 1993). For reviews, see Rushen and Mason (2006) and Mason (1991a). Studies in applied ethology have generated important insights regarding abnormal behaviour, linking its nature and severity to environmental factors (Mason, 2006; Radkowska

* Corresponding author at: Biosciences Institute, Newcastle University, Newcastle upon Tyne, UK
E-mail address: veraannevinken@gmail.com (V. Vinken).

https://doi.org/10.1016/j.beproc.2023.104893
Received 20 October 2022; Received in revised form 16 May 2023; Accepted 18 May 2023
Available online 19 May 2023
0376-6357/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

et al., 2020), noting that they often are species-specific (Mason and Mendl, 1997; Radkowska et al., 2020), observing there can be substantial individual variation within species such as rodents and primates (Mason and Mendl, 1997; Bashaw et al., 2001; Mason et al., 2007), and showing that they are difficult to eliminate or reduce once established (Mason and Latham, 2004; Swaisgood and Shepherdson, 2006). Suggested explanations or causes for abnormal behaviour often appeal to either difference between the captive and natural environment, or to neurophysiological changes in the animal. The former relates to the fact that these abnormal behaviours are absent in nature, which has been suggested to indicate a mismatch between the captive and the natural environment that the animal's behaviour system is adapted to (Bergeron et al., 2006; Clubb and Vickery, 2006). This mismatch can result in stress and frustration when the behavioural and motivational needs of the animal are not met (Mason, 2006). The latter explanation considers neurophysiological changes in the animal which lead to dysfunction in the CNS and prevent the animal from functioning 'normally' (Garner, 2006; McBride and Hemmings, 2009; Dìez-León et al., 2019). In addition to these theories, a suggested underlying mechanism of abnormal behaviours is associative learning through reinforcement. It is proposed that the arrival of primary reinforcers (e.g. food, social interaction or foraging opportunities) can inadvertently reinforce abnormal behaviours. For example, the arrival of food can reinforce the behaviours exhibited just before the feeding takes place (Mason, 1993; Mellor, 2020; Anderson et al., 2020). In domestic animals, inadvertent reinforcement through gaining caregiver attention can cause unwanted behaviour (Mills and Luescher, 2006). Additionally, stereotypic behaviour itself could have a reinforcing effect if it allows the animal to better 'cope' with stressors in the environment (Mason, 1991a; Würbel, 2006; Mason, 2006). In this paper, we aim to further elaborate on this research by using a computational approach and modelling the role of conditioned reinforcement in the development and persistence of abnormal behaviour.

A conditioned reinforcer is an initially neutral stimulus that derives value from having previously predicted primary reinforcers such as food and water (Pierce and Cheney, 2013). For example, repeatedly sounding a bell before delivering food turns the bell into a conditioned reinforcer (in addition to resulting in overt behavioural conditioning; Pavlov, 1927, Skinner, 1938). That is, the bell becomes capable, by itself, to reinforce Pavlovian and instrumental behaviour (Mackintosh, 1983; Williams, 1994). Conditioned reinforcement is widely exploited in animal training and husbandry, such as in clicker training (McGreevy and Boakes, 2011). Two circumstances suggest that conditioned reinforcement may contribute to abnormal behaviour. Firstly, a conditioned reinforcer itself can establish other conditioned reinforcers. For example, a light that predicts a bell that predicts food will also become a conditioned reinforcer. Consequently, animals can learn to respond to stimuli that have never been directly paired with primary reinforcement. Secondly, animals will respond to conditioned reinforcers with their species-specific behavioural repertoire, which can be counterproductive in captive environments. For example, animals are likely to respond with feeding behaviours (e.g. chewing) when responding to a stimulus that has been directly associated with food.

We use new models of associative learning (see section below) to explore the role of conditioned reinforcement in the development and persistence of abnormal behaviours. Our models include two assumptions regarding underlying causes of abnormal behaviour: 1) There is a mismatch between the captive environments and inborn predispositions (Bergeron et al., 2006; Clubb and Vickery, 2006), and 2) Animals can learn sequences of behaviour (i.e. sequences of the same one or more actions repeated) due to the establishment of conditioned reinforcers that support behaviour (Enquist et al., 2016).

First, we will introduce "A-learning", a general model of associative learning that includes conditioned reinforcement, genetic predispositions and internal motivation (Enquist et al., 2016; Ghirlanda et al., 2020; Enquist et al., 2023). Then, we use A-learning to model 3 idealised scenarios to further develop the hypothesis that conditioned reinforcement affects the development of abnormal behaviour. Our first scenario examines the question 'How can abnormal behaviours, such as locomotor stereotypies, arise from certain spatial locations acquiring conditioned reinforcement value?'. Our second scenario examines 'How can conditioned reinforcement support abnormal behaviour in response to stimuli that regularly precede food or other reinforcers?'. Finally, our third scenario examines the question 'How can abnormal behaviour be the result of a mismatch between the motivational system, which is adapted to the natural environment, and the captive environment, which offers a different temporal structure?'. To investigate the cumulative effects of long sequences of experiences, we use computer simulations of learning agents in virtual environments (Jonsson et al., 2021). Finally, we develop novel predictions that can form a basis for future modelling and empirical work. Please note that the scenarios we consider simplify reality to better highlight how conditioned reinforcement can yield a diversity of abnormal behaviour, in interaction with inborn predispositions. For each scenario, we mention specific empirical examples that fit the model qualitatively. A quantitative evaluation, however, will require detailed modelling of specific aspects in each empirical case, which we leave to future work (see also the Discussion).

2. Associative learning with conditioned reinforcement

We use computational models from machine learning (Sutton and Barto, 2018) that are providing us with a more complete understanding of associative learning, including how sequences of behaviour can be acquired (Enquist et al., 2016; Ghirlanda et al., 2020). Our own version of these models, called "A-learning", augments the successful Rescorla and Wagner (1972) model with conditioned reinforcement, genetic predispositions, and motivation. Here we present the model informally, referring to Supplementary Materials, Enquist et al. (2016), and Ghirlanda et al. (2020) for further details.

The model consists of a decision-making rule and two learning processes. Given a particular stimulus situation '*S*', the decision-making equation first calculates the support (Eq. 1) for each of the available behavioural responses. The support for a particular response collects all relevant causal factors into one variable (McFarland and Houston, 1981). For instance, the support for responding with behaviour '*B*' towards stimulus '*S*' could look like this:

$$Support[S{\rightarrow}B] = Memory[S{\rightarrow}B] + Inborn\ predispositions\ [S{\rightarrow}B] + Internal\ motivation\ [S{\rightarrow}B] \quad (1)$$

where the stimulus-response memory ("*Memory*" in Eq. 1) stores past experiences of responding with *B* towards *S*. Inborn predispositions are species-specific fixed values that are finetuned to the environment through genetic evolution. They either increase or decrease the support for *B*. Motivation includes internal factors that promote or inhibit *B* ("*Internal motivation*" in Eq. 1). It is straightforward to include more detailed stimulus control, for example, by stimulus compounds or variations in stimulus intensity (see Supplementary Materials). In functional terms, the support for performing a behaviour can be said to represent the subjective value the animal attributes to the response. For instance, if *B* is a feeding behaviour, the support would increase with hunger (higher feeding motivation) and with memories of obtaining food by responding to *S* with *B*.

Once the support values are calculated, the actual response is determined by competition among behaviours. For example, in a situation with only two behavioural options, '*B1*' and '*B2*', the probability '*PR*' of *B1* is calculated as:

$$PR(B1) = \frac{Support[S{\rightarrow}B1]}{Support[S{\rightarrow}B1] + Support[S{\rightarrow}B2]} \quad (2)$$

and similarly for *Pr(B2)*. That each behaviour is assigned a probability means that even behaviours with less support are sometimes

performed. This strikes a compromise between immediate gains (choosing the response with the highest support) and exploration. This is essential for most learning, as without exploration the animal can get stuck with a particular behaviour and never perform potentially better options. Eq. (2) describes animal decision-making as a function of the value of alternative behaviours and can be further refined for improved fit (Bridle, 1990; Herrnstein, 1961, 1970; Baum, 1974). For example, consider a T-maze with two arms, 1 and 2. If both arms contain the same reward, say a food pellet, the support for going to each arm will be the same, resulting in $Pr(B1) = Pr(B2) = 0.5$, or an indifferent choice. When the reward in arm 1 is larger, the animal will learn a higher support value for arm 1, resulting in more frequent visits to this arm (we will explain shortly how this learning occurs). If the difference between the arms becomes substantial, the animal will eventually choose arm 1 nearly every time. The effect of internal motivation can be included when calculating support. For example, if arm 1 of a T-maze leads to food and arm 2 to water, hungry animals will learn to go to arm 1 and thirsty animals to arm 2.

Learning results in changed support for each behaviour, reflecting positive and negative experiences, thus altering what decisions are made. There are two learning processes in the A-learning model. The first one, stimulus-response (S-R) learning, directly modifies stimulus-response associations ('*Memory[S->B]*' in Eq. 1). A-learning uses the same learning rule as the Rescorla and Wagner (1972) model, which leads the value of *Memory[S->B]* to reflect the perceived value of the next stimulus. For example, if the animal repeatedly experiences the sequence *S->B->food*, then *Memory[S->B]* will eventually equal the value of the food. The second learning process, stimulus value learning, describes conditioned reinforcement and influences decision-making indirectly, by changing the perceived value of stimuli. Thus, through stimulus value learning, the stimulus *S* in the sequence *S->R->food* will also acquire the value of food. Afterward, the stimulus *S* will itself be perceived as a reinforcer, and it will be capable of increasing the support for behaviours that precede *S*. Together, S-R learning and stimulus value learning enable the model to reproduce many findings about Pavlovian and instrumental learning (Ghirlanda et al., 2020; Enquist et al., 2023).

As the animal learns about an environment, it experiences many sequences of stimuli, which can result in the value of primary reinforcers spreading to other stimuli. These can include stimuli that occur much earlier in time than the primary reinforcement, as long as they reliably predict it. For example, a sequence of stimuli like *S1->S2->S3->Food* will first result in *S3* becoming a conditioned reinforcer, then *S2* (supported by the conditioned reinforcement value of *S3*), and eventually *S1*. At this point, behaviour that produces *S1* will be reinforced even if *S1* itself is not a primary reinforcer. Skinner (1938) referred to this process as "chaining" (see also Mackintosh, 1974; Williams, 1994; Pierce and Cheney, 2006). In natural environments, this mechanism encodes knowledge about the environment and enables animals to learn behavioural sequences that eventually lead to valuable outcomes (Enquist et al., 2016). However, as we discuss below, the same mechanism can backfire in captive environments.

The following is a concrete example of learning a behavioural sequence through conditioned reinforcement. To drink from an automatic water system, a cow must first Approach the *Water bowl* and then Push the *Paddle* with her nose, which will open a valve and fill the bowl with water. This forms the following sequence (stimuli in italic):

Water bowl → Approach → *Paddle* → Push → *Water*

At first, the cow may approach the water bowl because she has seen water in it, because she has seen other cows in its vicinity, or by chance. Through S-R learning, however, the cow learns to Push the *Paddle*, by initially performing this behaviour by chance, and then experiencing the positive outcome; *Water*. This results in an increased value of Memory [*Paddle*->Push], and, additionally, in an increased value of the *Paddle* stimulus through stimulus-value learning. At this point, the cow can learn to Approach the *Water bowl* because this stimulus now has a conditioned reinforcement value and thus can affect the value of Memory[*Water bowl*->Approach]. In this way, animals can learn sequences of behaviour that include steps, like approaching the water bowl, that are not intrinsically rewarding.

Below we apply A-learning to specific situations that arise in captive environments further developing the hypothesis that conditioned reinforcement can contribute to the development of abnormal behaviour. All models are simulated using version 1.1 of a general software for simulating learning phenomena (Jonsson et al., 2021) available at www.learningsimulator.org/. Simulation details and simulation scripts are included as Supplementary Materials. The simulations include parameters that describe, for example, the primary value of stimuli, the speed of learning, and the strength of genetic predisposition for performing particular behaviours. Changing these values affects simulation outcomes in a way that matches current knowledge of associative learning and decision-making.

2.1. Model 1: Transfer of value across spatial locations

Here, we show how animals can learn to perform non-productive responses through the transfer of value across spatial locations. As an example, we explore the possibility that movement stereotypies arise due to animals having a predisposition to move when in an unsatisfying environment, in combination with the delivery of primary reinforcement. Examples of these dissatisfactions could be hunger, restriction or social isolation. Movement stereotypies are common in captive animals (e.g. Bashaw et al., 2001; Clubb and Vickery, 2006; Roberts et al., 2017). This type of behaviour is especially observed in carnivores with naturally large home-ranges that are kept in small enclosures (Bashaw et al., 2001; Clubb and Mason, 2003; Clubb and Mason, 2007; Kroshko et al., 2016) or social species when kept in isolation (Cooper and Mason, 1998; Cooper et al., 2000). These stereotypies may persist if the enclosure is enlarged or otherwise enriched (Hansen et al., 1994; Hansen and Jeppesen, 2000; Mason and Latham, 2004; Swaisgood and Shepherdson, 2006).

In this model, we focus on how the predisposition to move, combined with a restricted enclosure, favours the transfer of reward value from locations where the animal receives reinforcement (e.g. food) to other locations, which in turn can reinforce visiting these locations. We consider a restricted environment with only two locations, labelled locations 0 and 1 (Fig. 1A). Food is delivered at location 0, at regular intervals. In between feeding opportunities, the subject has the option of being still or moving between locations. Moving has a small cost, whereas staying in the location has no cost. As mentioned above, we assume that individuals have an inborn predisposition to move: everything else being equal, the decision to move is more frequent than the decision to stay in the current location.

Fig. 2 shows that, under these assumptions, associative learning can give rise to stereotypic movement between the two locations. Fig. 2A shows how the probability of moving between the two locations increases with time. This occurs despite the fact that the food is always delivered in location 0. The initial movement away from the location where food is received (location 0) is due to the animals' predisposition to move. Fig. 2 B shows how location 1 acquires significant stimulus value over time and thus becomes a conditioned reinforcer that rewards moving from location 0 to location 1. A similar development occurs for location 0, resulting in movement back and forth between locations. One can say the subject falsely "believes" that moving between the locations produces food. Fig. 2C illustrates the importance of the predisposition to move, in the absence of which much less movement develops. Fig. 2D shows the importance of conditioned reinforcement, as in its absence (pure stimulus-response learning) less movement develops (compare to Fig. 2A). S-R learning cannot produce conditioned reinforcers on its own, and this explains why much more movement emerges under A-learning (see Fig. 2). Fig. 2E illustrates the effect of increasing

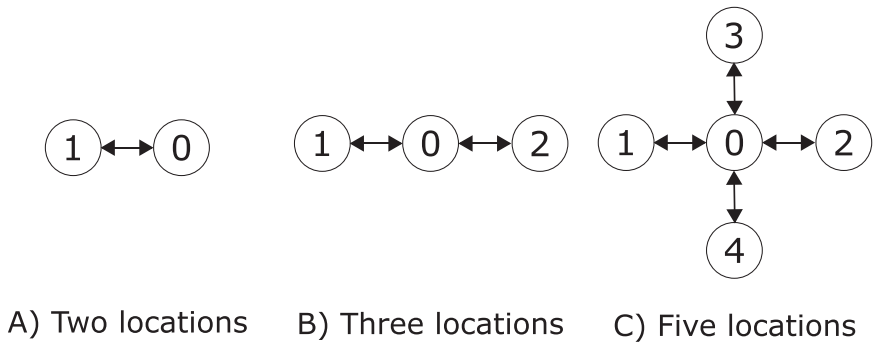

Fig. 1. The environments used in this simulation. The circles indicate locations the animal has access to, and the arrows indicate available movements. Food is delivered only at location 0. The left environment is the most restricted environment used with just two locations, the middle one a slightly more complex environment with one additional location, and the left one an even more complex environment with a total of five locations.

environmental complexity (environments in Fig. 1B and 1C), after initial experiences with the environment in Fig. 1A. Such an increase has little effect: the subject continues to move between locations 0 and 1, rarely entering the new locations. The first bar in Fig. 2E shows the acquisition, where the subjects have access to 2 locations and the behaviour is established. The second bar shows that movement between location 0 and 1 continues,

even when environmental complexity is added. The third bar suggests that adding more complexity, with five locations instead of 3, does not affect the behaviour differently. As a control, the rightmost bar in Fig. 2E shows that, when placed in a three-location environment from the start (Fig. 1B), subjects learn to visit locations 1 and 2 equally. Fig. 2 F is theoretically important. It shows that, in the absence of conditioned reinforcement, increasing enclosure complexity does reduce the stereotypy: after some time, subjects visit all available locations equally often (bar 2 and 3 in Fig. 2F).

2.2. Model 2: Transfer of value to temporally distant stimuli

This model builds upon the previous one ('Model 1′ in the section above), as we consider animals that receive food, or other primary reinforcement, and learn that performing predisposed behaviours leads to this reward. Here, we take into account sequences of stimuli that are predictive of the upcoming primary reward. This is common in captive environments, where there are often stimuli that predict feeding (e.g. sounds, the appearance of staff or smells in the environment). We show how the delivery of a primary reward can affect the response to stimuli that occur well in advance of such reinforcement. We suggest that, in artificial environments, this can trigger periods of unproductive predisposed behaviours. A possible example of such anticipatory behaviour occurs in mink kept in captivity, where the animals start moving in their cage well in advance of feeding, possibly in response to sound stimuli from the food truck starting to move along the cages at some distance (Mason and Mendl, 1997; Axelsson et al., 2009).

For our model we focus on how conditioned reinforcement can result in animals responding to stimuli that occur well in advance of the delivery of food or other primary reinforcement. We consider subjects exposed to sequences of stimuli composed of one or several repetitions of stimulus *S3*, followed by repetition of a stimulus *S2*, followed by a single stimulus *S1*. To each stimulus, the animal can respond with foraging behaviour or ignore the stimulus. Responding to *S1* results in a food reward, while responding to *S3* and *S2* has no consequence. We also assume that there is an inborn predisposition to perform the movement (as in Model 1) and that responding is associated with a small cost.

We have explored two different sequences of the above kind starting with:

$S3 \rightarrow 5 * S2 \rightarrow S1 \rightarrow Reward$ (Sequence 1)

where *5 * S2* means five repetitions of *S2*. These need not be actual repetitions, but can also signify a longer duration of *S2*. The result of simulating exposure to this sequence is displayed in the upper panels in Fig. 3. Fig. 3A shows how responding towards both *S3* and the repeated *S2*-stimuli develops.

despite the response being costly and unproductive. Fig. 3B shows that these responses are supported by *S2* accruing stimulus value, that is by S2 becoming a conditioned reinforcer that can reward responses to *S3* and to itself. The latter arises because *S2* follows itself part of the time. The response toward *S3* and *S2* depends on the response predisposition, when this is removed, no responding toward *S3* and *S2* develops (Fig. 3C).

Consider now the following sequence:

$5 * S3 \rightarrow 100 * S2 \rightarrow S1 \rightarrow Reward$ (Sequence 2)

in which *S3* is repeated a few times, and *S2* many times. This sequence results in significant responding towards *S3* but, contrary to the Sequence 1, little responding toward *S2* (Fig. 3D). This illustrates that a stimulus can act as a conditioned reinforcer without being responded to (compare Fig. 3D and 3E). Because *S2* is repeated 100 times, the detriment of responding can be up to 100 times the cost of a single response. This cost is so large that the subject learns to ignore *S2*, even in the presence of moderately strong predispositions. Nevertheless, *S2* becomes a conditioned reinforcer because it precedes *S1*, which can reinforce the less costly response to *S3* (which is repeated only a few times). As in the case of Sequence 1, responding to *S3* does not develop when the predisposition for responding is removed (Fig. 3F).

2.3. Model 3: Mismatches between the motivational system and captive environment

In this section we show that associative learning models can account for internal stimuli that arise from motivational systems, and that conditioned reinforcement may modulate the effect of such stimuli. Motivational systems are tuned to each species' natural environment (McFarland and Houston, 1981), which could cause abnormal behaviour in artificial environments. An example that might apply here are herbivores that forage on plant material with low nutrient density, such that they are motivated to feed for long stretches of time. In captivity, however, herbivores such as cattle, sheep, and horses are often fed highly nutritious food in a short time. Thus, there may be times in which feeding motivation is high, but food is absent, which may result in the animal directing foraging behaviour to non-food stimuli (Bergeron et al., 2000; Bergeron et al., 2006; Radkowska et al., 2020).

An internal motivational state can be entered into the A-learning model as an internal stimulus that influences decisions. In the following example, we refer to feeding for ease of presentation, but our arguments apply to other motivational systems. We assume that feeding behaviour

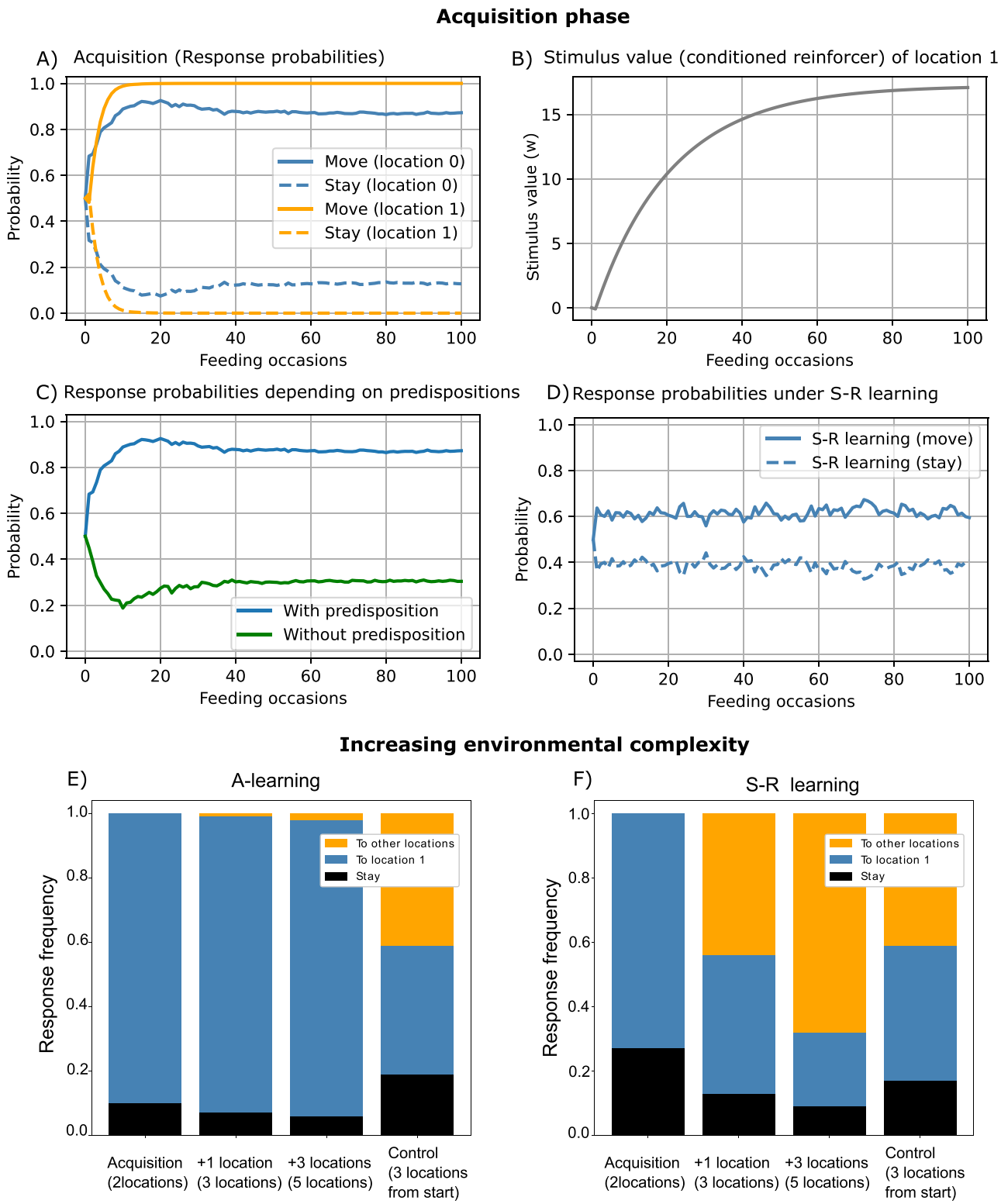

Fig. 2. Acquisition of a movement stereotypy and responses to increased environmental complexity. A) Shows how repetitive movement emerges with increasing probabilities of moving between the two locations (see environment A in Fig. 1) although food is always delivered in location 0. B) Shows how location 1 acquires significant stimulus value and becomes a potent conditioned reinforcer. C) Illustrates how the probability to move is influenced by the predisposition for moving. D) Shows that S-R learning also produces movements between the two locations, but less than under associative learning that includes conditioned reinforcement (see text). E) and F) illustrates the consequences of subsequent increased complexity for associative learning (as in the A-learning) and S-R learning respectively. The first bar in both panels shows the behaviour in the environment with two locations at the end of the acquisition phase. The two bars in the middle show the responses to increased complexity when one respectively three locations are added. The last bar is the result of a control simulation in which subjects are placed in an environment with three locations from the beginning. Parameter values used: Reward value 20, cost of moving = 0.05, The interval between feeding occasions is 50-time steps. The data shown are the average of 200 simulated subjects. For more information see the supplementary materials.

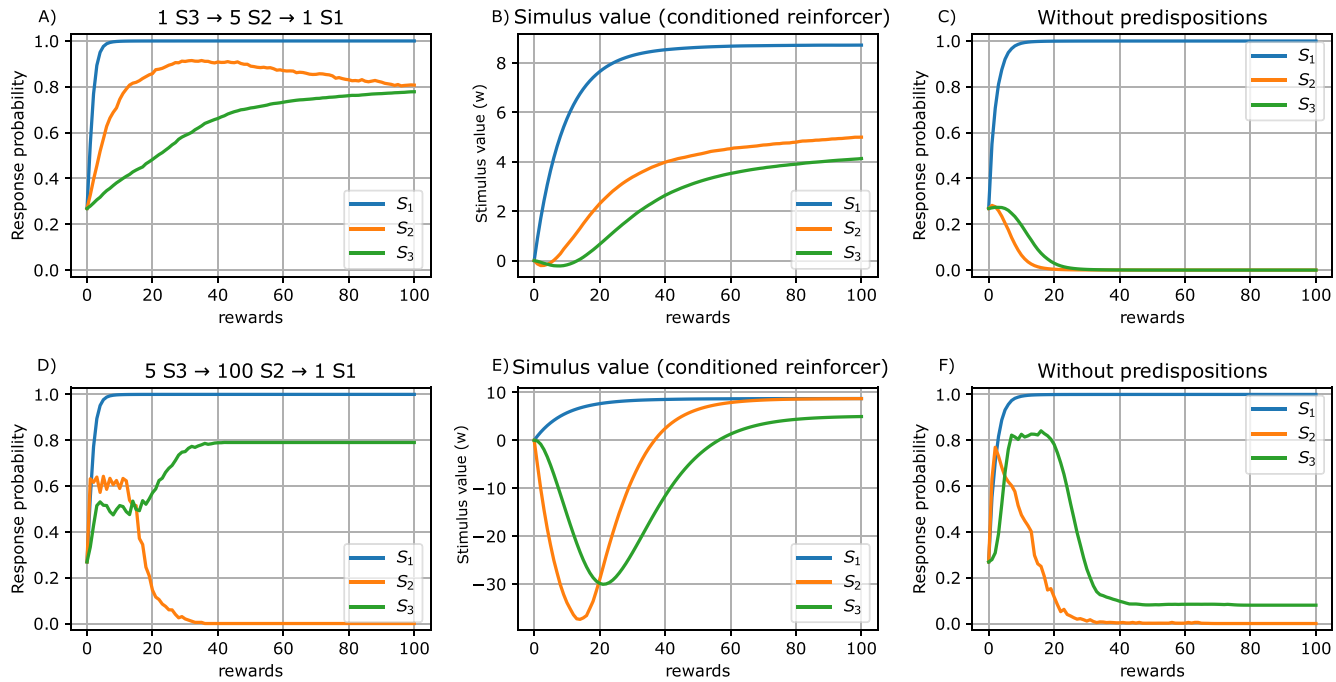

Fig. 3. Some consequences of temporal dependencies among stimuli. Upper panels refer to sequence 1 and the lower ones to sequence 2 (see text). A and D illustrate the development of responding for sequence 1 and 2 respectively. Panel B and E show the development of stimulus values (conditioned reinforcers), and panel C and F, responding in the absence of the predisposition. The data shown are the average of 200 simulated subjects. Parameter values used: Reward value = 10, response cost = 1. For more information see supplementary materials.

is predisposed (that is, especially easy to elicit) during a time window of high feeding motivation that is adapted to long feeding times in the wild. We also assume that feeding times in captivity are much shorter, causing feeding motivation to persist after food is consumed. We speculate that this mismatch may encourage feeding behaviour directed towards non-food stimuli, especially those bearing a degree of similarity with food

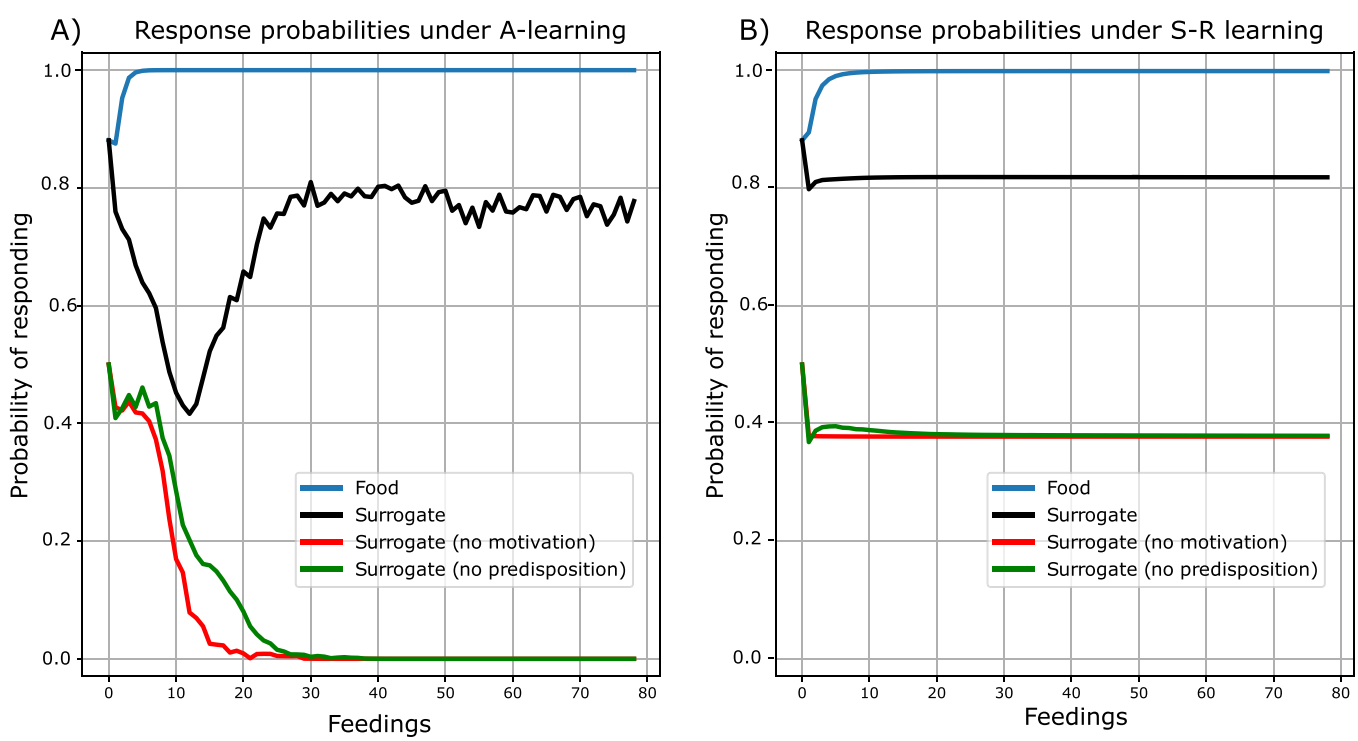

Fig. 4. An example of the mismatch between motivational systems and captive environments. Panel A shows how a combination of internal motivation and inborn predispositions can cause the development of unproductive responses towards a surrogate stimulus. Panel B shows what responses Stimulus-Response learning alone would predict. Parameter values used: Reward value = 5, response cost = 0.5, and inborn predispositions for responding when motivated = 2. The time window is 50 timesteps long and individuals are fed during the first 5 of these. The data shows an average of 200 simulated subjects in both cases. For more information see the supplementary materials.

stimuli. For example, non-food stimuli may resemble food stimuli visually, or by providing similar tactile sensations when chewing or suckling. We model this idea by assuming that the food stimulus is a compound of two perceptual elements, one of which is shared with a non-food stimulus in the environment. We refer to the non-food stimulus as the "surrogate" stimulus. Lastly, we assume that responding to the food stimulus is rewarding while responding to the surrogate stimulus carries a small cost.

The outcome of this scenario is illustrated in Fig. 4A. The subject performs feeding behaviour in response to the food stimulus, as is appropriate, but feeding responses also develop to the surrogate stimulus. Responding to the surrogate is lower than to the food, but it persists indefinitely. The predisposition to feed in the high-motivation state is necessary in order for the feeding response to the surrogate stimulus to develop (green vs. black line). Fig. 4A also shows that, under A-learning, responding to the surrogate when motivation is low, is suppressed (red line). These are learned responses. Fig. 4B shows simulation without conditioned reinforcement, that is, under simple stimulus-response learning. Responding to the food and surrogate stimuli is similar, but stimulus-response learning is less capable of suppressing feeding behaviour to the surrogate stimulus when feeding motivation is low.

3. Discussion

Our models show how abnormal behaviour can arise when the establishment of conditioned reinforcers is paired with genetic predispositions that do not match the captive environment. Whilst our work needs to be integrated with other theories, these models can help us understand the establishment and persistence of abnormal behaviour. Note that our models explore general learning and memory processes but that applying them to specific scenarios or behaviours will require more detailed modelling, taking into account the combination of species-specific predispositions, reinforcement history and the temporal structure of the environment (McGreevy and Boakes, 2011). Therefore, to validate our models and link the predictions suggested below to empirical observations, future work needs to include models of specific scenarios with more detail.

3.1. Model 1: Transfer of value across spatial locations

In model 1, we show how location 1, which is never directly paired with food, becomes a conditioned reinforcer that increases the movement response in the subjects. This response, to move when hungry, or dissatisfied is genetically predisposed and therefore particularly prone to exacerbation by conditioned reinforcers simply because it is more likely to occur than other behaviour. In nature, these predispositions, for example moving when searching for prey/patrolling a territory, are generally functional. However, they become counterproductive in many captive environments where access to food, space or social contact depends on established feeding schedules and cage sizes, instead of the individual's own locomotive efforts. Our model predictions show that these behaviours are difficult to alleviate once they are established. We also showed that, once acquired, unproductive movements are difficult to alleviate, mirroring the observation that environmental enrichment is often ineffective (Garner, 2006; Mason, 2006). While we have focused on locomotor stereotypies arising from food reinforcement, the model can be adapted to other types of repetitive behaviour, and to other primary reinforcers.

This model makes some predictions that can be developed and tested in future work. First, species with large home ranges and territories relative to their size, that rely on moving in their natural environment, should be more vulnerable to developing movement stereotypies, because of a stronger predisposition to move when in suboptimal environments. Additionally, we would expect that the repetitive behaviours that include multiple spatial locations, involve the location where the animal is fed. This would be necessary, in order for the other locations to become conditioned reinforcers. Lastly, we expect that placing obstacles along a stereotypical route may alter the route but not extinguish the stereotypy, as the animal would eventually learn to go around the obstacle to reach other locations with conditioned reinforcement value.

3.2. Model 2: Transfer of value to temporally distant stimuli

With model 2 we show that food-predicting stimuli that occur far away in time from primary reinforcement, can acquire conditioned reinforcement value and strengthen the occurrence of predisposed behaviour, even if the subject does not respond to the conditioned reinforcer itself. This shows that it is not only response sequences that need to be considered when studying abnormal behaviours. Other aspects must be taken into account, such as sequences of stimuli that are generated by the environment. In practice, we believe the predictability and repetitiveness of many captive environments is important here, with the same events predicting feeding (e.g. food delivery machines being turned on), for example in farmed mink (Axelsson et al., 2009; Olofsson and Lidfors, 2012) and laboratory rabbits (Lidfors, 1997). The situation in Sequence 2, suggesting a behaviourally silent transfer of value from a reward to a temporally remote stimulus, is currently a novel theoretical prediction of A-learning. Nevertheless, this result follows logically from what we know about conditioned reinforcement and has also some empirical support. For example, Cronin (1980) showed that, in laboratory experiments, long-duration stimuli can support responses to preceding short-duration stimuli, using sequences similar to Sequence 2 (Enquist et al., 2016).

Following our model predictions, we suggest two interesting hypotheses to be tested in future work. Firstly, in situations described by Sequence 2, responding to *S2* may develop initially before eventually disappearing (Fig. 3D), because the learning mechanism may need some time to learn that responding to *S2* is costly. Secondly, our model predicts that responding to *S3* is more likely to develop when the environment contains few stimuli and is highly predictable. Here, the animal is likely to receive the same sequences of events repeated without variation as there are only a few possible combinations of stimuli and behaviours. These are the most favourable circumstances for primary reinforcement value to spread to temporally distant stimuli. Even though the result of studies investigating the effect of predictability on abnormal behaviours vary (e.g. Bloomsmith and Lambeth, 1995; Johannesson and Ladewig, 2000; Gottlieb et al., 2013), a review by Bassett and Buchanan-Smith (2007) concludes that unpredictable feeding schedules are beneficial, as long as animals are provided with a unique reliable signal before feeding and when other less reliable stimuli that are further away in time, are removed. In line with this research and our models, we predict that, when the predictive stimuli around feeding time are removed, especially those far apart in time, the development of abnormal behaviour triggered by precursory stimuli might be prevented.

3.3. Model 3: Mismatches between the motivational system and captive environment

In model 3 we introduce motivation, showing how high motivation, combined with artificial environments can result in behaviour directed towards unnatural targets through conditioned reinforcement. Responding to the surrogate stimulus arises because the surrogate shares one perceptual element with the food stimulus and this element acquires reinforcement value. The animal will then learn to respond to this element and therefore the surrogate stimulus. The state of high motivation is important because feeding behaviour is only reinforcing in this state. The food itself remains a stronger stimulus for feeding than the surrogate, because all of its perceptual elements become associated with feeding. As a consequence, the subject performs feeding behaviour towards the food as long as food is present and starts directing feeding behaviour towards the surrogate once the food has been consumed.

These results fit with some results from empirical literature. For example, temporally restricted feeding has been suggested to increase chain chewing and bar biting in pigs (Spoolder et al., 1995; Whittaker et al., 1998) and short feeding times have been suggested to increase cross-sucking or sucking of inanimate objects in dairy calves (Loberg and Lidfors, 2001; de Passillé et al., 2011). We also showed that S-R learning cannot suppress responses to the surrogate in the same way as A-learning, because it only takes into account the small cost of one response to the surrogate stimulus, while conditioned reinforcement enables the subject to estimate the total cost of all responses to the surrogate (the surrogate attains negative conditioned reinforcement value). This is consistent with A-learning performing generally better than S-R learning in environments with sequential structure, because A-learning can take into account the future costs and benefits of responses (Enquist et al., 2016; Sutton and Barto, 2018). Nevertheless, inborn predispositions can prevent A-learning from correctly estimating costs and benefits in some environments, resulting in excess responding to the surrogate stimulus.

Thus far, we have assumed that responses to the surrogate stimulus are unproductive. However, some abnormal behaviours have beneficial consequences, such as nutrient ingestion, feelings of satiety or lowering of stress hormones (Bergeron et al., 2006). Our model predicts that, in these cases, the behaviours will be established even more easily, as they would yield both conditioned and primary reinforcement. Future work could investigate this further by developing more specific models including these primary reinforcers. Similarly, our model predicts that higher motivation would also lead to stronger responses toward the surrogate stimulus, for example, when animals are hungry rather than sated. The model also predicts that feeding time should affect responding to the surrogate, because feeding behaviour is preferentially directed to the appropriate food stimulus as long as this is present. Thus, lengthening feeding time may result in extinguishing the motivation to feed before a significant amount of responding to the surrogate occurs.

3.4. General discussion

Other studies considering the role of associative learning in abnormal behaviour, have often focused on inadvertent reinforcement of behaviours through primary rewards such as food or social interaction (Mellor, 2020; Anderson et al., 2020; Mason, 1993). Our work complements these theories by considering how sequences of experiences can alter the landscape of reinforcement by creating conditioned (learned) reinforcers. Conditioned reinforcers reflect the value of forthcoming events—up to several hours—and can be as powerful as primary (inborn) reinforcers. Our results indicate that understanding abnormal behaviours requires considering events that occur over extended periods of time.

Besides reinforcement learning, other potential factors contributing to abnormal behaviours include stress, frustration, neurophysiological changes that lead to CNS dysfunctions and coping strategies (Koolhaas et al., 1999; Clubb and Vickery, 2006; Rushen and Mason, 2006). We believe these concepts could complement our model as they affect which stimuli become conditioned reinforcers, what behaviours are likely to be selected and how persistent these behaviours might become. The 'coping hypothesis' is specifically interesting as it has been suggested that abnormal behaviour could produce outcomes that are inherently rewarding and function as a reaction strategy (Mason, 1991a; Würbel, 2006; Mason, 2006). 'Coping' refers to behavioural responses, active or passive, an animal might display in an attempt to control or change stressful situations (for review see Wechsler, 1995). These can include learned behaviours that help avoid stressors, increase well-being, or are inherently rewarding in another way, thereby reinforcing the behaviour itself (Würbel, 2006). Additionally, these are suggested to include unlearned behaviours that are thought to have a neurobiological origin (Koolhaas et al., 1999; Cabib, 2006) and are evolutionarily adapted behaviours (Wechsler, 1995). Future work exploring how coping and these other factors may be included in our model will, however, require that they are first formalized mathematically, by specifying how they influence learning and decision making. For example, the concept of coping could be incorporated in learning models by specifying the reinforcement value of 'coping' behaviours.

Our work stresses the importance of theoretical work in applied ethology and animal welfare, as the developmental processes that underlie abnormal behaviours are difficult to understand from empirical work alone. It is often hard to identify what reinforcers are operating in a given situation, and what motivational processes are engaged as controlling these variables empirically is challenging (Baragli et al., 2015). Furthermore, the sequential nature of learning means that abnormal behaviours can arise as long-term outcomes of incremental processes that are not easily observed, such as a build-up of stimulus value that can later reinforce maladaptive behaviour. Even though conditioned reinforcement and chaining are well-known phenomena, their long-term consequences can be counterintuitive and difficult to predict. Theoretical models and computer simulations ameliorate these difficulties because they can isolate potential causal factors and compute for us how animals might behave when exposed to specific sequences of internal and external stimuli. The latter includes understanding gene-environment interactions, as we have attempted above in simple cases, by running simulations with and without inborn predispositions. Future work could explore the predicted effects of enrichment and other environmental changes on abnormal behaviours.

Of potential relevance to animal welfare, learning models may be able to explain why abnormal behaviour is remarkably difficult to extinguish (Mason and Latham, 2004; Swaisgood and Shepherdson, 2006; Garner, 2006;). In fact, large associative strengths can be maintained by conditioned reinforcement even after primary reinforcement is withdrawn, leading to the animal getting 'stuck' in performing unproductive behaviour without exploring alternative ones. This is described in the section about the A-learning model, where the probability of choosing a behaviour increases with the underlying associative strength (Roper, 1983; Ghirlanda et al., 2020), which reflects the history of reinforcement for that behaviour (Herrnstein, 1961; Baum, 1981; Houston et al., 2021). Even behaviours with low associative strength are sometimes performed, which is essential in order to explore an environment effectively. However, a behaviour can become dominant if the corresponding associative strength becomes much larger than the others, and the animal becomes 'stuck' in performing the behaviour. This is also referred to as 'habitual' behaviour, and can be caused by overtraining (Dickinson, 1985; Balleine and Dickinson, 1998; Balleine et al., 2009; Ghirlanda et al., 2020). In our first model this is demonstrated as the decision to move can acquire a large associative strength as a consequence of location 1 acquiring stimulus value. Hence, the decision to move will be maintained by stimulus value far beyond the initial predisposition to move. This agrees with the suggestion that, over time, stereotypic behaviour may be maintained by different mechanisms than those which elicit it in the first place (Cronin, 1985). Note that, if large associative strengths are maintained by conditioned reinforcement (learned stimulus value), then abnormal behaviours may persist even when primary reinforcement contingencies are altered. Because conditioned reinforcement can arise from longer-term contingencies, altering it may require significant changes to the temporal structure of an animal's experience.

In conclusion, we have provided a proof of concept that the combination of conditioned reinforcement and genetic predispositions that mismatch the captive environment may be important in the development of abnormal behaviour. Reinforcement learning models can provide a general framework to study empirical cases for developing more detailed models that can guide future empirical work.

Funding

This work was supported by the Knut and Alice Wallenberg

Foundation (grant 2015.0005) and the Marianne and Marcus Wallenberg Foundation (grant 2021.0039).

CRediT authorship contribution statement

Vera Vinken: Conceptualization, Methodology, Writing – original draft, Visualisation, Writing – review, Editing. **Lena Lidfors**: Conceptualization, Writing – original draft, Writing – review, Editing. **Jenny Loberg:** Conceptualization, Writing – original draft, Writing – review, Editing. **Anna Lundberg**: Conceptualization, Writing – original draft, Writing – review, Editing. **Johan Lind:** Writing – review, Editing. **Markus Jonsson:** Software, Writing – review, Editing. **Stefano Ghirlanda**: Software, Writing – review, Editing. **Magnus Enquist:** Conceptualization, Methodology, Formal analysis, Writing – original draft, Software, Visualisation, Supervision, Writing – review, Editing.

Data Availability

No data was used for the research described in the article.

Acknowledgements

We thank Georgia Mason, Lindsey Kitchenham and an anonymous reviewer for their very valuable comments and suggestions. They helped us greatly improve this paper, especially with their thoughtful comments on the possible applications of our models.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.beproc.2023.104893.

References

Anderson, C., Von Keyserlingk, M.A.G., Lidfors, L.M., Weary, D.M., 2020. Anticipatory behaviour in animals: a critical review. Anim. Welf. 29 (3), 231–238. https://doi.org/10.7120/09627286.29.3.231.

Axelsson, H.M.K., Aldén, E., Lidfors, L., 2009. Behaviour in female mink housed in enriched standard cages during winter. Appl. Anim. Behav. Sci. 121 (3–4), 222–229. https://doi.org/10.1016/j.applanim.2009.09.013.

Balleine, B.W., Dickinson, A., 1998. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology 37 (4–5), 407–419. https://doi.org/10.1016/S0028-3908(98)00033-1.

Balleine, B.W., Liljeholm, M., Ostlund, S.B., 2009. The integrative function of the basal ganglia in instrumental conditioning. Behav. Brain Res. 199 (1), 43–52. https://doi.org/10.1016/j.bbr.2008.10.034.

Baragli, P., Padalino, B., Telatin, A., 2015. The role of associative and non-associative learning in the training of horses and implications for the welfare (a review). Ann. dell'Istituto Super. di Sanita 51, 40–51. https://doi.org/10.4415/ANN_15_01_08.

Bashaw, M.J., Tarou, L.R., Maki, T.S., Maple, T.L., 2001. A survey assessment of variables related to stereotypy in captive giraffe and okapi. Appl. Anim. Behav. Sci. 73 (3), 235–247. https://doi.org/10.1016/S0168-1591(01)00137-X.

Bassett, Lois, Buchanan-Smith, Hannah M., 2007. Effects of predictability on the welfare of captive animals. Appl. Anim. Behav. Sci. 102 (3–4), 223–245. https://doi.org/10.1016/j.applanim.2006.05.029.

Baum, W.M., 1974. On two types of deviation from the matching law: Bias and undermatching 1. J. Exp. Anaysis Behav. 22 (1), 231–242. https://doi.org/10.1901/jeab.1974.22-231.

Baum, W.M., 1981. Optimization and the matching law as accounts of instrumental behavior. J. Exp. Anal. Behav. 36 (3), 387–403. https://doi.org/10.1901/jeab.1981.36-387.

Bergeron, R., Bolduc, J., Ramonet, Y., Robert, S., Meunier-Salaün, M.-C., 2000. Feeding motivation and stereotypies in pregnant sows fed increasing levels of fibre and/or food. Appl. Anim. Behav. Sci. 70, 27–40. https://doi.org/10.1016/S0168-1591(00)00142-8.

Bergeron, R., Badnell-Waters, A.J., Lambton, S., Mason, G., 2006. Stereotypic oral behaviour in captive ungulates: foraging, diet and gastrointestinal function. In: Mason, G., Rushen, J. (Eds.), Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare, second ed. CABI. https://doi.org/10.1079/9780851990040.0019.

Bloomsmith, Mollie A., Lambeth, Susan P., 1995. Effects of predictable versus unpredictable feeding schedules on Chimpanzee behavior. Appl. Anim. Behav. Sci. 44, 65–74. https://doi.org/10.1016/0168-1591(95)00570-I.

Bridle, J.S., 1990. Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition. Neurocomputing: Algorithms, Archit. Appl. 68, 227–236. https://doi.org/10.1007/978-3-642-76153-9_28.

Broom, D.M., Fraser, A.F., 2015. Domestic Animal Behaviour and Welfare, fifth ed. CABI. https://doi.org/10.1079/9781780645391.0000.

Cabib, S., 2006. The Neurobiology of Stereotypy II: the Role of Stress. In: Mason, G., Rushen, J. (Eds.), Stereotypic animal behaviour: fundamentals and applications to welfare, second ed. CABI. https://doi.org/10.1079/9780851990040.0227.

Clubb, R., Mason, G., 2003. Captivity effects on wide-range carnivores. Nature 425 (6957), 473–474. https://doi.org/10.1038/425473a.

Clubb, R., Mason, G.J., 2007. Natural behavioural biology as a risk factor in carnivore welfare: How analysing species differences could help zoos improve enclosures. Appl. Anim. Behav. Sci. 102 (3–4), 303–328. https://doi.org/10.1016/j.applanim.2006.05.033.

Clubb, R., Vickery, S., 2006. Locomotory stereotypies in carnivores: does pacing stem from hunting, ranging, or frustrated escape. In: Mason, G., Rushen, J. (Eds.), Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare, second ed. CABI. https://doi.org/10.1079/9780851990040.0058.

Cooper, J.J., Mason, G., 1998. The identification of abnormal behaviour and behavioural problems in stabled horses and their relationship to horse welfare: a comparative review. Equinc I'Yi. J. Suppl. 21, 5–9. https://doi.org/10.1111/j.2042-3306.1998.tb05136.x.

Cooper, J.J., McDonald, L., Mills, D.S., 2000. The effect of increasing visual horizons on stereotypic weaving: implications for the social housing of stabled horses. Appl. Anim. Behav. Sci. 69, 67–83. https://doi.org/10.1016/S0168-1591(00)00115-5.

Cronin, G.M., 1985. The development and significance of abnormal stereotyped behaviours in tethered sows. PhD thesis, Wageningen University.

Cronin, P.B., 1980. Reinstatement of post-response stimuli prior to reward in delayed reward discrimination learning by pigeons. Anim. Learn. Behav. 8 (3), 352–358.

Dickinson, A., 1985. Actions and habits: the development of behavioural autonomy. Philos. Trans. R. Soc. Lond. 308 (1135), 67–78.

Dìez-León, M., Kitchenham, L., Duprey, R., Bailey, C.D., Choleris, E., Lewis, M., Mason, G., 2019. Neurophysiological correlates of stereotypic behaviour in a model carnivore species. Behav. Brain Res. 373 (July), 112056 https://doi.org/10.1016/j.bbr.2019.112056.

Enquist, M., Lind, J., Ghirlanda, S., 2016. The power of associative learning and the ontogeny of optimal behaviour. R. Soc. Open Sci. 3 (11), 160734 https://doi.org/10.1098/rsos.160734.

Enquist, M., Ghirlanda, S., Lind, J., 2023. The Human Evolutionary Transition: From Animal Intelligence to Culture. Princeton University Press,. https://doi.org/10.1515/9780691240763.

Garner, J., 2006. Perseveration and stereotypy - systems-level insights from clinical psychology. In: Mason, G., Rushen, J. (Eds.), Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare, second ed. CABI. https://doi.org/10.1079/9780851990040.0121.

Ghirlanda, S., Lind, J., Enquist, M., 2020. A-learning: a new formulation of associative learning theory. Psychon. Bull. Rev. 27, 1166–1194. https://doi.org/10.3758/s13423-020-01749-0.

Gottlieb, Daniel H., Kristine, Coleman, Brenda, Mc.Cowan, 2013. The effects of predictability in daily husbandry routines on captive rhesus macaques (Macaca mulatta. Appl. Anim. Behav. Sci. 143 (2–4), 117–127. https://doi.org/10.1016/j.applanim.2012.10.010.

Gygax, L., 1993. Spatial movement patterns and behaviour of two captive bottlenose dolphins (Tursiops truncatus): absence of stereotyped behaviour or lack of definition. Appl. Anim. Behav. Sci. 38 (3–4), 337–344. https://doi.org/10.1016/0168-1591(93)90031-J.

Hansen, C.P.B., Jeppesen, L.L., 2000. Short term behavioural consequences of denied access to environmental facilities in mink. Agric. Food Sci. Finl. 9 (2), 149–155. https://doi.org/10.23986/afsci.5656.

Hansen, S.W., Hansen, B.K., Berg, P., 1994. The effect of cage environment and ad libitum feeding on the circadian rhythm, behaviour and feed intake of farm mink. Acta Agric. Scand. A: Anim. Sci. 44 (2), 120–127. https://doi.org/10.1080/09064709409410190.

Herrnstein, R.J., 1961. Relative and absolute strength of response as a function of frequency of reinforcement. J. Exp. Anal. Behav. 4 (3), 267–272 https://doi.org/10.1901%2Fjeab.1961.4-267.

Herrnstein, R.J., 1970. On the law of effect 1. J. Exp. Anal. Behav. 13 (2), 243–266. https://doi.org/10.1901/jeab.1970.13-243.

Houston, A.I., Trimmer, P.C., McNamara, J.M., 2021. Matching behaviours and rewards. Trends Cog. Sci. 25 (5), 403–415. https://doi.org/10.1016/j.tics.2021.01.011.

Johannesson, Torfi, Ladewig, Jan, 2000. The effect of irregular feeding times on the behaviour and growth of dairy calves. Appl. Anim. Behav. Sci. 69 (2), 103–111. https://doi.org/10.1016/S0168-1591(00)00127-1.

Jonsson, M., Ghirlanda, S., Lind, J., Vinken, V., Enquist, M., 2021. Learning simulator: a simulation software for animal and human learning. J. Open Source Softw. 6 (58), 2891. https://doi.org/10.21105/joss.02891.

Koolhaas, J., Korte, S., De Boer, S., Van Der Vegt, B., Van Reenen, C., Hopster, H., De Jong, I., Ruis, M., Blokhuis, H., 1999. Coping styles in animals: current status in behavior and stress-physiology. Neurosci. Biobehav. Rev. 23 (7), 925–935. https://doi.org/10.1016/S0149-7634(99)00026-3.

Kroshko, J., Clubb, R., Harper, L., Mellor, E., Moehrenschlager, A., Mason, G., 2016. Stereotypic route tracing in captive Carnivora is predicted by species-typical home range sizes and hunting styles. Anim. Behav. 117, 197–209. https://doi.org/10.1016/j.anbehav.2016.05.010.

Lidfors, L., 1997. Behavioural effects of environmental enrichment for individually caged rabbits. Appl. Anim. Behav. Sci. 52 (1–2), 157–169. https://doi.org/10.1016/S0168-1591(96)01141-0.

Loberg, J., Lidfors, L., 2001. Effect of milkflow rate and presence of a floating nipple on abnormal sucking between dairy calves. Appl. Anim. Behav. Sci. 72 (3), 189–199.
Mackintosh, N., 1983. Conditioning and Associative Learning. Oxford University Press.
Mackintosh, N.J., 1974. The Psychology of Animal Learning. Academic Press.
Mason, G., 2006. Stereotypic behaviour in captive animals: fundamentals and implications for welfare and beyond. In: Mason, G., Rushen, J. (Eds.), Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare, second ed. CABI. https://doi.org/10.1079/9780851990040.0325.
Mason, G., Latham, N., 2004. Can't stop, won't stop: is stereotypy a reliable welfare indicator. Anim. Welf. 13, 57–69. https://doi.org/10.1017/S096272860001438X.
Mason, G., Mendl, M., 1997. Do the stereotypies of pigs, chickens and mink reflect adaptive species differences in the control of foraging. Appl. Anim. Behav. Sci. 53 (1–2), 45–58. https://doi.org/10.1016/S0168-1591(96)01150-1.
Mason, G., Clubb, R., Latham, N., Vickery, S., 2007. Why and how should we use environmental enrichment to tackle stereotypic behaviour. Appl. Anim. Behav. Sci. 102 (3–4), 163–188. https://doi.org/10.1016/j.applanim.2006.05.041.
Mason, G.J., 1991a. Stereotypies: a critical review. Anim. Behav. 41 (March 1990), 1015–1037. https://doi.org/10.1016/S0003-3472(05)80640-2.
Mason, G.J., 1991b. Stereotypies and suffering. Behav. Proc. 25 (2–3), 103–115. https://doi.org/10.1016/0376-6357(91)90013-P.
Mason, G.J., 1993. Age and context affect the stereotypies of caged mink. Behaviour 127 (3–4), 191–229. https://doi.org/10.1163/156853993×00029.
McBride, S., Hemmings, A., 2009. A neurologic perspective of equine stereotypy. J. Equine Vet. Sci. 29 (1), 10–16. https://doi.org/10.1016/j.jevs.2008.11.008.
McFarland, D., Houston, A., 1981. Quantitative ethology: the state space approach. Pitman advanced publishing program.
McGreevy, P., Boakes, R., 2011. Carrots and sticks: principles of animal training. Darlington Press.
Mellor, E., Brilot, B., Collins, S., 2018. Abnormal repetitive behaviours in captive birds: a Tinbergian review. Appl. Anim. Behav. Sci. 198, 109–120. https://doi.org/10.1016/j.applanim.2017.09.011.
Mellor, E.L., 2020. Does natural foraging niche influence captive animal health and welfare? Doctoral dissertation, University of Bristol.
Mills, D., Luescher, A., 2006. Veterinary and pharmacological approaches to abnormal repetitive behaviour. In: Mason, G., Rushen, J. (Eds.), Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare, second ed. CABI. https://doi.org/10.1079/9780851990040.0286.
Odberg, F.O., 1987. The influence of cage size and environmental enrichment on the development of stereotypies in bank voles (Clethrionomys glareolus). Behav. Proc. 14 (2), 155–173. https://doi.org/10.1016/0376-6357(87)90042-8.
Olofsson, L., Lidfors, L., 2012. Abnormal behaviour in Swedish farm mink during winter. Proceedings of the Xth International Scientific Congress in fur animal production. Springer, pp. 426–432.
de Passillé, A.M., Borderas, F., Rushen, J., 2011. Cross-sucking by dairy calves may become a habit or reflect characteristics of individual calves more than milk allowance or weaning. Appl. Anim. Behav. Sci. 133, 137–143. https://doi.org/10.1016/j.applanim.2011.04.020.
Pavlov, I.P., 1927. Conditioned Reflexes. Oxford University Press, Oxford.
Pierce, D.W., Cheney, C.D., 2013. Behavior Analysis and Learning. Psychology Press.
Radkowska, I., Godyn, D., Fic, K., 2020. Stereotypic behaviour in cattle, pigs and horses-a review. Anim. Sci. Pap. Rep. 38 (4), 303–319.
Rescorla, R.A., Wagner, A.R., 1972. A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. Classical Conditioning: Current Research and Theory. Appleton-Century-Crofts, pp. 64–99.
Roberts, K., Hemmings, A.J., McBride, S.D., Parker, M.O., 2017. Causal factors of oral versus locomotor stereotypy in the horse. J. Vet. Behav.: Clin. Appl. Res. 20, 37–43. ⟨https://doi.org/10.1016/j.jveb.2017.05.003⟩.
Roper, T., 1983. Learning as a biological phenomenon. Anim. Behav. 3, 178–212.
Rushen, J., Mason, G., 2006. A decade-or-more's progress in understanding stereotypic behaviour. In: Mason, G., Rushen, J. (Eds.), Stereotypic animal behaviour: fundamentals and applications to welfare, 2nd ed..,. CABI. https://doi.org/10.1079/9780851990040.0001.
Sarrafchi, A., Blokhuis, H.J., 2013. Equine stereotypic behaviors: causation, occurrence, and prevention. J. Vet. Behav.: Clin. Appl. Res. 8 (5), 386–394. https://doi.org/10.1016/j.jveb.2013.04.068.
Skinner, B.F., 1938. The Behavior of Organisms: an Experimental Analysis. Copley Publishing Group Acton, Massachusetts.
Spoolder, H.A.M., Burbidge, J.A., Edwards, S.A., Simmins, P.H., Lawrence, A.B., 1995. Provision of straw as a foraging substrate reduces the development of excessive chain and bar manipulation in food restricted sows. Appl. Anim. Behav. Sci. 43, 249–262. https://doi.org/10.1016/0168-1591(95)00566-B.
Sutton, R.S., Barto, A.G., 2018. Reinforcement Learning: an Introduction. MIT press.
Swaisgood, R., Shepherdson, D., 2006. Environmental enrichment as a strategy for mitigating stereotypies in zoo animals: a literature review and meta-analysis. In: Mason, G., Rushen, J. (Eds.), Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare, second ed. CABI. https://doi.org/10.1079/9780851990040.0256.
Wechsler, B., 1995. Coping and coping strategies: a behavioural view. Appl. Anim. Behav. Sci. 43 (2), 123–134. https://doi.org/10.1016/0168-1591(95)00557-9.
Whittaker, X., Spoolder, H.A., Edwards, S.A., Lawrence, A.B., Corning, S., 1998. The influence of dietary fibre and the provision of straw on the development of stereotypic behaviour in food restricted pregnant sows. Appl. Anim. Behav. Sci. 61 (2), 89–102. https://doi.org/10.1016/S0168-1591(98)00183-X.
Wiepkema, P.R., 1985. Abnormal behaviour in farm animals: ethological implications. Neth. J. Zool. 35, 279–299.
Williams, B.A., 1994. Conditioned reinforcement: experimental and theoretical issues. Behav. Anal. 17 (2), 261–285. https://doi.org/10.1007/BF03392675.
Würbel, H., 2006. The motivational basis of caged rodents' stereotypies. In: Mason, G., Rushen, J. (Eds.), Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare, second ed. CABI. https://doi.org/10.1079/9780851990040.0086.